Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: January 28, 2021

H. Bidgoli, Ed.
Nokia
V. Voyer
Bell Canada
P. Parekh
Cisco System
Z. Zhang
Juniper Networks
July 27, 2020

P2MP Policy Ping
draft-hb-pim-p2mp-policy-ping-00

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 January 28, 2021.

Table of Contents

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 PI to allow global optimization of the CP via Make Before Break procedures between the active PI and the newly setup and optimized PI.

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) from the root to a set of leaves.

This draft defines two new sub-TLVs to identifier a P2MP Policy and reuses concepts from [RFC6425]

2. Conventions used in this document

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 RFC 2119 [RFC2119].

3. Motivation

A P2MP Policy and its corresponding Replication Segments are usually setup via a PCE, after PCE has calculated the tree from the root to a set of 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 utilize to detect faults on the path of the tree and its associated replication segments. 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.

3.1. P2MP Policy Ping and Traceroute

The P2MP Policy Ping and Traceroute detect failure on the path of P2MP Tree and its corresponding replication segments. Ping/ Traceroute packets are forwarded on the P2MP Policy, on a specific CP or the CP's PI toward the leaf routers. They are replicated at the replication point based on the replication segment information on the corresponding transit node. The packet are processed accordingly when their TTL expires or when the egress router (leaf) is reached. The appropriate respond is send back to the root as per [RFC6425]

The implementation should take into account that there can be many CP under the P2MP Policy, one active CP and many inactive CPs. The implementation should allow each CP and its corresponding PIs to be tested via Ping and Trace route. The Ping and Traceroute packet is forwarded via that specific CP and its corresponding replication segments.

3.2. Packet format and new TLVs

The packet format used is as per [RFC4379] section 3.

3.2.1. Identifying a P2MP Policy

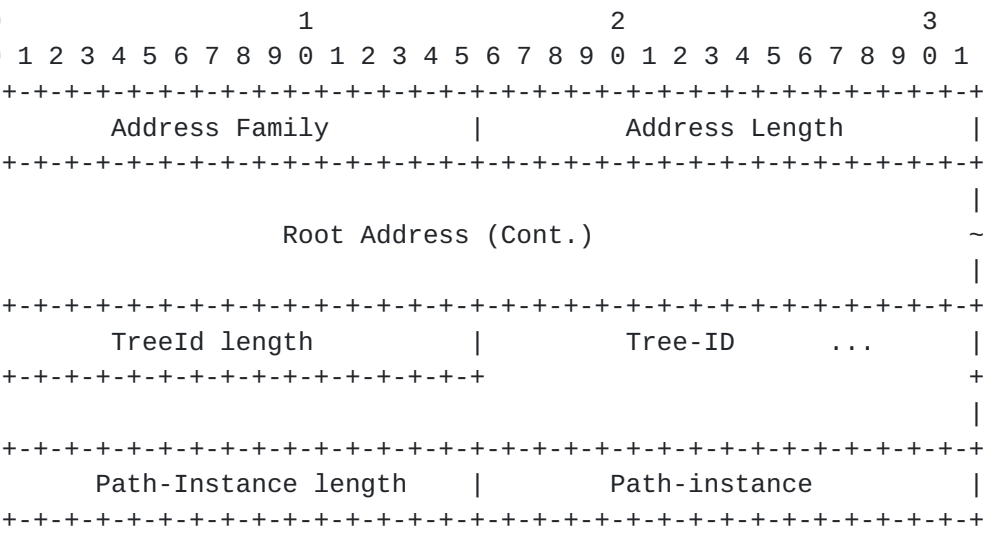
The new sub-TLV identifiers are assigned, as follow:

artwork

Sub-Type	Length	Value Field
38	20	P2MP Policy IPv4 CP
39	32	P2MP Policy IPv6 CP

3.2.1.1. P2MP Policy Sub-TLVs

artwork



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

- Address Length: Length of the Root LSR Address in octets.

- Root Addrt: 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. This should be set to 2.

o path-instance: path instance ID to be tested

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. Operation of P2MP Policy Ping

The P2MP Policy Ping should be able to test the path for all the CPs (active or inactive) or the CPs corresponding PIs for a P2MP policy on the root router. In addition it should test all the Out going interface for the replication segment connecting the root to the set of leaves for this CP or PI. To do so the concepts in [RFC4379] and [RFC6425] are extended to this draft.

The following sections will explain any new concepts for P2MP policy Ping

4.1. Replication Segments Connected via a Unicast SR domain

Two replication segment can be connected via a unicast SR domain. This means the traffic will be steered out of the upstream replication segment based on the unicast SR SID List (as an example node or adjacency SIDs). In this case the replication SID TTL is subtracted by one on the current node. If the TTL hasn't expired, the SR SID List will be placed on top of the replication segment, with replication segment being at the bottom of the stack. The SR SID List TTL should be set to 255, as an example pipe mode. When in SR domain the TTL of SR SIDs will be decremented accordingly. The TTL of the replication SID will be untouched in the SR domain. The replication SID TTL will be examined again on the next replication segment after the unicast SR SID List is removed from the label stack. The appropriate actions will be taken on this replication segment for P2MP Policy PING and Traceroute packet. To detect failures in SR domain is beyond the scope of this draft.

5. IANA Consideration

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

6. Security Considerations

TBD

7. Acknowledgments

8. References

8.1. 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>.

8.2. Informative References

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

[draft-voyer-spring-sr-replication-segment-04] "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
USA

Email: riparekh@cisco.com

Zhaohui Zhang
Juniper Networks
Boston
USA

Email: zzhang@juniper.com