Network Working Group
Internet-Draft
Intended status: Informational
Expires: May 2, 2020

M. Koldychev
S. Sivabalan
Cisco Systems, Inc.
T. Saad
V. Beeram
Juniper Networks, Inc.
H. Bidgoli
Nokia
B. Yadav
Ciena

October 30, 2019

PCEP Extensions for Signaling Multipath Information
draft-koldychev-pce-multipath-00

Abstract

This document introduces a mechanism to communicate multipath information in PCEP as a set of Explicit Route Objects (EROs). A special object is defined to carry per ERO attributes. This mechanism is applicable to SR-TE and RSVP-TE.

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 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 May 2, 2020.

Table of Contents

1. Introduction

Path Computation Element (PCE) Communication Protocol (PCEP) [RFC5440] enables the communication between a Path Computation Client (PCC) and a Path Control Element (PCE), or between two PCEs based on the PCE architecture [RFC4655].

PCEP Extensions for the Stateful PCE Model [RFC8231] describes a set of extensions to PCEP to enable active control of Multiprotocol Label Switching Traffic Engineering (MPLS-TE) and Generalized MPLS (GMPLS) tunnels. [RFC8281] describes the setup and teardown of PCE-initiated

LSPs under the active stateful PCE model, without the need for local configuration on the PCC, thus allowing for dynamic centralized control of a network.

PCEP Extensions for Segment Routing [I-D.ietf-pce-segment-routing] specifies extensions to the Path Computation Element Protocol (PCEP)

that allow a stateful PCE to compute and initiate Traffic Engineering (TE) paths, as well as a PCC to request a path subject to certain constraint(s) and optimization criteria in SR networks.

Segment Routing Policy for Traffic Engineering [I-D.ietf-spring-segment-routing-policy] details the concepts of SR Policy and approaches to steering traffic into an SR Policy. In particular, it describes the SR candidate-path as a collection of one or more Segment-Lists. The current PCEP standards only allow for signaling of one Segment-List per Candidate-Path. PCEP extension to support Segment Routing Policy Candidate Paths [I-D.barth-pce-segment-routing-policy-cp] specifically avoids defining how to signal multipath information, and states that this will be defined in another document (this one).

2. Terminology

In this document, the key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" are to be interpreted as described in BCP 14, RFC 2119 [RFC2119].

2.1. Terms and Abbreviations

The following terms are used in this document:

Endpoint:

The IPv4 or IPv6 endpoint address of the SR policy in question, as described in [I-D.ietf-spring-segment-routing-policy].

PCEP Tunnel:

The object identified by the PLSP-ID, as per [I-D.koldychev-pce-operational].

Tunnel Instance:

The object identified by the LSP-Identifiers TLV, as per [I-D.koldychev-pce-operational].

3. Motivation

This extension is motivated by the use-cases described below.

3.1. Signaling Multiple Segment-Lists of an SR Candidate-Path

The Candidate-Path of an SR Policy corresponds to a PCEP Tunnel, see [I-D.barth-pce-segment-routing-policy-cp]. Each Candidate-Path can contain multiple Segment-Lists and each Segment-List is encoded by one SR-ERO object. However, each Tunnel Instance can contain only a single ERO object, which prevents us from encoding multiple Segment-Lists within the same SR Candidate-Path.

With the help of the protocol extensions defined in this document, this limitation is overcome.

3.2. Splitting of Requested Bandwidth

A PCC may request a path with 100 Gbit of bandwidth, but all links in the network have only 50 Gbit capacity. The PCE can return two paths, that can each carry 50 Gbit. The PCC can then equally or unequally split the incoming 100 Gbit of traffic among the two 50 Gbit paths. Section 4.3 introduces a new TLV that carries the ERO path weight that allows distributing of incoming traffic on to the multiple ERO path(s).

3.3. Providing Backup ERO for Protection

It is desirable for the PCE to compute and signal to the PCC a backup ERO path that is used to protect a primary ERO path. In this case, an indication specify a primary or backup.

When multipath is used, a backup ERO path may protect one or more primary ERO path. For this reason, a primary and backup path identifiers are needed to indicate which backup ERO path(s) protect which primary ERO path(s). Section 4.4 introduces a new TLV that carries the required information.

4. Protocol Extensions

4.1. Multipath Capability TLV

We define the MULTIPATH-CAP TLV that MAY be present in the OPEN object and/or the LSP object. The purpose of this TLV is two-fold:

1. From PCC: it tells how many multipaths the PCC can install in forwarding.

2. From PCE: it tells that the PCE supports this standard and how many multipaths the PCE can compute.

Only the first instance of this TLV can be processed, subsequent instances SHOULD be ignored.

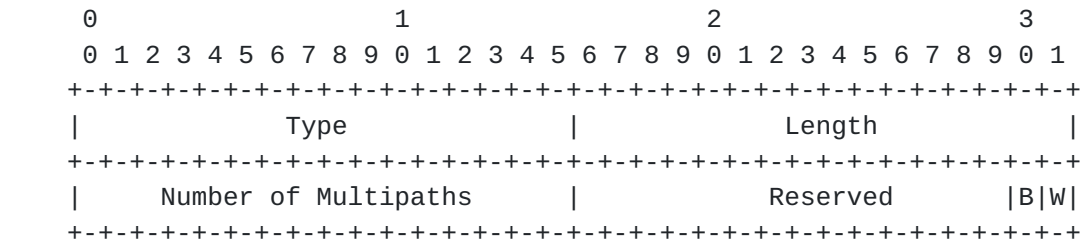


Figure 1: MULTIPATH-CAP TLV format

Type: TBD1 for "MULTIPATH-CAP" TLV.

Length: 4.

Number of Multipaths: the maximum number of multipaths that a PCE can return. The value 0 indicates unlimited number.

B-flag: whether MULTIPATH-BACKUP-TLV is supported.

W-flag: whether MULTIPATH-WEIGHT-TLV is supported.

Reserved: zero on transmit, ignore on receipt.

4.2. ERO Attributes Object

We define the ERO-ATTRIB object that is used to carry per-ERO information and to act as a separator between several ERO objects. The ERO-ATTRIB object always precedes the ERO that it applies to. If multiple ERO objects are present, then each ERO object MUST be preceded by an ERO-ATTRIB object that describes it.

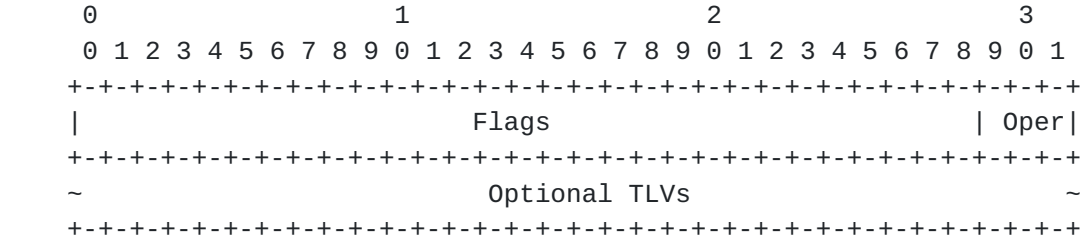


Figure 2: ERO-ATTRIB object format

Flags: to be extended in the future.

Oper: operational state of the ERO, same values as the identically named field in the LSP object.

4.3. Multipath Weight TLV

We define the MULTIPATH-WEIGHT TLV that MAY be present in the ERO_ATTRIBUTES object.

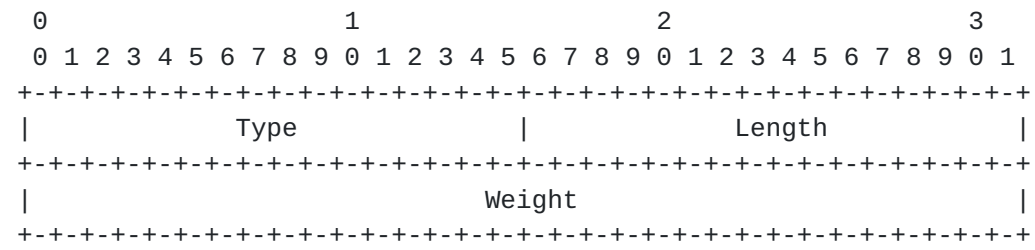


Figure 3: MULTIPATH-WEIGHT TLV format

Type: TBD2 for "MULTIPATH-WEIGHT" TLV.

Length: 4.

Weight: weight of this path within the multipath, if W-ECMP is desired. The fraction of flows a specific ERO carries is derived from the ratio of its weight to the sum of all other multipath ERO weights.

4.4. Multipath Backup TLV

This document introduces a new MULTIPATH-BACKUP TLV that is optional and can be present in the ERO_ATTRIBUTES object.

This TLV is used to indicate the presence of a backup ERO path that is used for protection in case of failure of the primary ERO path. The format of the MULTIPATH-BACKUP TLV is:

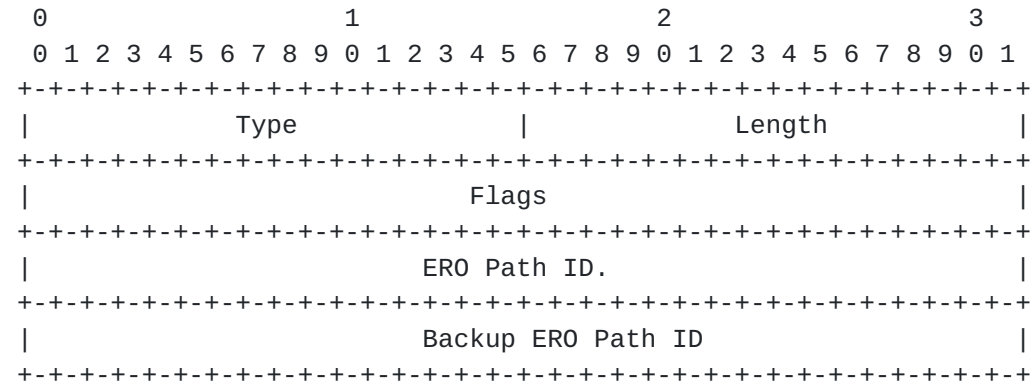
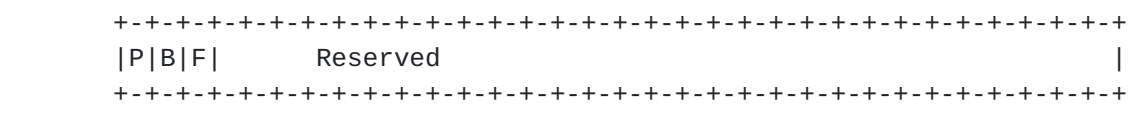


Figure 4: MULTIPATH-BACKUP TLV format

Type: TBD3 for "MULTIPATH-BACKUP" TLV

Length: 8

Flags:



P: If set, indicates the ERO is for a primary path

B: If set, indicates the ERO is for a backup path

F: If set, indicates this primary ERO is also protected. The backup ERO Path ID indicates the ERO of the backup path.

ERO Path ID: an identifier that identifies a primary path in the set of ERO(s)

Backup ERO Path ID: an identifier that identifies the backup path ERO in the set of ERO(s)

5. Operation

When the PCC wants to indicate to the PCE that it wants to get multipaths instead of a single path, it can do one or both of the following:

1. Send the MULTIPATH-CAP TLV in the OPEN object during session establishment. This applies to all PCEP Tunnels on the PCC, unless overridden by PCEP Tunnel specific information.

2. Send the MULTIPATH-CAP TLV in the LSP object for a particular PCEP Tunnel in the PCRpt message. This applies to the specified PCEP Tunnel and overrides the information from the OPEN object.

When PCE computes the path for a PCEP Tunnel, it MUST NOT return more multipaths than the corresponding value of "Number of Multipaths" from the MULTIPATH-CAP TLV. If this TLV is absent (from both OPEN and LSP objects), then the "Number of Multipaths" is assumed to be 1.

If the PCE supports this standard, then it MUST include the MULTIPATH-CAP TLV in the OPEN object. This tells the PCC that it can report multiple ERO objects to this PCE. If the PCE does not include the MULTIPATH-CAP TLV in the OPEN object, then the PCC MUST assume that the PCE does not support this standard and fall back to reporting only a single ERO.

6. IANA Considerations

IANA is requested to make the assignment of a new value for the existing "PCEP TLV Type Indicators" registry as follows:

TLV Type Value	TLV Name	Reference
TBD1	MULTIPATH-CAP	This document
TBD2	MULTIPATH-WEIGHT	This document
TBD3	MULTIPATH-BACKUP	This document

7. Security Considerations

None at this time.

8. Acknowledgement

Thanks to Dhruv Dhody for ideas and discussion.

9. References

9.1. Normative References

[I-D.barth-pce-segment-routing-policy-cp]
Koldychev, M., Sivabalan, S., Barth, C., and C. Li, "PCEP extension to support Segment Routing Policy Candidate Paths", draft-barth-pce-segment-routing-policy-cp-03 (work in progress), July 2019.

[I-D.ietf-pce-segment-routing]
Sivabalan, S., Filsfils, C., Tantsura, J., Henderickx, W., and J. Hardwick, "PCEP Extensions for Segment Routing", draft-ietf-pce-segment-routing-16 (work in progress), March 2019.

[I-D.ietf-spring-segment-routing-policy]
Filsfils, C., Sivabalan, S., daniel.voyer@bell.ca, d., bogdanov@google.com, b., and P. Mattes, "Segment Routing Policy Architecture", draft-ietf-spring-segment-routing-policy-03 (work in progress), May 2019.

[I-D.koldychev-pce-operational]
Koldychev, M., Sivabalan, S., Negi, M., Achaval, D., and H. Kotni, "PCEP Operational Clarification", draft-koldychev-pce-operational-00 (work in progress), July 2019.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC5440] Vasseur, JP., Ed. and JL. Le Roux, Ed., "Path Computation Element (PCE) Communication Protocol (PCEP)", RFC 5440, DOI 10.17487/RFC5440, March 2009, <https://www.rfc-editor.org/info/rfc5440>.

[RFC8231] Crabbe, E., Minei, I., Medved, J., and R. Varga, "Path Computation Element Communication Protocol (PCEP) Extensions for Stateful PCE", RFC 8231, DOI 10.17487/RFC8231, September 2017, <https://www.rfc-editor.org/info/rfc8231>.

[RFC8281] Crabbe, E., Minei, I., Sivabalan, S., and R. Varga, "Path Computation Element Communication Protocol (PCEP) Extensions for PCE-Initiated LSP Setup in a Stateful PCE Model", RFC 8281, DOI 10.17487/RFC8281, December 2017, <https://www.rfc-editor.org/info/rfc8281>.

9.2. Informative References

[RFC4655] Farrel, A., Vasseur, J., and J. Ash, "A Path Computation Element (PCE)-Based Architecture", RFC 4655, DOI 10.17487/RFC4655, August 2006, <https://www.rfc-editor.org/info/rfc4655>.

Authors' Addresses

Mike Koldychev
Cisco Systems, Inc.

Email: mkoldych@cisco.com

Siva Sivabalan
Cisco Systems, Inc.

Email: msiva@cisco.com

Tarek Saad
Juniper Networks, Inc.

Email: tsaad@juniper.net

Vishnu Pavan Beeram
Juniper Networks, Inc.

Email: vbeeram@juniper.net

Hooman Bidgoli
Nokia

Email: hooman.bidgoli@nokia.com

Bhupendra Yadav
Ciena

Email: byadav@ciena.com