Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: September 6, 2015

J. Dong
M. Chen
D. Dhody
Huawei Technologies
J. Tantsura
Ericsson

March 5, 2015

BGP Extensions for Path Computation Element (PCE) Discovery
draft-dong-pce-discovery-proto-bgp-02

Abstract

In networks where Path Computation Element (PCE) is used for centralized path computation, it is desirable for 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. This document specifies the BGP extensions for PCE discovery in these scenarios. 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].

Table of Contents

1. Introduction

In network scenarios where Path Computation Element (PCE) is used for centralized path computation, it is desirable for 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 PCE discovery mechanism based on Interior Gateway Protocol (IGP). Those IGP based mechanisms may not work in scenarios where the PCEs do not participate in the IGP, and it is difficult for PCEs to participate in IGP of multiple domains where PCE discovery is needed.

In some other scenarios, Backward Recursive Path Computation (BRPC) [RFC5441] can be used by cooperating PCEs to compute inter-domain path, in which case these cooperating PCEs should be known to each other. In case of inter-AS network where the PCEs do not participate in a common IGP, the existing IGP discovery mechanism cannot be used to discover the PCEs in other domains.

In the Hierarchical PCE scenario [RFC6805], the child PCEs need to know the address of the parent PCEs. This cannot be achieved through IGP based discovery, as normally the child PCEs and the parent PCE are under different administration and reside in different domains.

Besides, as BGP could be used for north-bound distribution of routing and Label Switched Path (LSP) information to PCE as described in [I-D.ietf-idr-ls-distribution] [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, some other PCE discovery mechanism is also needed.

A detailed set of requirements for a PCE discovery mechanism are provided in [RFC4674].

This document proposes to extend BGP for PCE discovery for the above scenarios. In networks where BGP-LS is already used for the north-bound routing information distribution to PCE, BGP based PCE discovery can reuse the existing BGP sessions and mechanisms to achieve PCE discovery. It should be noted that, in IGP domain, the IGP based PCE discovery mechanism may be used in conjunction with the BGP based PCE discovery. Thus the BGP based PCE discovery is complementary to the existing IGP based mechanisms.

```
                           +-----------+
                           |    PCE    |
                           +-----------+
                                 |
                                 v
                           +-----------+
                           |    BGP    |               +-----------+
                           |  Speaker  |               |    PCE    |
                           +-----------+               +-----------+
                             |   |   |                       |
                             |   |   |                       |
             +---------------+   |   +-------------------+   |
             v                   v                       v   v
       +-----------+       +-----------+           +-----------+
       |    BGP    |       |    BGP    |           |    BGP    |
       |  Speaker  |       |  Speaker  |    . . .  |  Speaker  |
       |   & PCC   |       |   & PCC   |           |           |
       +-----------+       +-----------+           +-----------+
                                                          |
                                                          |   via
                                                          |   IGP
                                                          v
                                                    +-----------+
                                                    |    PCC    |
                                                    +-----------+
```

Figure 1: BGP for PCE discovery

As shown in the network architecture in Figure 1, BGP is used for both routing information distribution and 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 between different PCEs. The PCCs maybe co-located with the BGP speakers as shown in Figure 1. The IGP based PCE discovery mechanism may be used for the distribution of PCE discovery information in IGP domain.

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 [I-D.ietf-idr-ls-distribution] are re-used, and 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:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+
|  Protocol-ID  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                           Identifier                          |
|                            (64 bits)                          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
~                  PCE-Address (4 or 16 octets)                 ~
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 2. PCE Discovery NLRI

The 'Protocol-ID' field do not apply to the PCE Discovery NLRI and SHOULD be set to 0 on transmission and be ignored upon receipt.

The 'Identifier' field is used to identify the "routing universe" where the PCE belongs, and the identifier values as below defined in [I-D.ietf-idr-ls-distribution] apply.

Identifier	Routing Universe
0	L3 packet topology
1	L1 optical topology

2.2. PCE Discovery TLVs

The detailed PCE discovery information is carried in BGP-LS attribute [I-D.ietf-idr-ls-distribution] 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:

```
  0                   1                   2                   3
  0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |              Type             |             Length            |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |                                                               |
 ~                 PCE Discovery Sub-TLVs (variable)             ~
 |                                                               |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

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 defined in [RFC5088] and [RFC5089]. The PATH-SCOPE TLV MUST always be carried in the BGP-LS Attribute if the NLRI is PCE Discovery NLRI. Other PCE Discovery TLVs are optional and may facilitate the PCE selection process.

Type	Length	Name
TBD	3	PATH-SCOPE sub-TLV
TBD	variable	PCE-CAP-FLAGS sub-TLV
TBD	variable	OSPF-PCE-DOMAIN sub-TLV
TBD	variable	IS-IS-PCE-DOMAIN sub-TLV
TBD	variable	OSPF-NEIG-PCE-DOMAIN sub-TLV
TBD	variable	IS-IS-NEIG-PCE-DOMAIN sub-TLV

More PCE Discovery sub-TLVs may be defined in future and the format SHOULD be in line with the new sub-TLVs defined for 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 the PCE selection on the PCCs would be performed based on the information carried in the PCE Discovery TLV.

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. Acknowledgements

The authors would like to thank Zhenbin Li and Hannes Gredler for their discussion and comments.

7. References

7.1. Normative References

[I-D.ietf-idr-ls-distribution]
Gredler, H., Medved, J., Previdi, S., Farrel, A., and S. Ray, "North-Bound Distribution of Link-State and TE Information using BGP", draft-ietf-idr-ls-distribution-10 (work in progress), January 2015.

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

[RFC4271] Rekhter, Y., Li, T., and S. Hares, "A Border Gateway Protocol 4 (BGP-4)", RFC 4271, January 2006.

[RFC4760] Bates, T., Chandra, R., Katz, D., and Y. Rekhter, "Multiprotocol Extensions for BGP-4", RFC 4760, January 2007.

[RFC5088] Le Roux, JL., Vasseur, JP., Ikejiri, Y., and R. Zhang, "OSPF Protocol Extensions for Path Computation Element (PCE) Discovery", RFC 5088, January 2008.

[RFC5089] Le Roux, JL., Vasseur, JP., Ikejiri, Y., and R. Zhang, "IS-IS Protocol Extensions for Path Computation Element (PCE) Discovery", RFC 5089, January 2008.

7.2. 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-02 (work in progress), January 2015.

[I-D.ietf-idr-te-pm-bgp]
Wu, Q., Previdi, S., Gredler, H., Ray, S., and J. Tantsura, "BGP attribute for North-Bound Distribution of Traffic Engineering (TE) performance Metrics", draft-ietf-idr-te-pm-bgp-02 (work in progress), January 2015.

[RFC4272] Murphy, S., "BGP Security Vulnerabilities Analysis", RFC 4272, January 2006.

[RFC4674] Le Roux, J., "Requirements for Path Computation Element (PCE) Discovery", RFC 4674, October 2006.

[RFC5441] Vasseur, JP., 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, April 2009.

[RFC6805] King, D. and A. Farrel, "The Application of the Path Computation Element Architecture to the Determination of a Sequence of Domains in MPLS and GMPLS", RFC 6805, November 2012.

[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, May 2013.

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
Leela Palace
Bangalore, Karnataka 560008
India

Email: dhruv.ietf@gmail.com

Jeff Tantsura
Ericsson
300 Holger Way
San Jose, CA 95134
US

Email: jeff.tantsura@ericsson.com