

PCE Working Group
Internet-Draft
Intended status: Informational
Expires: August 23, 2019

C. Li
X. Zhang
Huawei Technologies
February 19, 2019

Stateful Path Computation Element (PCE) Inter-domain Considerations draft-dhody-pce-stateful-pce-interdomain-08

Abstract

A stateful Path Computation Element (PCE) maintains information about Label Switched Path (LSP) characteristics and resource usage within a network in order to provide traffic engineering path calculations for its associated Path Computation Clients (PCCs). Furthermore, PCEs are used to compute shortest constrained Traffic Engineering Label Switched Paths (TE LSPs) in Multiprotocol Label Switching (MPLS) and Generalized MPLS (GMPLS) networks across multiple domains.

This document describes general considerations for the deployment of stateful PCE(s) in inter-domain scenarios including inter-area and inter-AS. The inter-layer considerations will be described in a separate document. This document does not specify any extensions to PCEP.

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 August 23, 2019.

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	2
2.	Overview	3
2.1.	LSP State Synchronization	4
3.	Stateful PCE Deployments	4
3.1.	Single Stateful PCE, Multiple Domains	5
3.2.	Multiple Stateful PCE, Multiple Domains	5
3.2.1.	Per Domain Path Computation	6
3.2.2.	Backward-Recursive PCE-based Computation	7
3.2.2.1.	Delegation	8
3.2.2.2.	PCE-initiated LSP	8
3.2.2.3.	LSP Stitching	8
3.2.3.	Hierarchical PCE	8
4.	Interworking between different signalling types	9
5.	Security Considerations	9
6.	Manageability Considerations	10
6.1.	Control of Function and Policy	10
6.2.	Information and Data Models	10
6.3.	Liveness Detection and Monitoring	10
6.4.	Verify Correct Operations	10
6.5.	Requirements On Other Protocols	10
6.6.	Impact On Network Operations	10
7.	IANA Considerations	10
8.	Acknowledgments	10
9.	References	11
9.1.	Normative References	11
9.2.	Informative References	11
Appendix A.	Contributor Addresses	14
	Authors' Addresses	14

[1.](#) Introduction

The Path Computation Element communication Protocol (PCEP) provides mechanisms for Path Computation Elements (PCEs) to perform path computations in response to Path Computation Clients' (PCCs) requests.

[RFC8051] describes general considerations for a stateful PCE deployment and examines its applicability and benefits, as well as its challenges and limitations through a number of use cases. [RFC8231] describes a set of extensions to PCEP to provide stateful control. A stateful PCE has access to not only the information carried by the network's Interior Gateway Protocol (IGP), but also the set of active paths and their reserved resources for its computations. The additional state allows the PCE to compute constrained paths while considering individual LSPs and their interactions.

The ability to compute shortest constrained TE LSPs in Multiprotocol Label Switching (MPLS) and Generalized MPLS (GMPLS) networks across multiple domains has been identified as a key motivation for PCE development. In this context, a domain is a collection of network elements within a common sphere of address management or path computational responsibility such as an Interior Gateway Protocol (IGP) area or an Autonomous Systems (AS).

This document presents general considerations for stateful PCE(s) deployment in multi-domain scenarios. Further, [I-D.ietf-pce-inter-area-as-applicability] examines the applicability of the PCE architecture, protocols, and protocol extensions for computing multi-area and multi-AS paths in MPLS and GMPLS networks with focus on the stateless PCE deployments.

2. Overview

A stateful PCE maintains two sets of information for use in path computation. The first is the Traffic Engineering Database (TED) which includes the topology and resource state in the network. The second is the LSP State Database (LSP-DB), in which a PCE stores attributes of all active LSPs in the network, such as their paths through the network, bandwidth/resource usage, switching types and LSP constraints. This state information allows the PCE to compute constrained paths while considering individual LSPs and their inter-dependency. [RFC8231] applies equally to MPLS-TE and GMPLS LSPs and distinguishes between an active and a passive stateful PCE. A passive stateful PCE uses LSP state information to optimize path computations but does not actively update LSP state. In contrast, an active stateful PCE may issue recommendations to the network. For example, an active stateful PCE may update LSP parameters for those LSPs that have been delegated control over to the PCE by its PCCs.

The capability to compute the routes of end-to-end inter-domain MPLS-TE LSPs is expressed as requirements in [RFC4105] and [RFC4216] and may be realized by PCE(s). PCEs may use one of the following mechanisms to compute end-to-end paths:

- o a per-domain path computation technique [[RFC5152](#)];
- o a Backward-Recursive PCE-based Computation (BRPC) mechanism [[RFC5441](#)];
- o a Hierarchical PCE mechanism [[RFC6805](#)];

This document examines the stateful PCE inter-domain considerations for all of these mechanisms.

[2.1.](#) LSP State Synchronization

The population of the LSP-DB using information received from PCCs (ingress LSR) is supported by the stateful PCE extensions defined in [[RFC8231](#)] , i.e., via LSP state report messages.

The inter-domain LSP state is synchronized to the ingress-PCE from the ingress LSR (PCC), but this PCC cannot synchronize to other PCEs (in transit or egress domains), thus other mechanism must be investigated for this purpose.

Either the boundary node of the other domains, would need to synchronize the state of LSP passing through it to the PCE, or a mechanism for synchronization of inter-domain LSPs between the PCEs is required. The former would require small change in the existing state synchronization and reporting where a border node acts as a PCC. The latter could use the mechanism described in [[I-D.litkowski-pce-state-sync](#)] can be used between the PCEs to synchronize the inter-domain LSP state between each other. Further sections provide various considerations for this choice.

[3.](#) Stateful PCE Deployments

There are multiple models to perform PCE-based inter-domain path computation:

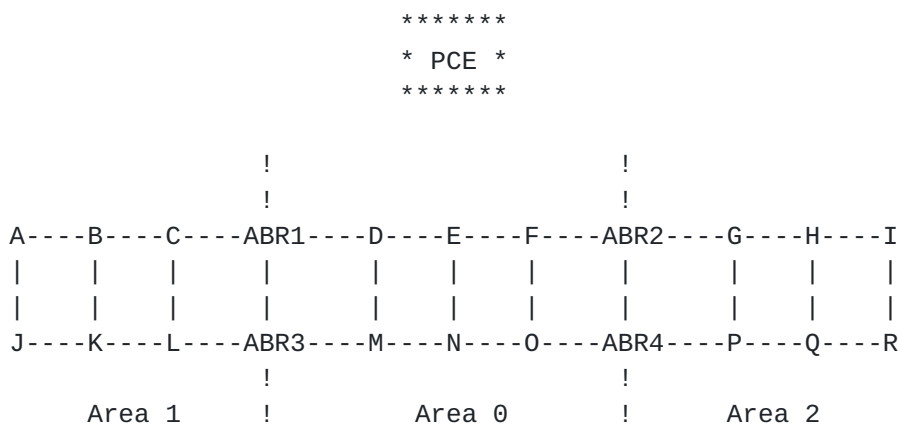
- o A single PCE
- o Multiple PCEs
 - * without inter-PCE communication
 - * with inter-PCE communication

This section describes stateful PCE considerations for each of these deployment models.

3.1. Single Stateful PCE, Multiple Domains

In this model, inter-domain path computation is performed by a single stateful PCE that has topology visibility into all domains. The inter-domain LSP state is synchronized to the PCE from the ingress LSR (PCC) itself. The PCC may also choose to delegate control over this LSP to the PCE. Thus this model is similar to a single domain in all aspects.

Following figure show an example of inter-area case comprising of Area 0,1 and 2. A single stateful PCE is deployed for all areas.



In this model, PCE has visibility into the topology (TED) of all domains as well as the state of all active LSPs (LSP-DB) including inter-domain LSPs. This model is thus well suited to take advantage of all stateful PCE capabilities.

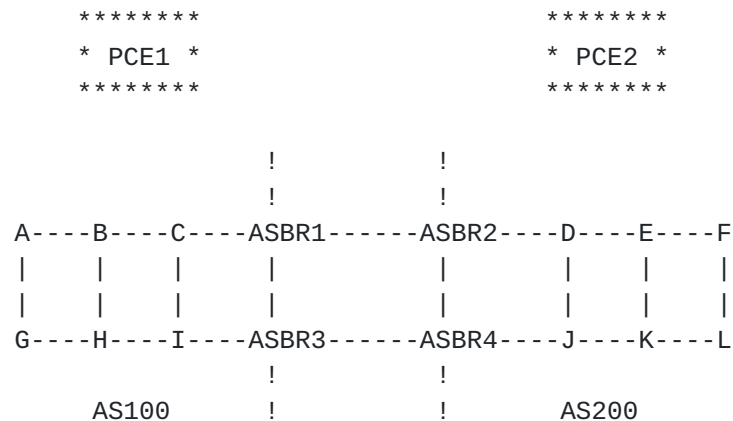
It should be noted that in some deployments, a single stateful PCE may not be possible because of administrative and confidentiality concerns.

3.2. Multiple Stateful PCE, Multiple Domains

In this model, there is at least one PCE per domain, and each PCE has topology (TED) visibility restricted to its own domain. The inter-domain LSP state is synchronized to the ingress-PCE from the ingress LSR (PCC), but this PCC may not be able to synchronize to other PCEs (in transit or egress domains). This PCC may also choose to delegate control over this LSP to the Ingress-PCE, which may issue inter-domain path computation or re-optimization request to other PCEs. An inter-domain LSP that originates in a domain, is synchronized to the PCE in that domain. A new procedure is needed to synchronize state of inter-domain LSP that do not originate in the domain. In other words, inter-domain LSP state should also be synchronized to transit

and egress PCEs as the inter-domain LSP traverse through those domains.

Following figure show an example of inter-AS case comprising of AS 100 and AS 200. A stateful PCE is deployed per AS.



In order to conceal the information, a PCE may use path-key based confidentiality mechanisms as per [\[RFC5520\]](#).

This section further describes considerations with respect to each of the inter-domain path computation techniques.

3.2.1. Per Domain Path Computation

The per domain path computation technique [\[RFC5152\]](#) is based on Multiple PCE Path Computation without Inter-PCE Communication Model as described in [\[RFC4655\]](#). It defines a method where the path is computed during the signaling process (on a per-domain basis). The entry Boundary Node (BN) of each domain is responsible for performing the path computation for the section of the LSP that crosses the domain, or for requesting that a PCE for that domain computes that piece of the path.

The ingress LSR would synchronize the state to the ingress PCE, further the entry boundary nodes should also synchronize the state of inter-domain LSP to transit and egress PCEs. Note that the BN on the path of an LSP can probably see the path (through the Record Route object in RSVP-TE signaling [\[RFC3209\]](#)) and knows the bandwidth reserved for the LSP. Thus each entry BN along the path could be made responsible to synchronize the LSP state to the transit/egress PCE(s).

Since the stateful PCE(s) do not communicate during this inter-domain path computation technique and each entry BN would perform path

computation via Path Computation Request (PCReq) and Reply (PCRep) messages, a passive stateful PCE is well suited for this case.

In case of delegation to the ingress PCE (active stateful PCE), it would be capable of loose path computation only and make updates to the ingress LSR with this limited visibility. The entry BN would perform path computation via Path Computation Request and Reply messages (and thus rely on the passive stateful mode). Thus the inter-domain LSP is delegated only to the ingress PCE.

3.2.2. Backward-Recursive PCE-based Computation

The BRPC [[RFC5441](#)] technique is based on Multiple PCE Path Computation with Inter-PCE Communication Model as described in [[RFC4655](#)]. It involves cooperation and communication between PCEs in order to compute an optimal end-to-end path across multiple domains. The sequence of domains to be traversed may be known before the path computation, but it can also be used when the domain path is unknown and determined during path computation.

As described in [Section 3.2.1](#), the entry boundary nodes may synchronize the state of inter-domain LSPs to transit and egress PCEs. An alternative approach may be for each PCE to synchronize the state along the path across domains, i.e., each PCE would report the state to the next PCE(s) in the adjacent domain along the domain sequence of the inter-domain path. A mechanism similar to state-sync described in [[I-D.litkowski-pce-state-sync](#)] may be utilized for this purpose.

Some path segment in the end to end path may also be hidden via path-key as per [[RFC5520](#)] during state synchronization.

In case of passive path computation request to the ingress PCE from the ingress LSR the BRPC path computation procedure is applied to compute end-to-end path by using PCReq and PCRep messages among stateful PCE(s) in passive mode.

In case of delegation to the ingress PCE (active stateful PCE), the ingress PCE may trigger the end-to-end path computation via the same BRPC procedure using the path computation request and reply messages among stateful PCE(s) (acting in passive mode). For re-optimization the ingress PCE still rely on the same BRPC procedure triggered by the ingress PCE. Ultimately the inter-domain LSP is delegated to the ingress PCE and only the ingress PCE can trigger end-to-end (E2E) path re-optimization with help of transit/egress PCE using the BRPC procedure, based on the result the ingress PCE would issue updates to the inter-domain LSP.

3.2.2.1. Delegation

As noted in this document, the inter-domain LSP is delegated to the ingress PCE and only the ingress PCE can issue updates to the inter-domain LSP. The ingress PCE is responsible to trigger E2E path re-optimization.

Thus the ingress PCE can recommend updation for all aspects of the inter-domain LSP including the segment of path in another domain (which it may have computed with the help of other cooperating PCEs). These interaction between PCEs for the inter-domain path computation are done using PCReq/PCRep messages (i.e., in a passive mode).

The transit/egress PCE cannot update any attribute of the inter-domain LSP on its own as it may not have any interaction with the ingress LSR. A mechanism may be developed for transit/egress PCE to inform the ingress PCE to trigger E2E re-optimization and choose to update the inter-domain LSP based on the result. Also the ingress PCE may use combination of local information and events along with some external mechanism (management / monitoring interface) to trigger E2E path re-optimization.

Though ingress PCE can recommend update for path segments in other domains, the entry boundary node of that domain can apply policy control during signaling as explained in [[RFC4105](#)] and [[RFC4216](#)].

3.2.2.2. PCE-initiated LSP

[RFC8281] describes setup, maintenance and teardown of PCE-initiated LSPs under the stateful PCE model, without the need for local configuration on the PCC. Similar to LSP updation, the inter-domain LSP can be initiated by the ingress PCE using the PCInitiate message to the ingress LSR. Note that per-domain LSP may also be initiated by respective domain's PCE and stitched together.

3.2.2.3. LSP Stitching

[I-D.dugeon-pce-stateful-interdomain] describes a proposal to combine a Backward Recursive method with PCInitiate message to setup independent paths per domain, and combine these different paths together in order to operated them as end-to-end inter-domain paths without the need of signaling session between AS border routers.

3.2.3. Hierarchical PCE

In H-PCE [[RFC6805](#)] architecture, the parent PCE is used to compute a multi-domain path based on the domain connectivity information. The

parent PCE may be requested to provide a end-to-end path or only the sequence of domains.

As described in [Section 3.2.1](#) and [Section 3.2.2](#), the entry boundary nodes may synchronize the state of inter-domain LSP to transit and egress child PCEs. In this case, it might not be possible to synchronize state to the parent PCE. If the parent PCE provides the sequence of domains and BRPC procedure is used to get the E2E path, each PCE may be responsible to synchronize the state along the path across domains similar to [Section 3.2.2](#). An alternative approach may be for ingress PCE to synchronize LSP state with the Parent PCE and it may further synchronize the state to the child PCE(s) along the path across domains, i.e. parent PCE would report the state to the child PCE(s) along the domain sequence.

Some path segment in the end to end path may also be hidden via path-key as per [\[RFC5520\]](#) during state synchronization.

In case of passive path computation request to the ingress PCE from the ingress LSR, the H-PCE path computation procedure is applied to compute sequence of domains or end-to-end path by using PCReq and PCRep messages among stateful PCE(s) in passive mode.

In case of delegation to the ingress PCE (active stateful PCE), the ingress child PCE may further delegate to parent PCE as per [\[I-D.ietf-pce-stateful-hpce\]](#). The parent PCE could update the path of the inter-domain LSP. Both per-domain stitched LSP as well as E2E contiguous LSP are possible. Further parent PCE could also initiate the creation of LSP for both per-domain stitched LSP to all child PCE or E2E contiguous LSP to ingress child PCE as described in [\[I-D.ietf-pce-stateful-hpce\]](#).

4. Interworking between different signalling types

Apart from the RSVP-TE signaling protocol, other TE path setup methods are possible within the PCE architecture, such as Segment Routing (SR) [\[I-D.ietf-pce-segment-routing\]](#) and PCECC [\[I-D.ietf-pce-pcep-extension-for-pce-controller\]](#). There is a possibility of where two domains may use different setup technique and coordination would be needed for inter-working. PCE can play an important in stitching per-domain heterogeneous LSPs.

5. Security Considerations

The security considerations are as per [\[RFC5440\]](#) and [\[RFC8231\]](#). Any multi-domain operation necessarily involves the exchange of information across domain boundaries. This may represent a significant security and confidentiality risk especially when the

domains are controlled by different commercial entities. PCEP allows individual PCEs to maintain confidentiality of their domain path information by using path-keys [[RFC5520](#)].

6. Manageability Considerations

6.1. Control of Function and Policy

Mechanisms defined in this document do not imply any new control of function and policy requirements.

6.2. Information and Data Models

[RFC7420] describes the PCEP MIB, there are no new MIB Objects for this document.

6.3. 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](#)].

6.4. Verify Correct Operations

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

6.5. Requirements On Other Protocols

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

6.6. Impact On Network Operations

Mechanisms defined in this document do not have any impact on network operations in addition to those already listed in [[RFC5440](#)].

7. IANA Considerations

This is an informational document and has no IANA considerations.

8. Acknowledgments

The authors would like to thank Young Lee, Haomian Zheng, Fatai Zhang for their comments.

9. References

9.1. Normative References

- [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>>.

9.2. Informative References

- [RFC3209] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V., and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", [RFC 3209](#), DOI 10.17487/RFC3209, December 2001, <<https://www.rfc-editor.org/info/rfc3209>>.
- [RFC4105] Le Roux, J., Ed., Vasseur, J., Ed., and J. Boyle, Ed., "Requirements for Inter-Area MPLS Traffic Engineering", [RFC 4105](#), DOI 10.17487/RFC4105, June 2005, <<https://www.rfc-editor.org/info/rfc4105>>.
- [RFC4216] Zhang, R., Ed. and J. Vasseur, Ed., "MPLS Inter-Autonomous System (AS) Traffic Engineering (TE) Requirements", [RFC 4216](#), DOI 10.17487/RFC4216, November 2005, <<https://www.rfc-editor.org/info/rfc4216>>.
- [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>>.
- [RFC5152] Vasseur, JP., Ed., Ayyangar, A., Ed., and R. Zhang, "A Per-Domain Path Computation Method for Establishing Inter-Domain Traffic Engineering (TE) Label Switched Paths (LSPs)", [RFC 5152](#), DOI 10.17487/RFC5152, February 2008, <<https://www.rfc-editor.org/info/rfc5152>>.

- [RFC5441] Vasseur, JP., Ed., Zhang, R., Bitar, N., and JL. Le Roux, "A Backward-Recursive PCE-Based Computation (BRPC) Procedure to Compute Shortest Constrained Inter-Domain Traffic Engineering Label Switched Paths", [RFC 5441](#), DOI 10.17487/RFC5441, April 2009, <<https://www.rfc-editor.org/info/rfc5441>>.
- [RFC5520] Bradford, R., Ed., Vasseur, JP., and A. Farrel, "Preserving Topology Confidentiality in Inter-Domain Path Computation Using a Path-Key-Based Mechanism", [RFC 5520](#), DOI 10.17487/RFC5520, April 2009, <<https://www.rfc-editor.org/info/rfc5520>>.
- [RFC6805] King, D., Ed. and A. Farrel, Ed., "The Application of the Path Computation Element Architecture to the Determination of a Sequence of Domains in MPLS and GMPLS", [RFC 6805](#), DOI 10.17487/RFC6805, November 2012, <<https://www.rfc-editor.org/info/rfc6805>>.
- [RFC7420] Koushik, A., Stephan, E., Zhao, Q., King, D., and J. Hardwick, "Path Computation Element Communication Protocol (PCEP) Management Information Base (MIB) Module", [RFC 7420](#), DOI 10.17487/RFC7420, December 2014, <<https://www.rfc-editor.org/info/rfc7420>>.
- [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>>.
- [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>>.
- [I-D.ietf-pce-inter-area-as-applicability]
King, D. and H. Zheng, "Applicability of the Path Computation Element to Inter-Area and Inter-AS MPLS and GMPLS Traffic Engineering", [draft-ietf-pce-inter-area-as-applicability-07](#) (work in progress), December 2018.
- [I-D.ietf-pce-stateful-hpce]
Dhody, D., Lee, Y., Ceccarelli, D., Shin, J., King, D., and O. Dios, "Hierarchical Stateful Path Computation Element (PCE).", [draft-ietf-pce-stateful-hpce-06](#) (work in progress), October 2018.

[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-15](#) (work in progress), February 2019.

[I-D.ietf-pce-pcep-extension-for-pce-controller]

Zhao, Q., Li, Z., Negi, M., and C. Zhou, "PCEP Procedures and Protocol Extensions for Using PCE as a Central Controller (PCECC) of LSPs", [draft-ietf-pce-pcep-extension-for-pce-controller-01](#) (work in progress), February 2019.

[I-D.litkowski-pce-state-sync]

Litkowski, S., Sivabalan, S., and D. Dhody, "Inter Stateful Path Computation Element communication procedures", [draft-litkowski-pce-state-sync-04](#) (work in progress), October 2018.

[I-D.dugeon-pce-stateful-interdomain]

Dugeon, O., Meuric, J., Lee, Y., Dhody, D., and D. Ceccarelli, "PCEP Extension for Stateful Inter-Domain Tunnels", [draft-dugeon-pce-stateful-interdomain-01](#) (work in progress), July 2018.

Appendix A. Contributor Addresses

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

EMail: dhruv.ietf@gmail.com

Udayasree Palle

EMail: udayasreereddy@gmail.com

Avantika

EMail: s.avantika.avantika@gmail.com

Authors' Addresses

Cheng Li
Huawei Technologies
Huawei Campus, No. 156 Beiqing Rd.
Beijing 100095
China

EMail: chengli13@huawei.com

Xian Zhang
Huawei Technologies
Bantian, Longgang District
Shenzhen 518129
P.R.China

EMail: zhang.xian@huawei.com

