PCE Working Group
Internet-Draft
Intended status: Standards Track
Expires: April 18, 2014

Q. Wu
D. Dhody
Huawei
S. Previdi
Cisco Systems, Inc
October 15, 2013

Extensions to Path Computation Element Communication Protocol (PCEP) for handling Link Bandwidth Utilization
draft-wu-pce-pcep-link-bw-utilization-00

Abstract

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.

Link bandwidth utilization considering the total bandwidth of a link in current use for the forwarding is an important factor to consider during path computation. This document describes extensions to PCEP to consider them as new constraints during path computation.

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 April 18, 2014.

Copyright Notice

Table of Contents

1. Introduction

Real time link bandwidth utilization is becoming critical in the path computation in some networks. It is important that link bandwidth utilization is factored in during path computation. PCC can request a PCE to provide a path such that it selects under-utilized links. This document extends PCEP [RFC5440] for this purpose.

Traffic Engineering Database (TED) as populated by Interior Gateway Protocol (IGP) contains Maximum bandwidth, Maximum reservable bandwidth and Unreserved bandwidth ([RFC3630] and [RFC3784]). [OSPF-TE-EXPRESS] and [ISIS-TE-EXPRESS] further populate Residual bandwidth and Available bandwidth. Further [ISIS-TE-EXPRESS] also define Bandwidth Utilization.

[Editors Note: [OSPF-TE-EXPRESS] should also be extended in future version for real time link bandwidth utilization]

The links in the path MAY be monitored for changes in the link bandwidth utilization, re-optimization of such path MAY be further requested.

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 following terminology is used in this document.

IGP: Interior Gateway Protocol. Either of the two routing protocols, Open Shortest Path First (OSPF) or Intermediate System to Intermediate System (IS-IS).

PCC: Path Computation Client: any client application requesting a path computation to be performed by a Path Computation Element.

PCE: Path Computation Element. An entity (component, application, or network node) that is capable of computing a network path or route based on a network graph and applying computational constraints.

PCEP: Path Computation Element Protocol.

RSVP: Resource Reservation Protocol

TE LSP: Traffic Engineering Label Switched Path.

3. Link Bandwidth Utilization (LBU)

The bandwidth utilization on a link, forwarding adjacency, or bundled link is populated in the TED (Bandwidth Utilization in [ISIS-TE-EXPRESS]). For a link or forwarding adjacency, bandwidth utilization represent the actual utilization of the link (i.e.: as measured in the router). For a bundled link, bandwidth utilization is defined to be the sum of the component link bandwidth utilization. This includes traffic for both RSVP and non-RSVP.

LBU Percentage is described as the (LBU / Maximum bandwidth) * 100.

4. Link Reserved Bandwidth Utilization (LRBU)

The reserved bandwidth utilization on a link, forwarding adjacency, or bundled link can be calculated from the TED. This includes traffic for only RSVP-TE LSPs.

LRBU can be calculated by using the Residual bandwidth, available bandwidth and LBU. The actual bandwidth by non-RSVP TE traffic can be calculated by subtracting Available Bandwidth from Residual Bandwidth. Once we have the actual bandwidth for non-RSVP TE traffic, subtracting this from LBU would result in LRBU.

LRBU Percentage is described as the (LRBU / (current reserved bandwidth)) * 100; where the current reserved bandwidth can be calculated by subtracting Residual bandwidth from Maximum bandwidth.

5. PCEP Requirements

Following requirements associated with bandwidth utilization are identified for PCEP:

1. PCE supporting this draft MUST have the capability to compute end-to-end path with bandwidth utilization constraints. It MUST also support the combination of bandwidth utilization constraint with existing constraints (cost, hop-limit...).

2. PCC MUST be able to request for bandwidth utilization constraint in PCReq message as the boundary condition that should not be crossed for each link in the path.

3. PCC MUST be able to request for bandwidth utilization constraint in PCReq message as an Objective function (OF) [RFC5541] to be

optimized.

4. PCEs are not required to support bandwidth utilization constraint. Therefore, it MUST be possible for a PCE to reject a PCReq message with a reason code that indicates no support for bandwidth utilization constraint.

5. PCEP SHOULD provide mechanism to handle bandwidth utilization constraint in multi-domain (e.g., Inter-AS, Inter-Area or Multi-Layer) environment.

6. PCEP Extensions

This section defines extensions to PCEP [RFC5440] for requirements outlined in Section 5. The proposed solution is used to consider bandwidth utilization during path computation.

6.1. BU Object

The BU (Bandwidth Utilization) is used to indicate the upper limit of the acceptable link bandwidth utilization percentage.

The BU object may be carried within the PCReq message and PCRep messages.

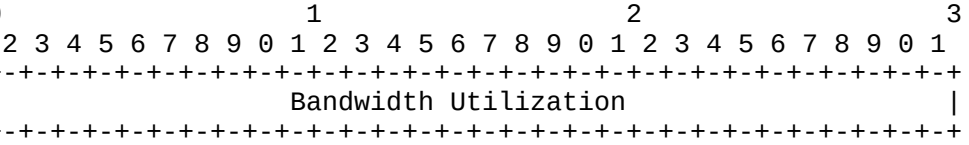
BU Object-Class is TBD.

Two Object-Type values are defined for the BU object:

- Link Bandwidth Utilization (LBU): BU Object-Type is 1.

- Link Reserved Bandwidth Utilization (LRBU): BU Object-Type is 2.

The format of the BU object body is as follows:



BU Object Body Format

Bandwidth utilization (32 bits): Represents the bandwidth utilization quantified as a percentage (as described in Section 3 and Section 4). The basic unit is 0.000000023%, with the maximum value 4,294,967,295 representing 98.784247785% (4,294,967,295 * 0.000000023%). This value is the maximum Bandwidth utilization

percentage that can be expressed.

The BU object body has a fixed length of 4 bytes.

6.1.1. Elements of Procedure

PCC SHOULD request the PCE to factor in the bandwidth utilization during path computation by including a BU object in the PCReq message.

Multiple BU objects MAY be inserted in a PCReq or a PCRep message for a given request but there MUST be at most one instance of the BU object for each object type. If, for a given request, two or more instances of a BU object with the same object type are present, only the first instance MUST be considered and other instances MUST be ignored.

BU object MAY be carried in a PCRep message in case of unsuccessful path computation along with a NO-PATH object to indicate the constraints that could not be satisfied.

If the P bit is clear in the object header and PCE does not understand or does not support bandwidth utilization during path computation it SHOULD simply ignore BU object.

If the P Bit is set in the object header and PCE receives BU object in path request and it understands the BU object, but the PCE is not capable of bandwidth utilization check during path computation, the PCE MUST send a PCErr message with a PCEP-ERROR Object Error-Type = 4 (Not supported object) [RFC5440]. The path computation request MUST then be cancelled.

If the PCE does not understand the BU object, then the PCE MUST send a PCErr message with a PCEP-ERROR Object Error-Type = 3 (Unknown object) [RFC5440].

6.2. New Objective Functions

This document defines two additional objective functions -- namely, MUP (Maximum Under-Utilized Path) and MRUP (Maximum Reserved Under-Utilized Path. Hence two new objective function codes have to be defined.

Objective functions are formulated using the following terminology:

o A network comprises a set of N links {Li, (i=1...N)}.

- A path P is a list of K links {Lpi,(i=1...K)}.
- Bandwidth Utilization on link L is denoted u(L).
- Reserved Bandwidth Utilization on link L is denoted ru(L).
- Maximum bandwidth on link L is denoted M(L).
- Current Reserved bandwidth on link L is denoted c(L).

The description of the two new objective functions is as follows.

Objective Function Code: TBD

Name: Maximum Under-Utilized Path (MUP)

Description: Find a path P such that (Min {(M(Lpi)- u(Lpi)) / M(Lpi), i=1...K }) is maximized.

Objective Function Code: TBD

Name: Maximum Reserved Under-Utilized Path (MRUP)

Description: Find a path P such that (Min {(c(Lpi)- ru(Lpi)) / c(Lpi), i=1...K }) is maximized.

These new objective function are used to optimize paths based on bandwidth utilization as the optimization criteria.

If the objective function defined in this document are unknown/ unsupported, the procedure as defined in [RFC5541] is followed.

6.3. PCEP Message

The new optional BU objects MAY be specified in the PCReq message. As per [RFC5541], an OF object specifying a new objective function MAY also be specified.

The format of the PCReq message (with [RFC5541] as a base) is updated as follows:

```
<PCReq Message> ::= <Common Header>
                    [<svec-list>]
                    <request-list>
where:
     <svec-list> ::= <SVEC>
                     [<OF>]
                     [<metric-list>]
                     [<svec-list>]

     <request-list> ::= <request> [<request-list>]

     <request> ::= <RP>
                   <END-POINTS>
                   [<LSPA>]
                   [<BANDWIDTH>]
                   [<bu-list>]
                   [<metric-list>]
                   [<OF>]
                   [<RRO>[<BANDWIDTH>]]
                   [<IRO>]
                   [<LOAD-BALANCING>]

and where:
     <bu-list>::=<BU>[<bu-list>]
     <metric-list> ::= <METRIC>[<metric-list>]
```

The BU objects MAY be specified in the PCRep message, in case of an unsuccessful path computation to indicate the bandwidth utilization as a reason for failure. The OF object MAY be carried within a PCRep message to indicate the objective function used by the PCE during path computation.

The format of the PCRep message (with [RFC5541] as a base) is updated as follows:

```
<PCRep Message> ::= <Common Header>
                    [<svec-list>]
                    <response-list>

where:

     <svec-list> ::= <SVEC>
                     [<OF>]
                     [<metric-list>]
                     [<svec-list>]

    <response-list> ::= <response> [<response-list>]

    <response> ::= <RP>
                   [<NO-PATH>]
                   [<attribute-list>]
                   [<path-list>]

    <path-list> ::= <path> [<path-list>]

    <path> ::= <ERO>
               <attribute-list>

and where:

    <attribute-list> ::= [<OF>]
                         [<LSPA>]
                         [<BANDWIDTH>]
                         [<bu-list>]
                         [<metric-list>]
                         [<IRO>]

        <bu-list>::=<BU>[<bu-list>]
    <metric-list> ::= <METRIC> [<metric-list>]
```

7. Other Considerations

7.1. Reoptimization Consideration

PCC can monitor the link bandwidth utilization of the setup LSPs and in case of drastic change, it MAY ask PCE for reoptimization as per [RFC5440].

7.2. Inter-domain Consideration

7.3. P2MP Consideration

7.4. Stateful PCE

8. IANA Considerations

TBD

9. Security Considerations

TBD

10. Security Considerations

This document defines a new BU object and OF codes which does not add any new security concerns beyond those discussed in [RFC5440].

11. Manageability Considerations

11.1. Control of Function and Policy

The only configurable item is the support of the new constraints on a PCE which MAY be controlled by a policy module. If the new constraints are not supported/allowed on a PCE, it MUST send a PCErr message as specified in Section 6.1.1.

11.2. Information and Data Models

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

11.3. 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].

11.4. Verify Correct Operations

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

11.5. Requirements On Other Protocols

PCE requires the TED to be populated with the bandwidth utilization. This mechanism is described in [OSPF-TE-EXPRESS] or [ISIS-TE-EXPRESS].

11.6. 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].

12. Acknowledgments

We would like to thank Alia Atlas, John E Drake, David Ward for their useful comments and suggestions.

13. References

13.1. Normative References

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

13.2. Informative References

[RFC3630] Katz, D., Kompella, K., and D. Yeung, "Traffic Engineering (TE) Extensions to OSPF Version 2", RFC 3630, September 2003.

[RFC3784] Smit, H. and T. Li, "Intermediate System to Intermediate System (IS-IS) Extensions for Traffic Engineering (TE)", RFC 3784, June 2004.

[RFC5440] Vasseur, JP. and JL. Le Roux, "Path Computation Element (PCE) Communication Protocol (PCEP)", RFC 5440, March 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.

[OSPF-TE-EXPRESS] Giacalone, S., Ward, D., Drake, J., Atlas, A., and S. Previdi, "OSPF Traffic Engineering (TE) Metric Extensions [draft-ietf-ospf-te-metric-extensions]", June 2013.

[ISIS-TE-EXPRESS] Previdi, S., Giacalone, S., Ward, D., Drake, J., Atlas, A., Filsfils, C., and W. Qin, "IS-IS Traffic Engineering (TE) Metric Extensions [draft-ietf-isis-te-metric-extensions-01]", October 2013.

[PCEP-MIB] Kiran Koushik, A S., Stephan, E., Zhao, Q., King, D., and J. Hardwick, "PCE communication protocol(PCEP) Management Information Base [draft-ietf-pce-pcep-mib]", Feb 2013.

Appendix A. Contributor Addresses

Udayasree Palle
Huawei Technologies
Leela Palace
Bangalore, Karnataka 560008
INDIA
EMail: udayasree.palle@huawei.com

Authors' Addresses

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

EMail: sunseawq@huawei.com

Dhruv Dhody
Huawei Technologies
Leela Palace
Bangalore, Karnataka 560008
INDIA

EMail: dhruv.ietf@gmail.com

Stefano Previdi
Cisco Systems, Inc
Via Del Serafico 200
Rome 00191
IT

EMail: sprevidi@cisco.com