PCE Working Group
Internet-Draft
Intended status: Standards Track
Expires: April 15, 2020

A. Raghuram
A. Goddard
AT&T
J. Karthik
S. Sivabalan
Cisco Systems, Inc.
M. Negi
Huawei Technologies
October 13, 2019

Ability for a Stateful Path Computation Element (PCE) to request and obtain control of a Label Switched Path (LSP) draft-ietf-pce-lsp-control-request-11

Abstract

A Stateful Path Computation Element (PCE) retains information about the placement of Multiprotocol Label Switching (MPLS) Traffic Engineering Label Switched Paths (TE LSPs). When a PCE has stateful control over LSPs it may send indications to LSP head-ends to modify the attributes (especially the paths) of the LSPs. A Path Computation Client (PCC) that has set up LSPs under local configuration may delegate control of those LSPs to a stateful PCE.

There are use-cases in which a stateful PCE may wish to obtain control of locally configured LSPs of which it is aware but that have not been delegated to the PCE.

This document describes an extension to the Path Computation Element communication Protocol (PCEP) to enable a PCE to make requests for such control.

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 April 15, 2020.

Copyright Notice

Copyright (c) 2019 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

Stateful Path Computation Element (PCE) communication Protocol (PCEP) extensions [RFC8231] specifies a set of extensions to PCEP [RFC5440] to enable stateful control of Traffic Engineering Label Switched

Paths (TE LSPs) between and across PCEP sessions in compliance with [RFC4657]. It includes mechanisms to synchronize LSP state between Path Computation Clients (PCCs) and PCEs, delegate control of LSPs to PCE, and PCE-control of timing and sequence of path computations within and across PCEP sessions. The stateful PCEP defines the following two useful network operations:

- Delegation: As per [RFC8051], an operation to grant a PCE temporary rights to modify a subset of LSP parameters on one or more LSPs of a PCC. LSPs are delegated from a PCC to a PCE and are referred to as "delegated" LSPs.

- Revocation: As per [RFC8231], an operation performed by a PCC on a previously delegated LSP. Revocation revokes the rights granted to the PCE in the delegation operation.

For Redundant Stateful PCEs (section 5.7.4. of [RFC8231]), during a PCE failure, one of the redundant PCE might want to request to take control over an LSP. The redundant PCEs may use a local policy or a proprietary election mechanism to decide which PCE would take control. In this case, a mechanism is needed for a stateful PCE to request control of one or more LSPs from a PCC, so that a newly elected primary PCE can request to take over control.

In case of virtualized PCEs (vPCEs) running in virtual network function (VNF) mode, as the computation load in the network increases, a new instance of vPCE could be instantiated to balance the current load. The PCEs could use a proprietary algorithm to decide which LSPs to be assigned to the new vPCE. Thus, having a mechanism for the PCE to request control of some LSPs is needed.

In some deployments, the operator would like to use stateful PCE for global optimization algorithms but would still like to keep the control of the LSP at the PCC. In such cases, a stateful PCE could request to take control during the global optimization and return the delegation once done.

Note that [RFC8231] specifies a mechanism for a PCC to delegate an orphaned LSP to another PCE. The mechanism defined in this document can be used in conjunction to [RFC8231]. Ultimately, it is the PCC that decides which PCE to delegate the orphaned LSP to.

This specification provides a simple extension: by using it a PCE can request control of one or more LSPs from any PCC over the stateful PCEP session. The procedures for granting and relinquishing control of the LSPs are specified in accordance with the specification [RFC8231] unless explicitly set aside in this document.

2. Terminology

This document uses the following terms defined in [RFC5440]:

PCC: Path Computation Client.

PCE: Path Computation Element.

PCEP: Path Computation Element communication Protocol.

This document uses the following terms defined in [RFC8231]:

PCRpt: Path Computation State Report message.

PCUpd: Path Computation Update Request message.

PLSP-ID: A PCEP-specific identifier for the LSP.

SRP: Stateful PCE Request Parameters.

Readers of this document are expected to have some familiarity with [RFC8231].

2.1. Requirements Language

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. LSP Control Request Flag

The Stateful PCE Request Parameters (SRP) object is defined in Section 7.2 of [RFC8231] and it includes a Flags field.

A new flag, the "LSP-Control Request Flag" (C) - TBD, is introduced in the SRP object. On a PCUpd message, a PCE sets the C Flag to 1 to indicate that it wishes to gain control of LSPs. The LSPs are identified by the PLSP-ID in the LSP object following the SRP object. A PLSP-ID of value other than 0 and 0xFFFFF is used to identify the LSP for which the PCE requests control. The PLSP-ID value of 0 indicates that the PCE is requesting control of all LSPs originating from the PCC that it wishes to delegate. The C Flag has no meaning in other PCEP messages that carry SRP objects and for which the C flag MUST be set to 0 on transmission and MUST be ignored on receipt.

4. Operation

During normal operation, a PCC that wishes to delegate the control of an LSP sets the D Flag (delegate, Section 7.3 of [RFC8231]) to 1 in all PCRpt messages pertaining to the LSP. The PCE confirms the delegation by setting D Flag to 1 in all PCUpd messages pertaining to the LSP. The PCC revokes the control of the LSP from the PCE by setting D Flag to 0 in PCRpt messages pertaining to the LSP. If the PCE wishes to relinquish the control of the LSP, it sets D Flag to 0 in all PCUpd messages pertaining to the LSP.

If a PCE wishes to gain control over an LSP, it sends a PCUpd message with C Flag set to 1 in SRP object. The LSP for which the PCE requests control is identified by the PLSP-ID in the associated LSP object. The PLSP-ID of 0 indicates that the PCE wants control over all LSPs originating from the PCC. An implementation of this feature needs to make sure to check for the LSP control feature (C flag set to 1) before any check for PLSP-ID (as prescribed in [RFC8231]). The D Flag and C Flag are mutually exclusive in a PCUpd message. The PCE MUST NOT send a control request for the LSP which is already delegated to the PCE, i.e. if the D Flag is set in the PCUpd message, then the C Flag MUST NOT be set. If a PCC receives a PCUpd message with D Flag set in the LSP object (i.e. LSP is already delegated) and the C Flag is also set (i.e. PCE is making a control request), the PCC MUST ignore the C Flag. A PCC can decide to delegate the control of the LSP at its own discretion. If the PCC grants or denies the control, it sends a PCRpt message with D Flag set to 1 and 0 respectively in accordance with stateful PCEP [RFC8231]. If the PCC does not grant the control, it MAY choose to not respond, and the PCE MAY choose to retry requesting the control preferably using exponentially increasing timer. Note that, if the PCUpd message with C Flag set is received for a currently non-delegated LSP (for which the PCE is requesting delegation), this MUST NOT trigger the error handling as specified in [RFC8231] (a PCErr with Error-type=19 (Invalid Operation) and error-value 1 (Attempted LSP Update Request for a non-delegated LSP)).

As per [RFC8231], a PCC cannot delegate an LSP to more than one PCE at any time. If a PCE requests control of an LSP that has already been delegated by the PCC to another PCE, the PCC MAY ignore the request, or MAY revoke the delegation to the first PCE before delegating it to the second. This choice is a matter of local policy.

It should be noted that a legacy implementation of PCC that does not support this extension would receive an LSP control request: PCUpd message with C flag set and D flag (delegate) unset, it would ignore the C flag and trigger the error condition for the D flag as

specified in [RFC8231] (a PCErr with Error-type=19 (Invalid Operation) and error-value 1 (Attempted LSP Update Request for a non-delegated LSP)). Further, in case of PLSP-ID of 0, the error condition as specified in [RFC8231] (a PCErr with Error-type=19 (Invalid Operation) and error-value 3 (Attempted LSP Update Request for an LSP identified by an unknown PSP-ID)) would be triggered.

[RFC8281] describes the setup, maintenance and teardown of PCE-initiated LSPs under the stateful PCE model. It also specifies how a PCE may obtain control over an orphaned LSP that was PCE-initiated. A PCE implementation can apply the mechanism described in this document in conjunction with those in [RFC8281].

5. Implementation Status

[Note to the RFC Editor - remove this section before publication, as well as remove the reference to RFC 7942.]

This section records the status of known implementations of the protocol defined by this specification at the time of posting of this Internet-Draft, and is based on a proposal described in [RFC7942]. The description of implementations in this section is intended to assist the IETF in its decision processes in progressing drafts to RFCs. Please note that the listing of any individual implementation here does not imply endorsement by the IETF. Furthermore, no effort has been spent to verify the information presented here that was supplied by IETF contributors. This is not intended as, and must not be construed to be, a catalog of available implementations or their features. Readers are advised to note that other implementations may exist.

According to [RFC7942], "this will allow reviewers and working groups to assign due consideration to documents that have the benefit of running code, which may serve as evidence of valuable experimentation and feedback that have made the implemented protocols more mature. It is up to the individual working groups to use this information as they see fit".

5.1. Huawei's Proof of Concept based on ONOS

The PCE function was developed in the ONOS open source platform. This extension was implemented on a private version as a proof of concept to enable multi-instance support.

- Organization: Huawei

- Implementation: Huawei's PoC based on ONOS

- Description: PCEP as a southbound plugin was added to ONOS. To support multi-instance ONOS deployment in a cluster, this extension in PCEP is used. Refer https://wiki.onosproject.org/display/ONOS/PCEP+Protocol

- Maturity Level: Prototype

- Coverage: Full

- Contact: satishk@huawei.com

6. Security Considerations

The security considerations listed in [RFC8231] and [RFC8281] apply to this document as well. However, this document also introduces a new attack vector. An attacker may flood the PCC with request to delegate all of its LSPs at a rate which exceeds the PCC's ability to process them, either by spoofing messages or by compromising the PCE itself. The PCC SHOULD be configured with a threshold rate for the delegation requests received from the PCE. If the threshold is reached, it is RECOMMENDED to log the issue.

A PCC is the ultimate arbiter of delegation. As per [RFC8231], a local policy at PCC is used to influence the delegation. A PCC can also revoke the delegation at any time. A PCC need not blindly trust the control requests and SHOULD take local policy and other factors into consideration before honoring the request.

Note that, a PCE may not be sure if a PCC supports this feature. A PCE would try sending a control request to a 'legacy' PCC, which would in turn respond with an error as described in Section 4. So a PCE would learn this fact only when it wants to take control over an LSP. A PCE might also be susceptible to a downgrade attacks by falsifying the error condition.

As per [RFC8231], it is RECOMMENDED that these PCEP extensions only be activated on authenticated and encrypted sessions across PCEs and PCCs belonging to the same administrative authority, using Transport Layer Security (TLS) [RFC8253], as per the recommendations and best current practices in BCP 195 [RFC7525] (unless explicitly excluded in [RFC8253]).

7. IANA Considerations

7.1. SRP Object Flags

IANA maintains a registry called the "Path Computation Element Protocol (PCEP) Numbers" registry. It contains a subregistry called the "SRP Object Flag Field" registry. This document requests IANA to allocate following code point in the "SRP Object Flag Field" subregistry.

Bit	Description	Reference
TBD	LSP-Control Request Flag	This document

8. Manageability Considerations

All manageability requirements and considerations listed in [RFC5440] and [RFC8231] apply to PCEP protocol extensions defined in this document. In addition, requirements and considerations listed in this section apply.

8.1. Control of Function and Policy

A PCC implementation SHOULD allow the operator to configure the policy based on which it honors the request to control the LSPs. This includes the handling of the case where an LSP control request is received for an LSP that is currently delegated to some other PCE. A PCC implementation SHOULD also allow the operator to configure the threshold rate based on which it accepts the delegation requests from the PCE. Further, the operator MAY be allowed to trigger the LSP control request for a particular LSP at the PCE. A PCE implementation SHOULD also allow the operator to configure an exponentially increasing timer to retry the control requests for which the PCE did not get a response.

8.2. Information and Data Models

The PCEP YANG module [I-D.ietf-pce-pcep-yang] could be extended to include mechanism to trigger the LSP control request.

8.3. Liveness Detection and Monitoring

Mechanisms defined in this document do not imply any new liveness detection and monitoring requirements in addition to those already listed in [RFC5440].

8.4. Verify Correct Operations

Mechanisms defined in this document do not imply any new operation verification requirements in addition to those already listed in [RFC5440] and [RFC8231].

8.5. Requirements On Other Protocols

Mechanisms defined in this document do not imply any new requirements on other protocols.

8.6. Impact On Network Operations

Mechanisms defined in [RFC5440] and [RFC8231] also apply to PCEP extensions defined in this document. Further, the mechanism described in this document can help the operator to request control of the LSPs at a particular PCE.

9. Acknowledgements

Thanks to Jonathan Hardwick to remind the authors to not use suggested values in IANA section.

Thanks to Adrian Farrel, Haomian Zheng and Tomonori Takeda for their valuable comments.

Thanks to Shawn M. Emery for security directorate's review.

Thanks to Francesca Palombini for GENART review.

Thanks to Benjamin Kaduk, Martin Vigoureux, Alvaro Retana, and Barry Leiba for IESG reviews.

10. References

10.1. 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>.

[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>.

[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>.

[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>.

10.2. Informative References

[RFC4657] Ash, J., Ed. and J. Le Roux, Ed., "Path Computation Element (PCE) Communication Protocol Generic Requirements", RFC 4657, DOI 10.17487/RFC4657, September 2006, <https://www.rfc-editor.org/info/rfc4657>.

[RFC7525] Sheffer, Y., Holz, R., and P. Saint-Andre, "Recommendations for Secure Use of Transport Layer Security (TLS) and Datagram Transport Layer Security (DTLS)", BCP 195, RFC 7525, DOI 10.17487/RFC7525, May 2015, <https://www.rfc-editor.org/info/rfc7525>.

[RFC7942] Sheffer, Y. and A. Farrel, "Improving Awareness of Running Code: The Implementation Status Section", BCP 205, RFC 7942, DOI 10.17487/RFC7942, July 2016, <https://www.rfc-editor.org/info/rfc7942>.

[RFC8051] Zhang, X., Ed. and I. Minei, Ed., "Applicability of a Stateful Path Computation Element (PCE)", RFC 8051, DOI 10.17487/RFC8051, January 2017, <https://www.rfc-editor.org/info/rfc8051>.

[RFC8253] Lopez, D., Gonzalez de Dios, O., Wu, Q., and D. Dhody, "PCEPS: Usage of TLS to Provide a Secure Transport for the Path Computation Element Communication Protocol (PCEP)", RFC 8253, DOI 10.17487/RFC8253, October 2017, <https://www.rfc-editor.org/info/rfc8253>.

[I-D.ietf-pce-pcep-yang] Dhody, D., Hardwick, J., Beeram, V., and J. Tantsura, "A YANG Data Model for Path Computation Element Communications Protocol (PCEP)", draft-ietf-pce-pcep-yang-12 (work in progress), July 2019.

Appendix A. Contributor Addresses

Dhruv Dhody
Huawei Technologies
Divyashree Techno Park, Whitefield
Bangalore, Karnataka 560066
India

EMail: dhruv.ietf@gmail.com

Jon Parker
Cisco Systems, Inc.
2000 Innovation Drive
Kanata, Ontario K2K 3E8
Canada

EMail: jdparker@cisco.com

Chaitanya Yadlapalli
AT&T
200 S Laurel Aevenue
Middletown NJ 07748
USA

EMail: cy098d@att.com

Authors' Addresses

Aswatnarayan Raghuram
AT&T
200 S Laurel Aevenue
Middletown, NJ 07748
USA

EMail: ar2521@att.com

Al Goddard
AT&T
200 S Laurel Aevenue
Middletown, NJ 07748
USA

EMail: ag6941@att.com

Jay Karthik
Cisco Systems, Inc.
125 High Street
Boston, Massachusetts 02110
USA

EMail: jakarthi@cisco.com

Siva Sivabalan
Cisco Systems, Inc.
2000 Innovation Drive
Kanata, Ontario K2K 3E8
Canada

EMail: msiva@cisco.com

Mahendra Singh Negi
Huawei Technologies
Divyashree Techno Park, Whitefield
Bangalore, Karnataka 560066
India

EMail: mahend.ietf@gmail.com