Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: December 26, 2016

J. Dong
M. Chen
D. Dhody
Huawei Technologies
J. Tantsura
Individual
K. Kumaki
KDDI Corporation
T. Murai
Furukawa Network Solution Corp.

June 24, 2016

BGP Extensions for Path Computation Element (PCE) Discovery
draft-dong-pce-discovery-proto-bgp-05

Abstract

In networks where Path Computation Element (PCE) is used for centralized path computation, it is desirable for the Path Computation Clients (PCCs) to automatically discover a set of PCEs and select the suitable ones to establish the PCEP session. RFC 5088 and RFC 5089 define the PCE discovery mechanisms based on Interior Gateway Protocols (IGP). This document describes several scenarios in which the IGP based PCE discovery mechanisms cannot be used directly. In such scenarios, BGP might be suitable, thus this document specifies the BGP extensions for PCE discovery. The BGP based PCE discovery mechanism is complementary to the existing IGP based mechanisms.

Requirements Language

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].

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 December 26, 2016.

Table of Contents

1. Introduction

In network scenarios where Path Computation Element (PCE) is used for centralized path computation, it is desirable for the Path Computation Clients (PCCs) to automatically discover a set of PCEs and select the suitable ones to establish the PCEP session. [RFC5088] and [RFC5089] define the PCE discovery mechanisms based on Interior Gateway Protocols (IGP).

The IGP based discovery mechanism requires the PCE participate in the IGP network, which usually requires that the PCE is directly adjacent to at least one of the IGP routers in the network. In some scenarios

such requirement cannot be satisfied. For example, a PCE may need to provide path computation service to some subsidiary networks of an operator, which typically locate in different geographical region (and not IGP adjacent). Also when PCE function is implemented in a central server running IGP on individual interfaces to each IGP area can be cumbersome.

The requirement on PCE discovery, as described in [RFC4674], also include the automatic discovery of the PCEs in other domains, as it is a desirable function in the case of inter-domain path computation. The IGP based discovery mechanisms cannot meet such requirement.

For example, Backward Recursive Path Computation (BRPC) [RFC5441] can be used by cooperating PCEs to compute an inter-AS path, in which case these cooperating PCEs should be known to each other in advance. In this case the PCEs belongs to different AS and do not participate in a common IGP, the IGP based discovery mechanisms are not applicable.

Another example is the hierarchical PCE scenario [RFC6805], in which the child PCEs need to know the information of the parent PCEs. This cannot be achieved via IGP based discovery, as the child PCEs and the parent PCE are usually in different domains.

In some BGP IP-VPN networks, an end-to-end TE LSP between the CEs (Customer Edges) of a particular VPN is required [RFC5824]. In this case, CEs need the information of the PCE which can perform the CE to CE path computation for that VPN. Since the PCE may locate in a VPN site different from the site of the requesting CE, the IGP based discovery mechanism is not directly applicable, and some BGP based discovery mechanism is required to distribute the per-VPN PCE information to the VPN sites.

Since BGP has been extended for north-bound distribution of routing and Label Switched Path (LSP) information to PCE [RFC7752] [I-D.ietf-idr-te-lsp-distribution] and [I-D.ietf-idr-te-pm-bgp], PCEs can obtain the routing information without participating in IGP. In this scenario, a new BGP based PCE discovery mechanism is needed.

This document proposes to extend BGP for PCE discovery in the above scenarios. In networks where BGP-LS is used for the north-bound routing information distribution to PCE, the BGP based PCE discovery can make use of the existing BGP sessions and mechanisms to achieve automatic PCE discovery. Further IGP may be used to redistribute remote PCE information, the detailed mechanism is out of the scope of this document. Thus the BGP based PCE discovery is complementary to the existing IGP based mechanisms.

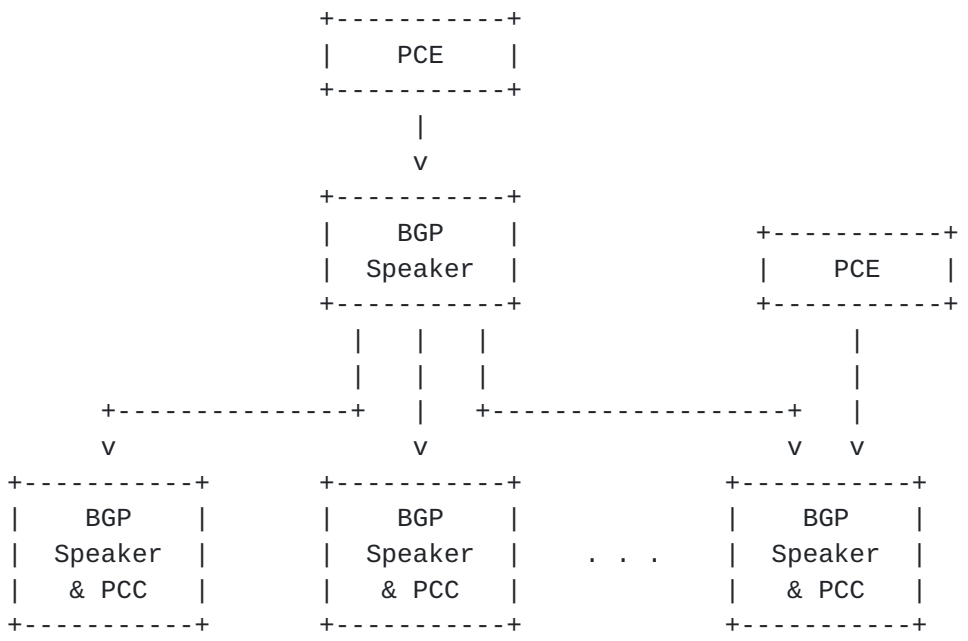

Figure 1: BGP for PCE discovery

As shown in the network architecture in Figure 1, BGP is used both for routing information distribution and for PCE information discovery. The routing information is collected from the network elements and distributed to PCE, while the PCE discovery information is advertised from PCE to PCCs, or among different PCEs. The PCCs maybe co-located with the BGP speakers as shown in Figure 1.

2. Carrying PCE Discovery Information in BGP

2.1. PCE Address Information

The PCE discovery information is advertised in BGP UPDATE messages using the MP_REACH_NLRI and MP_UNREACH_NLRI attributes [RFC4760]. The AFI and SAFI defined in [RFC7752] are re-used. For the PCEs in public network, the AFI / SAFI pair is 16388 / 71, while for the PCEs of a particular VPN, the AFI / SAFI pair is set to 16388 / 72.

A new NLRI Type is defined for PCE discovery information as below:

o Type = TBD: PCE Discovery NLRI

The format of PCE Discovery NLRI is shown in the following figure:

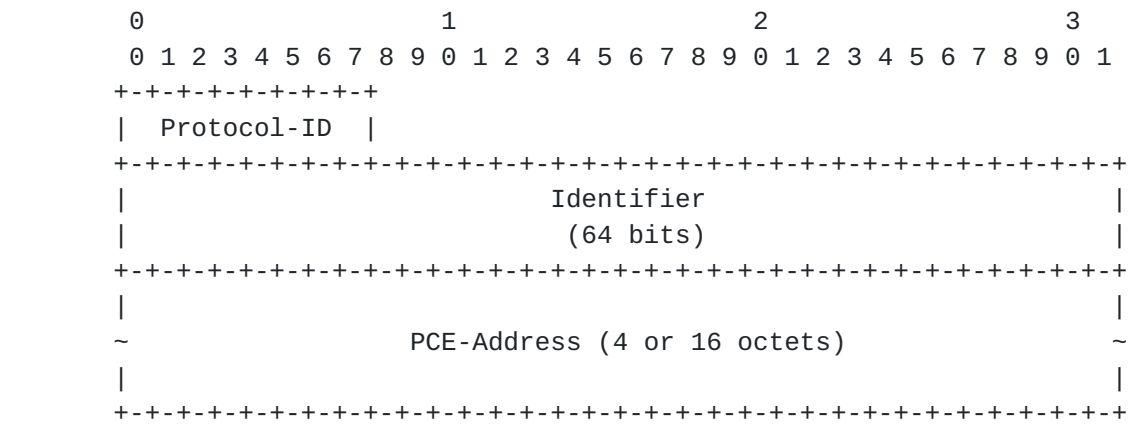

Figure 2. PCE Discovery NLRI

The 'Protocol-ID' field SHOULD be set to the appropriate value which indicates the source of the PCE discovery information. If BGP speaker and PCE are co-located, the Protocol-ID SHOULD be set to "Direct". In other cases, it is RECOMMENDED that the Protocol-ID value be set to "Static configuration".

As defined in [RFC7752], the 64-Bit 'Identifier' field is used to identify the "routing universe" where the PCE belongs.

2.2. PCE Discovery TLVs

The detailed PCE discovery information is carried in the BGP-LS attribute [RFC7752] with a new "PCE Discovery TLV", which contains a set of sub-TLVs for specific PCE discovery information. The PCE Discovery TLV and sub-TLVs SHOULD only be used with the PCE Discovery NLRI.

The format of the PCE Discovery TLV is shown as below:

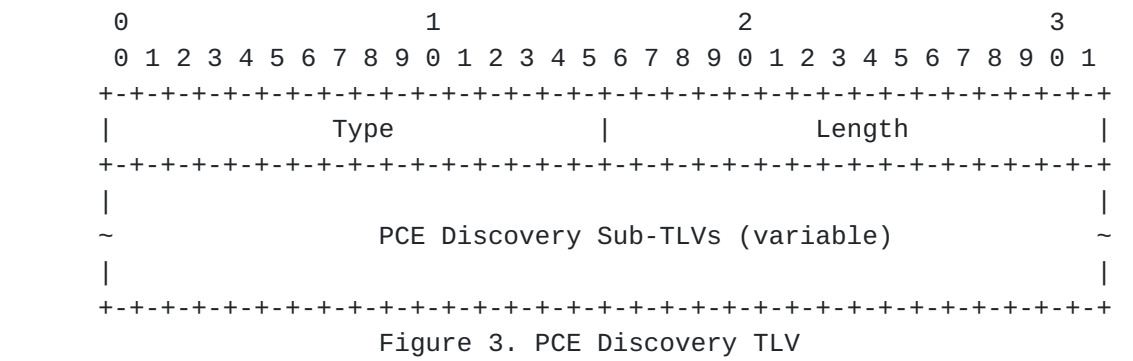

Figure 3. PCE Discovery TLV

The PCE Discovery sub-TLVs are listed as below. The format of the PCE Discovery sub-TLVs are consistent with the IGP PCED sub-TLVs as defined in [RFC5088] and [RFC5089]. The PATH-SCOPE sub-TLV MUST always be carried in the PCE Discovery TLV. Other PCE Discovery sub-

TLVs are optional and may facilitate the PCE selection process on the PCCs.

Type	Length	Name
1	3	PATH-SCOPE sub-TLV
2	variable	PCE-CAP-FLAGS sub-TLV
3	variable	OSPF-PCE-DOMAIN sub-TLV
4	variable	IS-IS-PCE-DOMAIN sub-TLV
5	variable	OSPF-NEIG-PCE-DOMAIN sub-TLV
6	variable	IS-IS-NEIG-PCE-DOMAIN sub-TLV

More PCE Discovery sub-TLVs may be defined in future. The format and semantic of new PCE Discovery sub-TLVs SHOULD be consistent in BGP and IGP based PCE discovery.

3. Operational Considerations

Existing BGP operational procedures apply to the advertisement of PCE discovery information. This information is treated as pure application level data which has no immediate impact on forwarding states. Normal BGP path selection can be applied to PCE Discovery NLRI only for the information propagation in the network, while on PCCs the PCE selection is based on the information carried in the PCE Discovery TLV. The PCE discovery information SHOULD be advertised only to the domains where such information is allowed to be used. This can be achieved by policy control on the ASBRs.

The PCE discovery information is considered relatively stable and does not change frequently, thus this information will not bring significant impact on the amount of BGP updates in the network.

4. IANA Considerations

IANA needs to assign a new NLRI Type for 'PCE Discovery NLRI' from the "BGP-LS NLRI-Types" registry.

IANA needs to assign a new TLV code point for 'PCE Discovery TLV' from the "node anchor, link descriptor and link attribute TLVs" registry.

IANA needs to create a new registry for "PCE Discovery Sub-TLVs". The registry will be initialized as shown in section 2.2 of this document.

5. Security Considerations

Procedures and protocol extensions defined in this document do not affect the BGP security model. See the 'Security Considerations' section of [RFC4271] for a discussion of BGP security. Also refer to [RFC4272] and [RFC6952] for analysis of security issues for BGP.

6. Contributors

The following individuals gave significant contributions to this document:

Takuya Miyasaka
KDDI Corporation
ta-miyasaka@kddi.com

7. Acknowledgements

The authors would like to thank Zhenbin Li, Hannes Gredler, Jan Medved, Adrian Farrel, Julien Meuric and Jonathan Hardwick for the valuable discussion and comments.

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, <http://www.rfc-editor.org/info/rfc2119>.

[RFC4271] Rekhter, Y., Ed., Li, T., Ed., and S. Hares, Ed., "A Border Gateway Protocol 4 (BGP-4)", RFC 4271, DOI 10.17487/RFC4271, January 2006, <http://www.rfc-editor.org/info/rfc4271>.

[RFC4760] Bates, T., Chandra, R., Katz, D., and Y. Rekhter, "Multiprotocol Extensions for BGP-4", RFC 4760, DOI 10.17487/RFC4760, January 2007, <http://www.rfc-editor.org/info/rfc4760>.

[RFC5088] Le Roux, JL., Ed., Vasseur, JP., Ed., Ikejiri, Y., and R. Zhang, "OSPF Protocol Extensions for Path Computation Element (PCE) Discovery", RFC 5088, DOI 10.17487/RFC5088, January 2008, <http://www.rfc-editor.org/info/rfc5088>.

[RFC5089] Le Roux, JL., Ed., Vasseur, JP., Ed., Ikejiri, Y., and R. Zhang, "IS-IS Protocol Extensions for Path Computation Element (PCE) Discovery", RFC 5089, DOI 10.17487/RFC5089, January 2008, <http://www.rfc-editor.org/info/rfc5089>.

[RFC7752] Gredler, H., Ed., Medved, J., Previdi, S., Farrel, A., and S. Ray, "North-Bound Distribution of Link-State and Traffic Engineering (TE) Information Using BGP", RFC 7752, DOI 10.17487/RFC7752, March 2016, <http://www.rfc-editor.org/info/rfc7752>.

8.2. Informative References

[I-D.ietf-idr-te-lsp-distribution] Dong, J., Chen, M., Gredler, H., Previdi, S., and J. Tantsura, "Distribution of MPLS Traffic Engineering (TE) LSP State using BGP", draft-ietf-idr-te-lsp-distribution-04 (work in progress), December 2015.

[I-D.ietf-idr-te-pm-bgp] Previdi, S., Wu, Q., Gredler, H., Ray, S., Tantsura, j., Filsfils, C., and L. Ginsberg, "BGP-LS Advertisement of IGP Traffic Engineering Performance Metric Extensions", draft-ietf-idr-te-pm-bgp-03 (work in progress), May 2016.

[RFC4272] Murphy, S., "BGP Security Vulnerabilities Analysis", RFC 4272, DOI 10.17487/RFC4272, January 2006, <http://www.rfc-editor.org/info/rfc4272>.

[RFC4674] Le Roux, J., Ed., "Requirements for Path Computation Element (PCE) Discovery", RFC 4674, DOI 10.17487/RFC4674, October 2006, <http://www.rfc-editor.org/info/rfc4674>.

[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, <http://www.rfc-editor.org/info/rfc5441>.

[RFC5824] Kumaki, K., Ed., Zhang, R., and Y. Kamite, "Requirements for Supporting Customer Resource ReSerVation Protocol (RSVP) and RSVP Traffic Engineering (RSVP-TE) over a BGP/MPLS IP-VPN", RFC 5824, DOI 10.17487/RFC5824, April 2010, <http://www.rfc-editor.org/info/rfc5824>.

[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, <http://www.rfc-editor.org/info/rfc6805>.

[RFC6952] Jethanandani, M., Patel, K., and L. Zheng, "Analysis of BGP, LDP, PCEP, and MSDP Issues According to the Keying and Authentication for Routing Protocols (KARP) Design Guide", RFC 6952, DOI 10.17487/RFC6952, May 2013, <http://www.rfc-editor.org/info/rfc6952>.

Authors' Addresses

Jie Dong
Huawei Technologies
Huawei Campus, No. 156 Beiqing Rd.
Beijing 100095
China

Email: jie.dong@huawei.com

Mach(Guoyi) Chen
Huawei Technologies
Huawei Campus, No. 156 Beiqing Rd.
Beijing 100095
China

Email: mach.chen@huawei.com

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

Email: dhruv.ietf@gmail.com

Jeff Tantsura
Individual
US

Email: jefftant.ietf@gmail.com

Kenji Kumaki
KDDI Corporation
Garden Air Tower, Iidabashi, Chiyoda-ku
Tokyo 102-8460
Japan

Email: ke-kumaki@kddi.com

Tomoki Murai
Furukawa Network Solution Corp.
5-1-9, Higashi-Yawata, Hiratsuka
Kanagawa 254-0016
Japan

Email: murai@fnsc.co.jp