Internet Engineering Task Force
Internet-Draft
Intended status: Standards Track
Expires: September 6, 2012

H. Chen
Huawei Technologies

March 5, 2012

The Applicability of the PCE to Computing Protection and Recovery Paths for Single Domain and Multi-Domain Networks

draft-chen-pce-protection-applicability-00.txt

Abstract

The Path Computation Element (PCE) provides path computation functions in support of traffic engineering in Multiprotocol Label Switching (MPLS) and Generalized MPLS (GMPLS) networks.

A link or node failure can significantly impact network services in large-scale networks. Therefore it is important to ensure the survivability of large scale networks which consist of various connections provided over multiple interconnected networks with varying technologies.

This document examines the applicability of the PCE architecture, protocols, and procedures for computing protection paths and restoration services, for single and multi-domain networks.

Status of this Memo

Copyright Notice

Table of Contents

1. Introduction

Network survivability remains a major concern for network operators and service providers, particularly as expanding applications such as private and Public Cloud drive increasingly more traffic across longer ranges, to a wider number of users. A variety of well-known pre-planned protection and post-fault recovery schemes have been developed for IP, MPLS and GMPLS networks.

The Path Computation Element (PCE) [RFC4655] can be used to perform complex path computation in large single domain, multi-domain and multi-layered networks. The PCE can also be used to compute a variety of restoration and protection paths and services.

This document examines the applicability of the PCE architecture, protocols, and protocol extensions for computing protection paths and restoration services.

1.1. Domains

A domain can be defined as a separate administrative, geographic, or switching environment within the network. A domain may be further defined as a zone of routing or computational ability. Under these definitions a domain might be categorized as an Antonymous System (AS) or an Interior Gateway Protocol (IGP) area (as per [RFC4726] and [RFC4655]), or specific switching environment.

In the context of GMPLS, a particularly important example of a domain is the Automatically Switched Optical Network (ASON) subnetwork [G-8080]. In this case, computation of an end-to-end path requires the selection of nodes and links within a parent domain where some nodes may, in fact, be subnetworks. Furthermore, a domain might be an ASON routing area [G-7715]. A PCE may perform the path computation function of an ASON routing controller as described in [G-7715-2].

It is assumed that the PCE architecture should be applied to small inter-domain topologies and not to solve route computation issues across large groups of domains, I.E. the entire Internet.

Most existing protocol mechanisms for network survivability have focused on single-domain scenarios. Multi-domain scenarios are much more complex and challenging as domain topology information is typically not shared outside each specific domain.

Therefore multi-domain survivability is a key requirement for today's complex networks. It is important to develop more adaptive multi-domain recovery solutions for various failure scenarios.

1.1.1. Inter-domain LSPs

Three signaling options are defined for setting up an inter-area or inter-AS LSP [RFC4726]:

- Contiguous LSP

- Stitched LSP

- Nested LSP

Three signaling options are defined for setting up an inter-area or inter-AS LSP [RFC4726]:

1.2. Recovery

Typically traffic-engineered networks such as MPLS-TE and GMPLS, use protection and recovery mechanisms based on the pre-established use of a packet or optical LSP and/or the availability of spare resources and the network topology.

1.3. Terminology

ABR: IGP Area Border Router, a router that is attached to more than one IGP area.

ASBR: Autonomous System Border Router, a router used to connect together ASs of a different or the same Service Provider via one or more inter-AS links.

CSPF: Constrained Shortest Path First.

Inter-area TE LSP: A TE LSP whose path transits through two or more IGP areas.

Inter-AS MPLS TE LSP: A TE LSP whose path transits through two or more ASs or sub-ASs (BGP confederations.

SRLG: Shared Risk Link Group.

TED: Traffic Engineering Database, which contains the topology and resource information of the domain. The TED may be fed by Interior Gateway Protocol (IGP) extensions or potentially by other means.

This document also uses the terminology defined in [RFC4655] and [RFC5440].

1.4. Conventions used in this document

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].

In this document, these words will appear with that interpretation only when in ALL CAPS. Lower case uses of these words are not to be interpreted as carrying RFC-2119 significance.

2. Path Computation Element Architecture Considerations

For the purpose of this document it is assumed that the path computation is the sole responsibility of the PCE as per the architecture defined in [RFC4655]. When a path is required the Path Computation Client (PCC) will send a request to the PCE. The PCE will apply the required constraints and compute a path and return a response to the PCC. In the context of this document it may be necessary for the PCE to co-operate with other PCEs in adjacent domains (as per BRPC [RFC5441]) or cooperate with the Parent PCE (as per [H-PCE]).

A PCE may be used to compute end-to-end paths across single or multiple domains. Multiple PCEs may be dedicated to each area to provide sufficient path computation capacity and redundancy for each domain.

During path computation [RFC5440], a PCC request may contain backup LSP requirements in order to setup in the same time the primary and backup LSPs. This request is known as dependent path computations. A typical dependent request for a primary and backup service would request that the computation assign a set of diverse paths, so both services are disjointed from each other.

2.1. Online Path Computation

Online path computation is performed on-demand as nodes in the network determine that they need to know the paths to use for services.

2.2. Offline Path Computation

Offline path computation is performed ahead of time, before the LSP setup is requested. That means that it is requested by, or performed as part of, a management application.

This method of computation allows the optimal placement of services

and explicit control of services. A CSP can plan where new protection services will be placed ahead of time. Furthermore by computing paths offline specific scenarios can be considered and a global view of network resources is available.

Finally, offline path computation provides a method to compute protection paths in the event of a single, or multiple, link failures. This allows the placement of backup services in the event of catastrophic network failures.

3. Protection Service Traffic Engineering

3.1. Path Computation

This document describes how the PCE architecture defined in [RFC4655] may be utilized to compute protection and recovery paths for critical network services. In the context of this document (inter-domain) it may be necessary for the PCE to co-operate with other PCEs in adjacent domains (as per BRPC [RFC5441]) or cooperate with the Parent PCE (as per [H-PCE]).

3.2. Bandwidth Reservation

3.3. Disjoint Path

Disjoint paths are required for end-to-end protection services. A backup service may be required to be fully disjoint from the primary service, link disjoint (allowing common nodes on the paths), or best-effort disjoint (allowing shared links or nodes when no other path can be found).

3.4. Service Preemption

3.5. Share Risk Link Groups

3.6. Multi-Homing

Networks constructed from multi-areas or multi-AS environments may have multiple interconnect points (multi-homing). End-to-end path computations may need to use different interconnect points to avoid single point failures disrupting primary and backup services.

Domain and path diversity may also be required when computing end-to-end paths. Domain diversity should facilitate the selection of paths that share ingress and egress domains, but do not share transit domains. Therefore, there must be a method allowing the inclusion or exclusion of specific domains when computing end-to-end paths.

3.6.1. Ingress and Egress Protection

An end-to-end primary service carried by a primary TE LSP from a primary ingress node to a primary egress node may need to be protected against the failures in the ingress and the egress. In this case, a backup ingress and a backup egress are required, which are different from the primary ingress and the primary egress respectively. The backup ingress should be in the same domain as the primary ingress, and the backup egress should be in the same domain as the primary egress.

A source of the service traffic may be sent to both the primary ingress and the backup ingress (dual-homing). The source may not be in the same domain as the primary ingress and the backup ingress. When the primary ingress fails, the service traffic is delivered through the backup ingress.

A receiver of the service traffic may be connected to both the primary egress and the backup egress (dual-homing). The receiver may not be in the same domain as the primary egress and the backup egress. When the primary egress fails, the receiver gets the service traffic from the backup egress.

4. Packet Protection Applications

Network survivability is a key objective for CSPs, particularly as expanding revenue services (cloud and data center applications) are increasing exponentially.

Pre-fault paths are pre-computed and protection resources are reserved a priory for rapid recovery. In the event of a network failure on the primary path, the traffic is fast switched to the backup path. These pre-provisioned mechanisms are capable of ensuring protection against single link failures.

Post-fault restoration schemes are reactive and require a reactive routing procedure to set up new working paths in the event of a failure. Post fault restoration can significantly impact network services as they are typically impacted by longer restoration delays and cannot guarantee recovery of a service. However, they are much more network resource efficient and are capable of handling multi-failure situations.

4.1. Single Domain Service Protection

A variety of pre-planned protection and post-fault restoration recovery schemes are available for single domain MPLS and GMPLS

networks, these include:

o Path Recovery

o Path Segment Recovery

o Local Recovery (Fast Reroute)

4.2. Multi-domain Service Protection

Typically network survivability has focused on single-domain scenarios. By contrast, broader multi-domain scenarios are much more challenging as no single entity has a global view of topology information. As a result, multi-domain survivability is very important.

A PCE may be used to compute end-to-end paths across multi-domain environments using a per-domain path computation technique [RFC5152]. The so called backward recursive path computation (BRPC) mechanism [RFC5441] defines a PCE-based path computation procedure to compute inter-domain constrained LSPs.

4.3. Backup Path Computation

A PCE can be used to compute backup paths in the context of fast reroute protection of TE LSPs. In this model, all backup TE LSPs protecting a given facility are computed in a coordinated manner by a PCE. This allows complete bandwidth sharing between backup tunnels protecting independent elements, while avoiding any extensions to TE LSP signaling. Both centralized and distributed computation models are applicable. In the distributed case each LSR can be a PCE to compute the paths of backup tunnels to protect against the failure of adjacent network links or nodes.

4.4. Fast Reroute (FRR) Path Computation

[RFC409] extends RSVP-TE to establish backup LSPs for the local repair of LSP tunnels. This extension allows CSPs to redirect traffic onto the backup LSP tunnels in 10s of milliseconds.

This local repair of the LSP is applicable to explicitly-routed LSPs. An FRR repair method is applicable to explicitly-routed LSPs across a single domain environment.

4.5. Point-to-Multipoint Path Protection

A PCE utilizing the extensions outlined in [RFC6006] (Extensions to PCEP for Point-to-Multipoint Traffic Engineering Label Switched

Paths), can be used to compute point-to-multipoint (P2MP) paths.

A PCC requesting path computation for a primary and backup path can request that these dependent computations use diverse paths. Furthermore, the specification also defines two new options for P2MP path dependent computation requests. The first option allows the PCC to request that the PCE should compute a secondary P2MP path tree with partial path diversity for specific leaves or a specific source-to-leaf (sub-path to the primary P2MP path tree. The second option, allows the PCC to request that partial paths should be link direction diverse

5. Optical Protection Applications

5.1. ASON Applicability

5.2. Multi-domain Restoration

6. Path and Service Protection Gaps

7. Manageability Considerations

This document does not describe any specific protocol, protocol extensions, or protocol usage, therefore no manageability considerations need to be discussed here.

8. Security Requirements

This document is informational and does not describe any new specific protocol, protocol extensions, or protocol usage. As such, it introduces no new security concerns.

9. IANA Considerations

This document makes no requests for IANA action.

10. Acknowledgement

The authors would like to thank Daniel King for his valuable contributions to this draft.

11. References

11.1. Normative References

11.2. Informative References

[RFC3473] Berger, L., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions", RFC 3473, January 2003.

[RFC4090] Pan, P., Swallow, G., and A. Atlas, "Fast Reroute Extensions to RSVP-TE for LSP Tunnels", RFC 4090, May 2005.

[RFC4655] Farrel, A., Vasseur, J., and J. Ash, "A Path Computation Element (PCE)-Based Architecture", RFC 4655, August 2006.

[RFC4726] Farrel, A., Vasseur, J., and A. Ayyangar, "A Framework for Inter-Domain Multiprotocol Label Switching Traffic Engineering", RFC 4726, November 2006.

[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.

[RFC6006] Zhao, Q., King, D., Verhaeghe, F., Takeda, T., Ali, Z., and J. Meuric, "Extensions to the Path Computation Element Communication Protocol (PCEP) for Point-to-Multipoint Traffic Engineering Label Switched Paths", RFC 6006, September 2010.

[G-8080] ITU-T Recommendation G.8080/Y.1304, Architecture for the automatically switched optical network (ASON).

[G-7715] ITU-T Recommendation G.7715 (2002), Architecture and Requirements for the Automatically Switched Optical Network (ASON).

[G-7715-2] ITU-T Recommendation G.7715.2 (2007), ASON routing architecture and requirements for remote route query.

[H-PCE] King, D. and A. Farrel, "The Application of the Path Computation Element Architecture to the Determination of a Sequence of Domains in MPLS & GMPLS", July

Author's Address

Huaimo Chen
Huawei Technologies
Boston, MA
USA

Email: Huaimochen@huawei.com