

PCE Working Group
Internet-Draft
Intended status: Experimental
Expires: March 19, 2015

D. Dhody
Q. Wu
U. Pallé
X. Zhang
Huawei Technologies
September 15, 2014

**PCE support for Domain Diversity
draft-dwpz-pce-domain-diverse-02**

Abstract

The Path Computation Element (PCE) may be used for computing path for services that traverse multi-area and multi-AS Multiprotocol Label Switching (MPLS) and Generalized MPLS (GMPLS) Traffic Engineered (TE) networks.

Path computation should facilitate the selection of paths with domain diversity. This document examines the existing mechanisms to do so and further propose some extensions to Path Computation Element Protocol (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 <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 March 19, 2015.

Copyright Notice

Copyright (c) 2014 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 (<http://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
1.1.	Requirements Language	3
2.	Terminology	3
3.	Domain Diversity	3
3.1.	Per Domain Path Computation	4
3.2.	Backward-Recursive PCE-based Computation	4
3.3.	Hierarchical PCE	5
3.3.1.	End to End Path	5
3.3.2.	Domain-Sequence	5
4.	Extension to PCEP	6
4.1.	SVEC Object	6
4.2.	Minimize Shared Domains	7
4.3.	Relationship between SVEC Diversity Flags and OF	7
5.	Other Considerations	8
5.1.	Transit Domain Identifier	8
5.2.	Diversity v/s Optimality	8
6.	Security Considerations	9
7.	Manageability Considerations	9
7.1.	Control of Function and Policy	9
7.2.	Information and Data Models	9
7.3.	Liveness Detection and Monitoring	9
7.4.	Verify Correct Operations	9
7.5.	Requirements On Other Protocols	9
7.6.	Impact On Network Operations	9
8.	IANA Considerations	9
9.	Acknowledgments	9
10.	References	10
10.1.	Normative References	10
10.2.	Informative References	10
Appendix A.	Contributor Addresses	12

[1.](#) Introduction

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 requirement. 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).

In a multi-domain environment, Domain Diversity is defined in [\[RFC6805\]](#). A pair of paths are domain-diverse if they do not traverse any of the same transit domains. Domain diversity may be maximized for a pair of paths by selecting paths that have the smallest number of shared domains. Path computation should facilitate the selection of domain diverse paths as a way to reduce the risk of shared failure and automatically helps to ensure path diversity for most of the route of a pair of LSPs.

The main motivation behind domain diversity is to avoid fate sharing, but it can also be because of some geo-political reasons and commercial relationships that would require domain diversity. for example, a pair of paths should choose different transit Autonomous System (AS) because of some policy considerations.

In case when full domain diversity could not be achieved, it is helpful to minimize the common shared domains. Also it is interesting to note that other scope of diversity (node, link, SRLG etc) can still be applied inside the common shared domains.

This document examine a way to achieve domain diversity with existing inter-domain path computation mechanism like per-domain path computation technique [\[RFC5152\]](#), Backward Recursive Path Computation (BRPC) mechanism [\[RFC5441\]](#) and Hierarchical PCE [\[RFC6805\]](#). This document also considers synchronized as well as non-synchronized dependent path computations. Since independent and synchronized path computation cannot be used to apply diversity, it is not discussed in this document.

1.1. 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 [\[RFC2119\]](#).

2. Terminology

The terminology is as per [\[RFC5440\]](#).

3. Domain Diversity

As described in [\[RFC6805\]](#), a set of paths are considered to be domain diverse if they do not share any transit domains, apart from ingress and egress domains.

Some additional parameters to consider would be -

Minimize shared domain: When a fully domain diverse path is not possible, PCE could be requested to minimize the number of shared transit domains. This can also be termed as maximizing partial domain diversity. Other scope of diversity (node, link, SRLG etc) can still be applied inside the common shared domains.

Boundary Nodes: Diversity in boundary node selection can be achieved by node diversity.

3.1. Per Domain Path Computation

The per domain path computation technique [[RFC5152](#)] 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.

Non-Synchronized Path Computation: Path computations are performed in a serialized and independent fashion. After the setup of primary path, a domain diverse path can be signaled by encoding the transit domain identifiers in exclude route object (XRO) or explicit exclusion route subobject (EXRS) using domain sub-objects defined in [[DOMAIN-SUBOBJ](#)] and [[RFC3209](#)] in RSVP-TE. Note that the head end LSR should be aware of transit domain identifiers of the primary path to be able to do so. Also a head end label switching router (LSR) can signal a path by using a domain diverse domain sequence known in priori and encoded in explicit route object (ERO) in path message.

Synchronized Path Computation: Not Applicable.

3.2. Backward-Recursive PCE-based Computation

The BRPC [[RFC5441](#)] technique 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 maybe known before the path computation, but it can also be used when the domain path is unknown and determined during path computation.

Non-Synchronized Path Computation: Path computations are performed in a serialized and independent fashion. After the path computation of the primary path, a domain diverse path computation request is sent by PCC to the PCE, by encoding the transit domain identifiers in XRO or EXRS using domain sub-objects defined in [[PCE-DOMAIN](#)] and [[RFC3209](#)] in PCEP. Note that the PCC should be

aware of transit domain identifiers of the primary path to be able to do so. Also a PCC can request a path by using a domain diverse domain sequence known in priori and encoded in include route object (IRO) in path request message.

Synchronized Path Computation: Not Applicable. [Since different transit domain PCEs may be involved , there is difficulty to achieve synchronization for domain diverse path computation]. Note that [[RFC5440](#)] describes other diversity parameters (node, link, SRLG etc) that may be applied.

3.3. 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.

3.3.1. End to End Path

Non-Synchronized Path Computation: Path computations are performed in a serialized and independent fashion. After the path computation of the primary path, a domain diverse path computation request is sent to the parent PCE, by encoding the transit domain identifiers in XRO or EXRS using domain sub-objects defined in [[PCE-DOMAIN](#)] and [[RFC3209](#)] in PCEP. Note that the PCC should be aware of transit domain identifiers of the primary path to be able to do so. The parent PCE should provide a domain diverse end to end path.

Synchronized Path Computation: Child PCE should be able to request dependent and synchronized domain diverse end to end paths from its parent PCE. A new flag is added in synchronized vectore (SVEC) object for this (Refer [Section 4.1](#)).

3.3.2. Domain-Sequence

Non-Synchronized Path Computation: Path computations are performed in a serialized and independent fashion. After the primary path computation using H-PCE (involving domain-sequence selection by parent PCE and end-to-end path computation via BRPC or Per-Domain mechanisms), a domain diverse path computation request is sent to the parent PCE, by encoding the transit domain identifiers in XRO or EXRS using domain sub-objects defined in [[PCE-DOMAIN](#)] and [[RFC3209](#)] in PCEP. Note that the PCC should be aware of transit domain identifiers of the primary path to be able to do so. The parent PCE should provide a diverse domain sequence.

Synchronized Path Computation: Child PCE should be able to request dependent and synchronized diverse domain-sequence(s) from it's parent PCE. A new flag is added in SVEC object for this (Refer [Section 4.1](#)). The parent PCE should reply with diverse domain sequence(s) encoded in ERO as described in [[PCE-DOMAIN](#)].

4. Extension to PCEP

[Editor's Note: It has been requested to move this section to the HPCE-Extension document - [draft-ietf-pce-hierarchy-extensions](#). This section would be removed from this document once that is done.]

4.1. SVEC Object

[RFC5440] defines SVEC object which includes flags for the potential dependency between the set of path computation requests (Link, Node and SRLG diverse). This document proposes a new flag 0 for domain diversity.

The format of the SVEC object body is as follows:

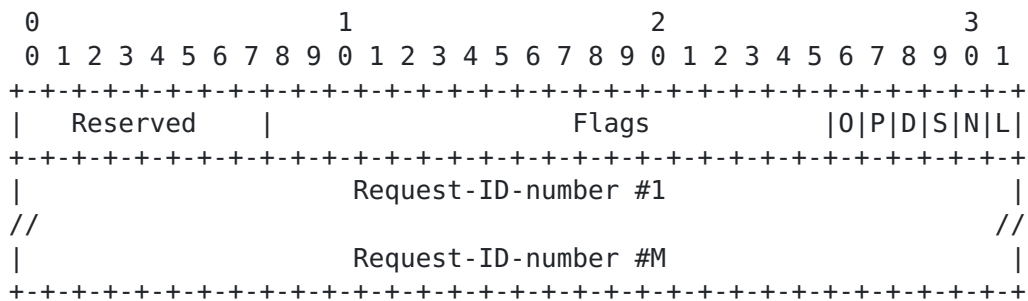


Figure 1: SVEC Body Object Format

Following new bit is added in the Flags field:

- * 0 (Domain diverse) bit: when set, this indicates that the computed paths corresponding to the requests specified by the following RP objects MUST NOT have any transit domain(s) in common.

The Domain Diverse 0-bit can be used in Hierarchical PCE path computation to compute synchronized domain diverse end to end path or diverse domain sequences as described in [Section 3.3](#).

When domain diverse 0 bit is set, it is applied to the transit domains. The other bit in SVEC object (N, L, S etc) is set, should still be applied in the ingress and egress domain.

4.2. Minimize Shared Domains

In case when full domain diversity could not be achieved, it maybe helpful to minimize the common shared domains. It's interesting to note that diversity (node, link etc) can still be applied inside the common shared transit domains as well as for ingress and egress domain via the bits in SVEC object (N, L, S etc).

A new Objective function (OF) [[RFC5541](#)] code for synchronized path computation requests is proposed:

MCTD

- * Name: Minimize the number of Common Transit Domains.
- * Objective Function Code: TBD
- * Description: Find a set of paths such that it passes through the least number of common transit domains.

The MCTD OF can be used in Hierarchical PCE path computation to request synchronized domain diverse end to end paths or diverse domain sequences as described in [Section 3.3](#).

[Editor's Note: A new document is created for the OF for minimizing shared node, links, SRLGs inside the domain - [[PCE-OF-DIVERSE](#)].]

For non synchronized diverse domain path computation the X bit in XRO or EXRS [[RFC5521](#)] sub-objects can be used, where X bit set as 1 indicates that the domain specified SHOULD be excluded from the path computed by the PCE, but MAY be included subject to PCE policy and the absence of a viable path that meets the other constraints and excludes the domain.

4.3. Relationship between SVEC Diversity Flags and OF

[RFC5440] uses SVEC diversity flag for node, link or SRLG to describe the potential disjointness between the set of path computation requests used in PCEP protocol. This document further extends by adding domain-diverse 0-bit in SVEC object and a new OF Code for minimizing the number of shared transit domain.

Further [[PCE-OF-DIVERSE](#)] defines three new OF codes to maximize diversity as much as possible, in other words, minimize the common shared resources (Node,Link or SRLG) between a set of paths.

It may be interesting to note that the diversity flags in the SVEC object and OF for diversity can be used together. Some example of usage are listed below -

- o SVEC object with domain-diverse bit=1 - ensure full domain-diversity.
- o SVEC object with domain-diverse bit=1 and node/link diverse bit=1 - ensure full domain-diversity, as well as node/link diverse in ingress and egress domain.
- o SVEC object with domain-diverse bit=0 and OF=MCTD - domain-diversity as much as possible.
- o SVEC object with domain-diverse bit=0; node/link diverse bit=1 and OF=MCTD - domain-diversity as much as possible, as well as node/link diverse in ingress, egress and shared transit domains.
- o SVEC object with domain-diverse bit=1 and OF=MCTD - ensure full domain-diversity.

5. Other Considerations

5.1. Transit Domain Identifier

In case of non-synchronized path computation, Ingress node (i.e. a PCC) should be aware of transit domain identifiers of the primary path. So during the path computation or signaling of the primary path, the transit domain should be identified.

A possible solution for path computation could be a flag in RP object requesting domain identifier to be returned in the PCEP path reply message.

[Editor's Note: There should be a mechanism in signaling and path computation to obtain the domain information. Further details - TBD]

5.2. Diversity v/s Optimality

In case of non-synchronized path computation, PCE may be requested to provide an optimal primary path first and then PCC requests for a backup path with exclusion. Note that this approach does not guarantee diversity compared to disjoint path computations for primary and backup path in a synchronized manner.

A synchronized path computation with diversity flags and/or objective function is used to make sure that both the primary path and the backup path can be computed simultaneously with full diversity or

optimized to be as diverse as possible. In the latter case we may sacrifice optimal path for diversity, thus there is a trade-off between the two.

An implementation may further choose to analyze the trade-off i.e. it may send multiple request to PCE asking to optimize based on diversity as well as say, cost and make an intelligent choice between them.

6. Security Considerations

TBD.

7. Manageability Considerations

7.1. Control of Function and Policy

TBD.

7.2. Information and Data Models

TBD.

7.3. Liveness Detection and Monitoring

TBD.

7.4. Verify Correct Operations

TBD.

7.5. Requirements On Other Protocols

TBD.

7.6. Impact On Network Operations

TBD.

8. IANA Considerations

TBD.

9. Acknowledgments

We would like to thank Qilei Wang for starting this discussion in the mailing list.

10. References

10.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", [BCP 14](#), [RFC 2119](#), March 1997.
- [RFC5152] Vasseur, JP., Ayyangar, A., and R. Zhang, "A Per-Domain Path Computation Method for Establishing Inter-Domain Traffic Engineering (TE) Label Switched Paths (LSPs)", [RFC 5152](#), February 2008.
- [RFC5440] Vasseur, JP. and JL. Le Roux, "Path Computation Element (PCE) Communication Protocol (PCEP)", [RFC 5440](#), March 2009.
- [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.

10.2. 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](#), December 2001.
- [RFC5521] Oki, E., Takeda, T., and A. Farrel, "Extensions to the Path Computation Element Communication Protocol (PCEP) for Route Exclusions", [RFC 5521](#), April 2009.
- [RFC5541] Le Roux, JL., Vasseur, JP., and Y. Lee, "Encoding of Objective Functions in the Path Computation Element Communication Protocol (PCEP)", [RFC 5541](#), June 2009.
- [DOMAIN-SUBOBJ] Dhody, D., Palle, U., Kondreddy, V., and R. Casellas, "Domain Subobjects for Resource ReserVation Protocol - Traffic Engineering (RSVP-TE). ([draft-ietf-ccamp-rsvp-te-domain-subobjects](#))", July 2014.

[PCE-DOMAIN]

Dhody, D., Palle, U., and R. Casellas, "Standard Representation Of Domain Sequence. ([draft-ietf-pce-pcep-domain-sequence](#))", July 2014.

[PCE-OF-DIVERSE]

Dhody, D. and Q. Wu, "PCE support for Maximizing Diversity. ([draft-dhody-pce-of-diverse](#))", June 2014.

Appendix A. Contributor Addresses

Ramon Casellas
CTTC - Centre Tecnologic de Telecomunicacions de Catalunya
Av. Carl Friedrich Gauss n7
Castelldefels, Barcelona 08860
Spain

EMail: ramon.casellas@cttc.es

Avantika
Huawei Technologies
Leela Palace
Bangalore, Karnataka 560008
India

EMail: avantika.sushilkumar@huawei.com

Authors' Addresses

Dhruv Dhody
Huawei Technologies
Leela Palace
Bangalore, Karnataka 560008
India

EMail: dhruv.ietf@gmail.com

Qin Wu
Huawei Technologies
101 Software Avenue, Yuhua District
Nanjing, Jiangsu 210012
China

EMail: bill.wu@huawei.com

Udayasree Palle
Huawei Technologies
Leela Palace
Bangalore, Karnataka 560008
India

EMail: udayasree.palle@huawei.com

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

EMail: zhang.xian@huawei.com