

PCE Working Group
Internet-Draft
Intended status: Standards Track
Expires: October 25, 2019

D. Dhody
U. Palle
Huawei Technologies
R. Singh
Juniper Networks
R. Gandhi
Cisco Systems, Inc.
L. Fang
Expedia, Inc.
April 23, 2019

**PCEP Extensions for MPLS-TE LSP Automatic Bandwidth Adjustment with
Stateful PCE
draft-ietf-pce-stateful-pce-auto-bandwidth-09**

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. The Stateful PCE extensions allow stateful control of Multi-Protocol Label Switching (MPLS) Traffic Engineering Label Switched Paths (TE LSPs) using PCEP.

The automatic bandwidth feature allows automatic and dynamic adjustment of the TE LSP bandwidth reservation based on the volume of traffic flowing through the LSP. This document describes PCEP extensions for automatic bandwidth adjustment when employing an Active Stateful PCE for both PCE-Initiated and PCC-Initiated LSPs.

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."

Copyright Notice

Copyright (c) 2019 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to [BCP 78](http://trustee.ietf.org/license-info) 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 | 3 |
| 2. | Conventions Used in This Document | 4 |
| 2.1. | Requirements Language | 4 |
| 2.2. | Abbreviations | 4 |
| 2.3. | Terminology | 5 |
| 3. | Requirements for PCEP Extensions | 7 |
| 4. | Architectural Overview | 8 |
| 4.1. | Auto-Bandwidth Overview | 8 |
| 4.2. | Auto-bandwidth Theory of Operation | 8 |
| 4.3. | Scaling Considerations | 10 |
| 5. | Extensions to the PCEP | 10 |
| 5.1. | Capability Advertisement | 10 |
| 5.1.1. | AUTO-BANDWIDTH-CAPABILITY TLV | 11 |
| 5.2. | AUTO-BANDWIDTH-ATTRIBUTES TLV | 11 |
| 5.2.1. | Sample-Interval sub-TLV | 13 |
| 5.2.2. | Adjustment Intervals | 13 |
| 5.2.2.1. | Adjustment-Interval sub-TLV | 13 |
| 5.2.2.2. | Down-Adjustment-Interval sub-TLV | 14 |
| 5.2.3. | Adjustment Thresholds | 15 |
| 5.2.3.1. | Adjustment-Threshold sub-TLV | 15 |
| 5.2.3.2. | Adjustment-Threshold-Percentage sub-TLV | 16 |
| 5.2.3.3. | Down-Adjustment-Threshold sub-TLV | 16 |
| 5.2.3.4. | Down-Adjustment-Threshold-Percentage sub-TLV | 17 |
| 5.2.4. | Minimum and Maximum Bandwidth Values | 18 |
| 5.2.4.1. | Minimum-Bandwidth sub-TLV | 18 |
| 5.2.4.2. | Maximum-Bandwidth sub-TLV | 18 |
| 5.2.5. | Overflow and Underflow Conditions | 19 |
| 5.2.5.1. | Overflow-Threshold sub-TLV | 19 |

| | | |
|----------|--|----|
| 5.2.5.2. | Overflow-Threshold-Percentage sub-TLV | 20 |
| 5.2.5.3. | Underflow-Threshold sub-TLV | 21 |
| 5.2.5.4. | Underflow-Threshold-Percentage sub-TLV | 21 |
| 5.3. | BANDWIDTH Object | 22 |
| 5.4. | The PCInitiate Message | 23 |
| 5.5. | The PCUpd Message | 23 |
| 5.6. | The PCRpt Message | 23 |
| 5.7. | The PCNtf Message | 24 |
| 6. | Security Considerations | 24 |
| 7. | Manageability Considerations | 25 |
| 7.1. | Control of Function and Policy | 25 |
| 7.2. | Information and Data Models | 25 |
| 7.3. | Liveness Detection and Monitoring | 25 |
| 7.4. | Verify Correct Operations | 25 |
| 7.5. | Requirements On Other Protocols | 25 |
| 7.6. | Impact On Network Operations | 25 |
| 8. | IANA Considerations | 26 |
| 8.1. | PCEP TLV Type Indicators | 26 |
| 8.2. | AUTO-BANDWIDTH-CAPABILITY TLV Flag Field | 26 |
| 8.3. | AUTO-BANDWIDTH-ATTRIBUTES Sub-TLV | 26 |
| 8.4. | Error Object | 27 |
| 8.5. | Notification Object | 27 |
| 9. | References | 28 |
| 9.1. | Normative References | 28 |
| 9.2. | Informative References | 28 |
| | Acknowledgments | 30 |
| | Contributors' Addresses | 30 |
| | Authors' Addresses | 31 |

1. Introduction

[RFC5440] describes the Path Computation Element Protocol (PCEP) as a communication mechanism between a Path Computation Client (PCC) and a Path Control Element (PCE), or between PCE and PCE, that enables computation of Multi-Protocol Label Switching (MPLS) Traffic Engineering Label Switched Paths (TE LSPs).

[RFC8231] specifies extensions to PCEP to enable stateful control of MPLS TE LSPs. It describes two mode of operations - Passive stateful PCE and Active stateful PCE. Further, [RFC8281] describes the setup, maintenance and teardown of PCE-Initiated LSPs for the stateful PCE model. In this document, the focus is on Active stateful PCE where the LSPs are controlled by the PCE.

Over time, based on the varying traffic pattern, an LSP established with a certain bandwidth may require to adjust the bandwidth reserved in the network dynamically. The head-end Label Switch Router (LSR) monitors the actual bandwidth demand of the established LSP and periodically computes new bandwidth. The head-end LSR adjusts the bandwidth reservation of the LSP based on the computed bandwidth automatically. This feature is commonly referred to as Auto-Bandwidth. The Auto-Bandwidth feature is described in detail in [Section 4](#) of this document.

In the model considered in this document, the PCC (head-end of the LSP) collects the traffic rate samples flowing through the LSP and calculates the new adjusted bandwidth. The PCC reports the calculated bandwidth to be adjusted to the PCE. This is similar to the Passive stateful PCE model, while the Passive stateful PCE uses path request/reply mechanism, the Active stateful PCE uses report/update mechanism. In case of PCE-Initiated LSP, the PCC is requested during the LSP initiation to monitor and calculate the new adjusted bandwidth. [\[RFC8051\]](#) describes the use-case for Auto-Bandwidth adjustment for Passive and Active stateful PCE.

This document defines the PCEP extensions needed to support Auto-Bandwidth feature in a Active stateful PCE model where the LSP bandwidth to be adjusted is calculated on the PCC (head-end of the LSP). The use of PCE to calculate the bandwidth to be adjusted is out of scope of this document.

[2.](#) Conventions Used in This Document

[2.1.](#) Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [BCP 14](#) [\[RFC2119\]](#) [\[RFC8174\]](#) when, and only when, they appear in all capitals, as shown here.

[2.2.](#) Abbreviations

PCC: Path Computation Client.

PCE: Path Computation Element.

PCEP: Path Computation Element Communication Protocol.

TE LSP: Traffic Engineering Label Switched Path.

2.3. Terminology

The reader is assumed to be familiar with the terminology defined in [\[RFC5440\]](#), [\[RFC8231\]](#), and [\[RFC8281\]](#).

The following auto-bandwidth terminology is defined in this document.

Maximum Average Bandwidth (MaxAvgBw): The maximum average bandwidth represents the current traffic bandwidth demand of the LSP during a time interval. This is the maximum value of the traffic bandwidth rate samples (Bandwidth-Samples) in a given time interval.

Adjusted Bandwidth: This is the Auto-Bandwidth computed bandwidth that is used to adjust the bandwidth reservation of the LSP.

Sample-Interval: The periodic time interval at which the measured traffic rate of the LSP is collected as a Bandwidth-Sample.

Bandwidth-Sample: The bandwidth sample of the measured traffic rate of the LSP collected at every Sample-Interval.

Maximum-Bandwidth: The maximum bandwidth that can be reserved for the LSP.

Minimum-Bandwidth: The minimum bandwidth that can be reserved for the LSP.

Up-Adjustment-Interval: The periodic time interval at which the bandwidth adjustment should be made using the MaxAvgBw, when MaxAvgBw is greater than the current bandwidth reservation of the LSP.

Down-Adjustment-Interval: The periodic time interval at which the bandwidth adjustment should be made using the MaxAvgBw, when MaxAvgBw is less than the current bandwidth reservation of the LSP.

Up-Adjustment-Threshold: This parameter is used to decide when the LSP bandwidth should be adjusted. If the percentage or absolute difference between the current MaxAvgBw and the current bandwidth reservation is greater than or equal to the threshold value, the LSP bandwidth is adjusted (upsized) to the current bandwidth demand (Adjusted Bandwidth) at the Up-Adjustment-Interval expiry.

Down-Adjustment-Threshold: This parameter is used to decide when the LSP bandwidth should be adjusted. If the percentage or absolute difference between the current bandwidth reservation and the

current MaxAvgBw is greater than or equal to the threshold value, the LSP bandwidth is adjusted (downsized) to the current bandwidth demand (Adjusted Bandwidth) at the Down-Adjustment-Interval expiry.

Overflow-Count: This parameter is used to decide when the LSP bandwidth should be adjusted when there is a sudden increase in traffic demand. This value indicates how many times consecutively, the percentage or absolute difference between the current MaxAvgBw and the current bandwidth reservation of the LSP is greater than or equal to the Overflow-Threshold value.

Overflow-Threshold: This parameter is used to decide when the LSP bandwidth should be adjusted when there is a sudden increase in traffic demand. If the percentage or absolute difference between the current MaxAvgBw and the current bandwidth reservation of the LSP is greater than or equal to the threshold value, the overflow condition is set to be met. The LSP bandwidth is adjusted to the current bandwidth demand bypassing the Up-Adjustment-Interval if the overflow condition is met consecutively for the Overflow-Count.

Underflow-Count: This parameter is used to decide when the LSP bandwidth should be adjusted when there is a sudden decrease in traffic demand. This value indicates how many times consecutively, the percentage or absolute difference between the current MaxAvgBw and the current bandwidth reservation of the LSP is greater than or equal to the Underflow-Threshold value.

Underflow-Threshold: This parameter is used to decide when the LSP bandwidth should be adjusted when there is a sudden decrease in traffic demand. If the percentage or absolute difference between the current MaxAvgBw and the current bandwidth reservation of the LSP is greater than or equal to the threshold value, the underflow condition is set to be met. The LSP bandwidth is adjusted to the current bandwidth demand bypassing the Down-Adjustment-Interval if the underflow condition is met consecutively for the Underflow-Count.

Minimum-Threshold: The increase or decrease of the LSP bandwidth should be at least or above the minimum-threshold represented as an absolute bandwidth value before the bandwidth adjustment for the LSP is made. This threshold can be seen as a suppression threshold that is used along with a percentage threshold to avoid unnecessary auto-bandwidth adjustments and re-signaling of the LSP at low bandwidth values.

3. Requirements for PCEP Extensions

The PCEP extensions required for auto-bandwidth are summarized in the following table as well as in Figure 1.

| | |
|-----------------------------|-----------------------------|
| +-----+-----+ | |
| PCC Initiated | PCE Initiated |
| +-----+-----+ | |
| | |
| PCC monitors the traffic | At the time of initiation, |
| and reports the calculated | PCE request PCC to monitor |
| bandwidth to be adjusted | the traffic and report the |
| to the PCE. | calculated bandwidth to be |
| | adjusted to the PCE. |
| | |
| Extension is needed for PCC | Extension is needed for PCE |
| to pass on the adjustment | to pass on the adjustment |
| parameters at the time of | parameters at the time of |
| LSP Delegation. | LSP Initiation. |
| | |
| +-----+-----+ | |

Table 1: Requirements for Auto-Bandwidth PCEP extensions

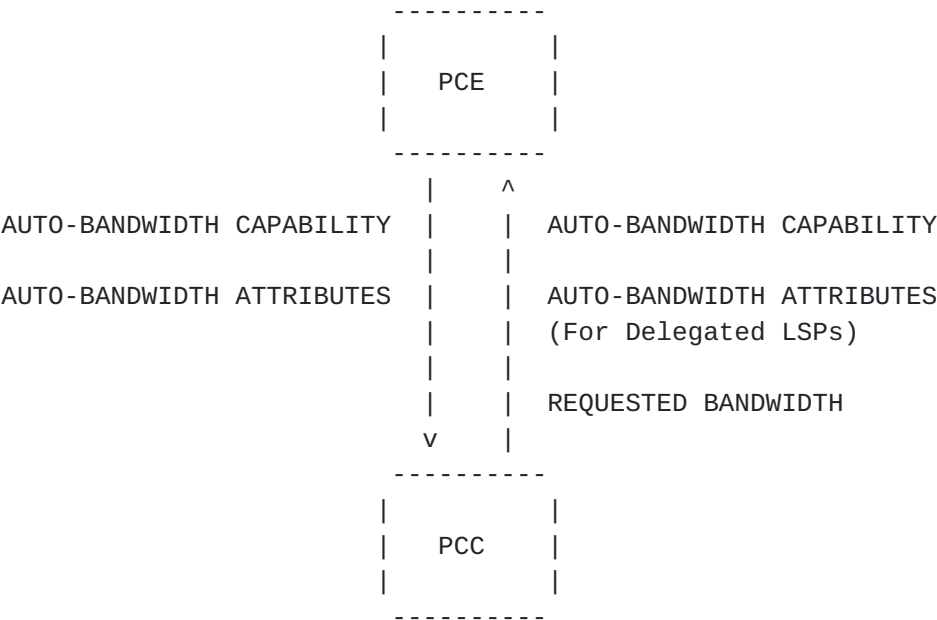


Figure 1: Overview of Auto-Bandwidth PCEP extensions

The PCEP speaker supporting this document must have a mechanism to advertise the automatic bandwidth adjustment capability for both PCC-

Initiated and PCE-Initiated LSPs.

Auto-bandwidth deployment considerations for PCEP extensions are summarized below:

- o It is required to identify and inform the PCC, which LSPs are enabled with Auto-Bandwidth feature. Not all LSPs in some deployments would like their bandwidth to be dependent on the real-time bandwidth usage but be constant as set by the operator.
- o In addition, an operator should be able to specify the auto-bandwidth adjustment parameters (i.e. configuration knobs) to control this feature (e.g. minimum/ maximum bandwidth range). The PCC should be informed about these adjustment parameters.

4. Architectural Overview

4.1. Auto-Bandwidth Overview

The Auto-Bandwidth feature allows automatic and dynamic adjustment of the reserved bandwidth of an LSP over time (i.e., without network operator intervention) to accommodate the varying traffic demand of the LSP. If the traffic flowing through the LSP is lower than the configured or current reserved bandwidth of the LSP, the extra bandwidth is being reserved needlessly and being wasted. Conversely, if the actual traffic flowing through the LSP is higher than the configured or current reserved bandwidth of the LSP, it can potentially cause congestion or packet loss in the network. The initial LSP bandwidth can be set to an arbitrary value (including zero), in practice, it can be operator expected value based on design and planning. The head-end Label Switch Router (LSR) monitors the actual traffic flowing through the LSP and uses that information to adjust the bandwidth reservation of the LSP in the network. Bandwidth adjustment must not cause disruption to the traffic flow carried by the LSP. One way to achieve this is to use the make-before-break (MBB) signaling method.

4.2. Auto-bandwidth Theory of Operation

This section describes the Auto-Bandwidth feature in a general way. When the Auto-Bandwidth feature is enabled, the measured traffic rate is periodically sampled at each Sample-Interval (which can be configured by an operator and the default value as 5 minutes) by the PCC, when the PCC is the head-end node of the LSP. The traffic rate samples are accumulated over the Adjustment-Interval period (in the Up or Down direction) (which can be configured by an operator and the

default value as 24 hours). The PCC, in-charge of calculating the bandwidth to be adjusted, can decide to adjust the bandwidth of the LSP to the highest traffic rate sample (MaxAvgBw) amongst the set of bandwidth samples collected over the adjustment-interval period (in the Up or Down direction) depending on the operator policy.

Note that the highest traffic rate sample could be higher or lower than the current LSP bandwidth. Only if the difference between the current bandwidth demand (MaxAvgBw) and the current bandwidth reservation is greater than or equal to the Adjustment-Threshold (percentage or absolute value) (which can be configured by an operator and the default as 5 percentage), the LSP bandwidth is adjusted (upsized) to the current bandwidth demand (MaxAvgBw). Similarly, if the difference between the current bandwidth reservation and the current bandwidth demand (MaxAvgBw) is greater than or equal to the Down-Adjustment-Threshold (percentage or absolute value), the LSP bandwidth is adjusted (downsized) to the current bandwidth demand (MaxAvgBw). Some LSPs are less eventful while other LSPs may encounter a lot of changes in the traffic pattern. The thresholds and intervals for bandwidth adjustment are configured based on the traffic pattern of the LSP.

In order to avoid frequent re-signaling, an operator may set a longer adjustment-interval value (Up and/or Down). However, longer adjustment-interval can result in an undesirable effect of masking sudden changes in traffic demands of an LSP. To avoid this, the Auto-Bandwidth feature may pre-maturely expire the adjustment-interval and adjust the LSP bandwidth to accommodate the sudden bursts of increase in traffic demand as an overflow condition or decrease in traffic demand as an underflow condition. An operator needs to configure appropriate values for the Overflow-Threshold and/or Underflow-Threshold parameters and they do not have default values defined in this document.

All thresholds in this document could be represented in both absolute value and percentage, and could be used together. This is provided to accommodate the cases where the LSP bandwidth reservation may become very large or very small over time. For example, an operator may use the percentage threshold to handle small to large bandwidth values and absolute values to handle very large bandwidth values. The auto-bandwidth adjustment is made when either one of the two thresholds, the absolute or percentage, is crossed.

When using the (adjustment/overflow/underflow) percentage thresholds, if the LSP bandwidth changes rapidly at very low values, it may trigger frequent auto-bandwidth adjustments due to the crossing of the percentage thresholds. This can lead to unnecessary re-signaling of the LSPs in the network. This is suppressed by setting the

minimum-threshold parameters along with the percentage thresholds. The auto-bandwidth adjustment is only made if the LSP bandwidth crosses both the percentage threshold and the minimum-threshold parameters.

4.3. Scaling Considerations

It should be noted that any bandwidth change requires re-signaling of an LSP, which can further trigger preemption of lower priority LSPs in the network. When deployed under scale, this can lead to a signaling churn in the network. The Auto-bandwidth application algorithm is thus advised to take this into consideration before adjusting the LSP bandwidth. Operators are advised to set the values of various auto-bandwidth adjustment parameters appropriate for the deployed LSP scale.

If a PCE gets overwhelmed, it can notify the PCC to temporarily suspend the reporting of the new LSP bandwidth to be adjusted (see [Section 5.7](#) of this document). Similarly, if a PCC gets overwhelmed due to signaling churn, it can notify the PCE to temporarily suspend new LSP setup requests.

5. Extensions to the PCEP

5.1. Capability Advertisement

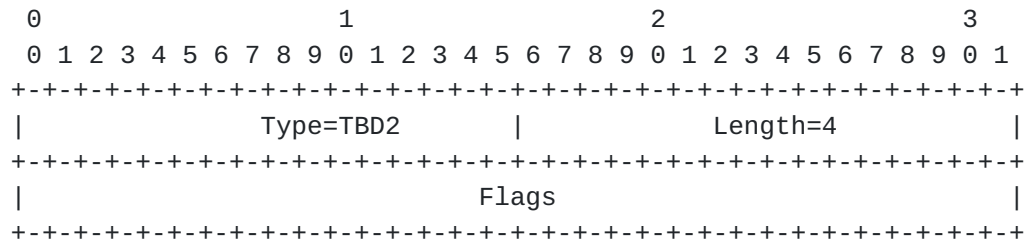
During PCEP Initialization Phase, PCEP speakers (PCE or PCC) advertise their support of Automatic Bandwidth adjustment feature. A PCEP speaker includes the "Auto-Bandwidth Capability" TLV, in the OPEN Object to advertise its support for PCEP Auto-Bandwidth extensions. The presence of the "Auto-Bandwidth Capability" TLV in the OPEN Object indicates that the Automatic Bandwidth feature is supported as described in this document.

- o The PCEP protocol extensions for Auto-Bandwidth adjustments MUST NOT be used if one or both PCEP speakers have not included the "Auto-Bandwidth Capability" TLV in their respective OPEN message.
- o The PCEP speaker that does not recognize the extensions defined in this document sends the PCErr message with error-type 2 (capability not supported) as per [Section 6.9 in \[RFC5440\]](#).
- o If the PCEP speaker that supports the extensions defined in this document but did not advertise this capability, then upon receipt of AUTO-BANDWIDTH-ATTRIBUTES TLV in the LSP Attributes (LSPA) object, it SHOULD generate a PCErr with error-type 19 (Invalid

Operation), error-value TBD4 (Auto-Bandwidth capability was not advertised) and ignore the AUTO-BANDWIDTH-ATTRIBUTES TLV.

5.1.1. AUTO-BANDWIDTH-CAPABILITY TLV

The AUTO-BANDWIDTH-CAPABILITY TLV is an optional TLV for use in the OPEN Object for Automatic Bandwidth Adjustment via PCEP capability advertisement. Its format is shown in the following figure:



AUTO-BANDWIDTH-CAPABILITY TLV format

The Type of the TLV is (TBD2) and it has a fixed Length of 4 octets.

The value comprises a single field - Flags (32 bits). No flags are defined for this TLV in this document.

Unassigned bits are considered reserved. They MUST be set to 0 on transmission and MUST be ignored on receipt.

Advertisement of the Auto-Bandwidth capability TLV implies support of auto-bandwidth adjustment, as well as the objects, TLVs and procedures defined in this document.

5.2. AUTO-BANDWIDTH-ATTRIBUTES TLV

The AUTO-BANDWIDTH-ATTRIBUTES TLV provides the 'configurable knobs' of the feature and it can be included as an optional TLV in the LSPA Object (as described in [[RFC5440](#)]).

For PCE-Initiated LSP [[RFC8281](#)], this TLV is included in the LSPA Object with the PCInitiate message. For the PCC-Initiated delegated LSPs, this TLV is carried in the PCRpt message in LSPA Object. This TLV is also carried in the LSPA object with the PCUpd message to direct the PCC (LSP head-end) to make updates to auto-bandwidth attributes such as Adjustment-Interval.

The TLV is encoded in all PCEP messages for the LSP while the auto-bandwidth adjustment feature is enabled, the absence of the TLV indicates the PCEP speaker wish to disable the feature. This TLV

Future specification can define additional sub-TLVs.

5.2.2.1. Adjustment-Interval sub-TLV

0 Down-Adjustment-Interval: The 4-octet time interval for downward bandwidth adjustments. The valid range is from 1 to 604800, in seconds. The default value equals the adjustment-interval. The down-adjustment-interval parameter MUST NOT be less than the sample-interval, otherwise the Sub-TLV MUST be ignored and the previous value is maintained.

5.2.3. Adjustment Thresholds

The sub-TLVs in this section are encoded to inform the PCEP peer of the adjustment threshold parameters. An implementation MAY include both sub-TLVs for the absolute value and the percentage, in which case the bandwidth is adjusted when either of the adjustment threshold conditions are met. The Adjustment-Threshold sub-TLV specifies the threshold for both upward and downward trend. If the operator would like to use a different adjustment threshold during the downward trend, the Down-Adjustment-Threshold sub-TLV is included. Similarly, the Adjustment-Threshold-Percentage sub-TLV specifies the threshold percentage for both upward and downward trend. If the operator would like to use a different adjustment threshold percentage during the downward trend, the Down-Adjustment-Threshold-Percentage sub-TLV is included. It is worth noting that regardless of how the threshold are set, the adjustment will not be made until at least one sample-interval simply because no sample will be made on which to base a comparison with a threshold.

5.2.3.1. Adjustment-Threshold sub-TLV

The Adjustment-Threshold sub-TLV is used to decide when the LSP bandwidth should be adjusted when MaxAvgBw is greater than or less than the current bandwidth reservation.

```

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=4           |           Length=4           |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|           Adjustment-Threshold           |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

Adjustment-Threshold sub-TLV format

The Type is 4, Length is 4 octets, and the value comprises of -

- o Adjustment-Threshold: The absolute Adjustment-Threshold bandwidth value, encoded in IEEE floating point format (see [\[IEEE.754.1985\]](#)), expressed in bytes per second. The default adjustment-threshold value is not set. Refer to [Section 3.1.2 of \[RFC3471\]](#) for a table of commonly used values.

If the modulus of difference between the current MaxAvgBw and the current bandwidth reservation is greater than or equal to the threshold value, the LSP bandwidth is adjusted to the current bandwidth demand (MaxAvgBw).




```

+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|           Type=6           |           Length=4           |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|           Down-Adjustment-Threshold           |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

Down-Adjustment-Threshold sub-TLV format

The Type is 6, Length is 4 octets, and the value comprises of -

- o Down-Adjustment-Threshold: The absolute Down-Adjustment-Threshold bandwidth value, encoded in IEEE floating point format (see [\[IEEE.754.1985\]](#)), expressed in bytes per second. The default value equals the adjustment-threshold. Refer to [Section 3.1.2 of \[RFC3471\]](#) for a table of commonly used values.

If the difference between current bandwidth reservation and the current MaxAvgBw is greater than or equal to the threshold value, the LSP bandwidth is adjusted to the current bandwidth demand (MaxAvgBw).

5.2.3.4. Down-Adjustment-Threshold-Percentage sub-TLV

The Down-Adjustment-Threshold-Percentage sub-TLV is used to decide when the LSP bandwidth should be adjusted when MaxAvgBw is lesser than the current bandwidth reservation. This parameter overwrites the Adjustment-Threshold-Percentage for the downward trend.

```

      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=7           |           Length=8           |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|           Reserved           | Percentage |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|           Minimum-Threshold           |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

Down-Adjustment-Threshold-Percentage sub-TLV format

The Type is 7, Length is 8 octets, and the value comprises of -

- o Reserved: SHOULD be set to zero on transmission and MUST be ignored on receipt.
- o Percentage: The Down-Adjustment-Threshold value (7 bits), encoded in percentage (an integer from 1 to 100). The value 0 is considered to be invalid. The default value equals the adjustment-threshold-percentage.

| 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 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| + | - | + | - | + | - | + | - | + | - | + | - | + | - | + | - | + | - | + | - | + | - | + | - | + | - | + | - | + | - | + | - | + | - | + | - | + | - | + | - |


```

|          Type=9          |          Length=4          |
+-+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+
|          Maximum-Bandwidth          |
+-+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+

```

Maximum-Bandwidth sub-TLV format

The Type is 9, Length is 4 octets, and the value comprises of -

- o Maximum-Bandwidth: The 4-octet bandwidth value encoded in IEEE floating point format (see [[IEEE.754.1985](#)]), expressed in bytes per second. The default maximum-bandwidth value is not set. Refer to [Section 3.1.2 of \[RFC3471\]](#) for a table of commonly used values.

5.2.5. Overflow and Underflow Conditions

The sub-TLVs in this section are encoded to inform the PCEP peer the overflow and underflow threshold parameters. An implementation MAY include sub-TLVs for an absolute value and/or a percentage for the threshold, in which case the bandwidth is immediately adjusted when either of the threshold conditions is met consecutively for the given count (as long as the difference in the bandwidth is at least or above the Minimum-Threshold). By default, the threshold values for overflow and underflow conditions are not set.

5.2.5.1. Overflow-Threshold sub-TLV

The Overflow-Threshold sub-TLV is used to decide if the LSP bandwidth should be adjusted immediately.

```

      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=10          |          Length=8          |
+-+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+
|          Reserved          |          Count          |
+-+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+
|          Overflow-Threshold          |
+-+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+

```

Overflow-Threshold sub-TLV format

The Type is 10, Length is 8 octets, and the value comprises of -

- o Reserved: SHOULD be set to zero on transmission and MUST be ignored on receipt.

- o Percentage: The Overflow-Threshold value (7 bits), encoded in percentage (an integer from 1 to 100). The value 0 is considered to be invalid. If the percentage increase of the current MaxAvgBw from the current bandwidth reservation is greater than or equal to the threshold percentage, the overflow condition is met.
- o Reserved: SHOULD be set to zero on transmission and MUST be ignored on receipt.
- o Count: The Overflow-Count value (5 bits), encoded in integer. The value 0 is considered to be invalid. The number of consecutive samples for which the overflow condition MUST be met for the LSP bandwidth to be immediately adjusted to the current bandwidth demand, bypassing the (up) adjustment-interval.




```

 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=13           |           Length=8           |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| Percentage |           Reserved           | Count |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|           Minimum-Threshold           |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

Underflow-Threshold-Percentage sub-TLV format

The Type is 13, Length is 8 octets, and the value comprises of -

- o Percentage: The Underflow-Threshold value (7 bits), encoded in percentage (an integer from 1 to 100). The value 0 is considered to be invalid. If the percentage decrease of the current MaxAvgBw from the current bandwidth reservation is greater than or equal to the threshold percentage, the underflow condition is met.
- o Reserved: SHOULD be set to zero on transmission and MUST be ignored on receipt.
- o Count: The Underflow-Count value (5 bits), encoded in integer. The value 0 is considered to be invalid. The number of consecutive samples for which the underflow condition MUST be met for the LSP bandwidth to be immediately adjusted to the current bandwidth demand, bypassing the down-adjustment-interval.
- o Minimum-Threshold: The absolute Minimum-Threshold bandwidth value, encoded in IEEE floating point format (see [[IEEE.754.1985](#)]), expressed in bytes per second. The decrease of the LSP bandwidth should be at least or above the minimum-threshold before the bandwidth adjustment is made.

5.3. BANDWIDTH Object

As per [[RFC5440](#)], the BANDWIDTH object (Object-Class value 5) is defined with two Object-Type values as following:

- o Requested Bandwidth: BANDWIDTH Object-Type value is 1.
- o Re-optimization Bandwidth: Bandwidth of an existing TE LSP for which a re-optimization is requested. BANDWIDTH Object-Type value is 2.

The PCC reports the calculated bandwidth to be adjusted (MaxAvgBw) to the Stateful PCE using the existing 'Requested Bandwidth' with BANDWIDTH Object-Type as 1. The reporting of the 're-optimization

bandwidth' with BANDWIDTH Object-Type as 2 is not required as the Stateful PCE is aware of the existing LSP bandwidth.

5.4. The PCInitiate Message

A PCInitiate message is a PCEP message sent by a PCE to a PCC to trigger LSP instantiation or deletion [[RFC8281](#)].

For the PCE-Initiated LSP with Auto-Bandwidth feature enabled, AUTO-BANDWIDTH-ATTRIBUTES TLV MUST be included in the LSPA object with the PCInitiate message.

The Routing Backus-Naur Format (RBNF) definition of the PCInitiate message [[RFC8281](#)] is unchanged by this document.

5.5. The PCUpd Message

A PCUpd message is a PCEP message sent by a PCE to a PCC to update the LSP parameters [[RFC8231](#)].

For PCE-Initiated LSPs with Auto-Bandwidth feature enabled, AUTO-BANDWIDTH-ATTRIBUTES TLV MUST be included in the LSPA object with the PCUpd message. The PCE can send this TLV to direct the PCC to change the auto-bandwidth parameters.

The RBNF definition of the PCUpd message [[RFC8231](#)] is unchanged by this document.

5.6. The PCRpt Message

The PCRpt message [[RFC8231](#)] is a PCEP message sent by a PCC to a PCE to report the status of one or more LSPs.

For PCE-Initiated LSPs [[RFC8281](#)], the PCC creates the LSP using the attributes communicated by the PCE, and using the local values for the unspecified parameters. After the successful instantiation of the LSP, PCC automatically delegates the LSP to the PCE and generates a PCRpt message to provide the status report for the LSP.

For both PCE-Initiated and PCC-Initiated LSPs, when the LSP is delegated to a PCE for the very first time as well as after the successful delegation, the BANDWIDTH object of type 1 is used to specify the requested bandwidth in the PCRpt message.

For all LSPs with Auto-Bandwidth feature enabled, AUTO-BANDWIDTH-ATTRIBUTES TLV MUST be included in the LSPA object of the PCRpt message.

The RBNF definition of the PCRpt message [[RFC8231](#)] is unchanged by this document.

5.7. The PCNtf Message

As per [[RFC5440](#)], the PCEP Notification message (PCNtf) can be sent by a PCEP speaker to notify its peer of a specific event.

A PCEP speaker (PCE or PCC) SHOULD notify its PCEP peer (PCC or PCE) when it is in overwhelmed state due to the auto-bandwidth feature. Upon receipt of such notification, the peer SHOULD NOT send any PCEP messages related to auto-bandwidth adjustment. If a PCEP message related to auto-bandwidth is received during in overwhelmed state, it MUST be ignored.

- o When a PCEP speaker is overwhelmed, it SHOULD notify its peer by sending a PCNtf message with Notification Type = TBD3 (Auto-bandwidth Overwhelm State) and Notification Value = 1 (Entering auto-bandwidth overwhelm state). Optionally, OVERLOADED-DURATION TLV [[RFC5440](#)] MAY be included that specifies the time period during which no further PCEP messages related to auto-bandwidth adjustment should be sent.
- o When the PCEP speaker is no longer in the overwhelm state and is available to process the auto-bandwidth adjustments, it SHOULD notify its peer by sending a PCNtf message with Notification Type = TBD3 (Auto-bandwidth Overwhelm State) and Notification Value = 2 (Clearing auto-bandwidth overwhelm state).

When Auto-Bandwidth feature is deployed, a PCE can send this notification to PCC when a PCC is reporting frequent auto-bandwidth adjustments. If a PCC is overwhelmed with re-signaling, it can also notify the PCE to not adjust the LSP bandwidth while in overwhelm state.

6. Security Considerations

This document defines AUTO-BANDWIDTH-CAPABILITY TLV and AUTO-BANDWIDTH-ATTRIBUTES sub-TLVs which do not add any new security concerns beyond those already discussed in [[RFC8231](#)] and [[RFC8281](#)] for stateful PCE operations. As per [[RFC8231](#)], it is RECOMMENDED that these PCEP extensions only be activated on authenticated and encrypted sessions across PCEs and PCCs belonging to the same administrative authority, using Transport Layer Security (TLS) [[RFC8253](#)], as per the recommendations and best current practices in [[RFC7525](#)] (unless explicitly set aside in [[RFC8253](#)]).

7. Manageability Considerations

7.1. Control of Function and Policy

The Auto-Bandwidth feature SHOULD be controlled per LSP (at PCC (head-end of the LSP) or PCE) and the values for auto-bandwidth parameters e.g. sample-interval, adjustment-interval (up/down), minimum-bandwidth, maximum-bandwidth, adjustment-threshold (up/down) SHOULD be configurable by an operator.

7.2. Information and Data Models

A Management Information Base (MIB) module for modeling PCEP is described in [[RFC7420](#)]. However, one may prefer the mechanism for configuration using YANG data model [[I-D.ietf-pce-pcep-yang](#)]. These SHOULD be enhanced to provide controls and indicators for support of auto-bandwidth feature. Support for various configuration knobs as well as counters of messages sent/received containing the TLVs defined in this document SHOULD be added.

7.3. Liveness Detection and Monitoring

The mechanisms defined in this document do not imply any new liveness detection and monitoring requirements in addition to those already listed in [[RFC5440](#)].

7.4. Verify Correct Operations

The mechanisms defined in this document do not imply any new operation verification requirements in addition to those already listed in [[RFC5440](#)].

7.5. Requirements On Other Protocols

The mechanisms defined in this document do not add any new requirements on other protocols.

7.6. Impact On Network Operations

In order to avoid any unacceptable impact on network operations, an implementation SHOULD allow a limit to be placed on the number of LSPs that can be enabled with auto-bandwidth feature. An implementation MAY allow a limit to be placed on the rate of auto-bandwidth related messages sent by a PCEP speaker and received by a peer. An implementation MAY also allow sending a notification when a PCEP speaker is overwhelmed or the rate of messages reach a threshold.

8. IANA Considerations

8.1. PCEP TLV Type Indicators

This document defines the following new PCEP TLVs; IANA is requested to make the following allocations from the "PCEP TLV Type Indicators" sub-registry of the PCEP Numbers registry, as follows:

| Value | Name | Reference |
|-------|---------------------------|-----------------|
| ----- | | |
| TBD2 | AUTO-BANDWIDTH-CAPABILITY | [This document] |
| TBD1 | AUTO-BANDWIDTH-ATTRIBUTES | [This document] |

8.2. AUTO-BANDWIDTH-CAPABILITY TLV Flag Field

IANA is requested to create a sub-registry to manage the Flag field of the AUTO-BANDWIDTH-CAPABILITY TLV.

New bit numbers are to be assigned by Standards Action [[RFC8126](#)]. Each bit should be tracked with the following qualities:

- o Bit number (counting from bit 0 as the most significant bit)
- o Capability description
- o Defining RFC

There is no bit defined for the AUTO-BANDWIDTH-CAPABILITY TLV Object flag field in this document.

8.3. AUTO-BANDWIDTH-ATTRIBUTES Sub-TLV

This document specifies the AUTO-BANDWIDTH-ATTRIBUTES Sub-TLVs. IANA is requested to create an "AUTO-BANDWIDTH-ATTRIBUTES Sub-TLV Types" sub-registry within the "Path Computation Element Protocol (PCEP) Numbers" registry to manage the type indicator space for sub-TLVs of the AUTO-BANDWIDTH-ATTRIBUTES TLV. The valid range of values in the registry is 0-65535. IANA is requested to initialize the registry with the following values. All other values in the registry should be marked as "Unassigned".

IANA is requested to set the registration procedure for this registry to read as follows:

| | |
|-------------|------------------|
| 0-65503 | IETF Review |
| 65504-65535 | Experimental Use |

This document defines the following types:

| Type | Name | Reference |
|-------|--|-----------------|
| 0 | Reserved | [This document] |
| 1 | Sample-Interval sub-TLV | [This document] |
| 2 | Adjustment-Interval sub-TLV | [This document] |
| 3 | Down-Adjustment-Interval sub-TLV | [This document] |
| 4 | Adjustment-Threshold sub-TLV | [This document] |
| 5 | Adjustment-Threshold-Percentage sub-TLV | [This document] |
| 6 | Down-Adjustment-Threshold sub-TLV | [This document] |
| 7 | Down-Adjustment-Threshold-Percentage sub-TLV | [This document] |
| 8 | Minimum-Bandwidth sub-TLV | [This document] |
| 9 | Maximum-Bandwidth sub-TLV | [This document] |
| 10 | Overflow-Threshold sub-TLV | [This document] |
| 11 | Overflow-Threshold-Percentage sub-TLV | [This document] |
| 12 | Underflow-Threshold sub-TLV | [This document] |
| 13 | Underflow-Threshold-Percentage sub-TLV | [This document] |
| 14- | Unassigned | [This document] |
| 65503 | | |

8.4. Error Object

This document defines a new Error-Value for PCErr message of Error-Type 19 (Invalid Operation) [[RFC8231](#)]. IANA is requested to allocate new error-value within the "PCEP-ERROR Object Error Types and Values" subregistry of the PCEP Numbers registry, as follows:

| Error-Type | Meaning & error values | Reference |
|------------|--|-----------------|
| 19 | Invalid Operations | |
| | Error-Value = TBD4: Auto-Bandwidth Capability was not Advertised | [This document] |

8.5. Notification Object

IANA is requested to allocate new Notification Type and Notification Values within the "Notification Object" sub-registry of the PCEP Numbers registry, as follows:

| Type | Meaning | Reference |
|------|--|-----------------|
| TBD3 | Auto-Bandwidth Overwhelm State | [This document] |
| | Notification-value=1: Entering Auto-Bandwidth overwhelm state | |
| | Notification-value=2: Clearing Auto-Bandwidth overwhelm state | |

9. References

9.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", [BCP 14](#), [RFC 2119](#), March 1997.
- [RFC5440] Vasseur, JP. and JL. Le Roux, "Path Computation Element (PCE) Communication Protocol (PCEP)", [RFC 5440](#), March 2009.
- [RFC8126] Cotton, M., Leiba, B., and T. Narten, "Guidelines for Writing an IANA Considerations Section in RFCs", [BCP 26](#), [RFC 8126](#), DOI 10.17487/RFC8126, June 2017, <<https://www.rfc-editor.org/info/rfc8126>>.
- [RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in [RFC 2119](#) Key Words", [BCP 14](#), [RFC 8174](#), DOI 10.17487/RFC8174, May 2017, <<https://www.rfc-editor.org/info/rfc8174>>.
- [RFC8231] Crabbe, E., Minei, I., Medved, J., and R. Varga, "Path Computation Element Communication Protocol (PCEP) Extensions for Stateful PCE", [RFC 8231](#), DOI 10.17487/RFC8231, September 2017, <<https://www.rfc-editor.org/info/rfc8231>>.
- [RFC8281] Crabbe, E., Minei, I., Sivabalan, S., and R. Varga, "Path Computation Element Communication Protocol (PCEP) Extensions for PCE-Initiated LSP Setup in a Stateful PCE, Model", [RFC 8281](#), DOI 10.17487/RFC8281, December 2017, <<https://www.rfc-editor.org/info/rfc8281>>.

9.2. Informative References

- [RFC3471] Berger, L., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Functional Description", [RFC 3471](#), January 2003.
- [RFC7420] Koushik, A., Stephan, E., Zhao, Q., King, D., and J. Hardwick, "Path Computation Element Communication Protocol (PCEP) Management Information Base (MIB) Module", [RFC 7420](#), December 2014.
- [RFC7525] Sheffer, Y., Holz, R. and P. Saint-Andre, "Recommendations for Secure Use of Transport Layer Security (TLS) and Datagram Transport Layer Security (DTLS)", [BCP 195](#), [RFC 7525](#), DOI 10.17487/RFC7525, May 2015.

- [RFC8051] Zhang, X. and I. Minei, "Applicability of a Stateful Path Computation Element (PCE)", [RFC 8051](#), January 2017.
- [RFC8253] Lopez, D., Dios, O., Wu, W., and D. Dhody, "PCEPS: Usage of TLS to Provide a Secure Transport for the Path Computation Element Communication Protocol (PCEP)", [RFC 8253](#), October 2017,
<<https://www.rfc-editor.org/info/rfc8253>>.
- [I-D.ietf-pce-pcep-yang] Dhody, D., Hardwick, J., Beeram, V., and J. Tantsura, "A YANG Data Model for Path Computation Element Communications Protocol (PCEP)", [draft-ietf-pce-pcep-yang](#) (work in progress).
- [IEEE.754.1985] Institute of Electrical and Electronics Engineers, "Standard for Binary Floating-Point Arithmetic", IEEE Standard 754, August 1985.

Acknowledgments

Authors would like to thank Robert Varga, Venugopal Reddy, Reeja Paul, Sandeep Boina, Avantika, JP Vasseur, Himanshu Shah and Adrian Farrel for their useful comments and suggestions.

Contributors' Addresses

He Zekun
Tencent Holdings Ltd,
Shenzhen P.R.China

Email: kinghe@tencent.com

Xian Zhang
Huawei Technologies
Research Area F3-1B,
Huawei Industrial Base,
Shenzhen, 518129
China

Phone: +86-755-28972645
Email: zhang.xian@huawei.com

Young Lee
Huawei Technologies
1700 Alma Drive, Suite 100
Plano, TX 75075
USA

Phone: +1 972 509 5599 x2240
Fax: +1 469 229 5397
Email: leeyoung@huawei.com

Authors' Addresses

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

Email: dhruv.ietf@gmail.com

Udayasree Palle
Huawei Technologies
Divyashree Techno Park, Whitefield
Bangalore, Karnataka 560037
India

Email: udayasreereddy@gmail.com

Ravi Singh
Juniper Networks
1194 N. Mathilda Ave.
Sunnyvale, CA 94089
USA

Email: ravis@juniper.net

Rakesh Gandhi
Cisco Systems, Inc.
Canada

Email: rgandhi@cisco.com

Luyuan Fang
Expedia, Inc.
USA

Email: luyuanf@gmail.com

