PCE Working Group
Internet-Draft
Intended status: Standards Track
Expires: October 19, 2017

D. Dhody
U. Palle
Huawei Technologies
R. Singh
Juniper Networks
R. Gandhi
Cisco Systems, Inc.
L. Fang
eBay

April 17, 2017

PCEP Extensions for MPLS-TE LSP Automatic Bandwidth Adjustment with Stateful PCE
draft-ietf-pce-stateful-pce-auto-bandwidth-04

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.

Automatic bandwidth 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.

Table of Contents

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).

[I-D.ietf-pce-stateful-pce] 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, [I-D.ietf-pce-pce-initiated-lsp] 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 certain bandwidth may require to adjust the bandwidth reserved in the network automatically. The head-end Label Switch Router (LSR) needs to monitor the actual bandwidth demand of the LSP and adjust the LSP bandwidth reservation periodically. This feature is commonly referred to as Auto-Bandwidth. Enabling Auto-Bandwidth feature on an LSP results in the head-end LSR automatically adjusting the LSP bandwidth reservation based on the traffic flowing through the LSP. 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. Once the LSP is set-up, the head-end monitors the traffic flow on the LSP and adjusts the bandwidth reservation periodically. The Auto-Bandwidth feature is described in detail in Section 4 of 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 a passive stateful PCE model, while the passive stateful PCE uses path request/reply mechanism, the active stateful PCE uses report/update mechanism to adjust the LSP bandwidth. 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.

The document [I-D.gandhi-pce-pm] describes the PCEP extensions for reporting the performance measurements to the PCE, and includes the real-time bandwidth utilization information of a TE LSP. Those extensions can be used to implement the auto-bandwidth feature on a stateful PCE, i.e. can be used to calculate the new bandwidth to be adjusted on the stateful PCE.

This document defines the extensions needed to support Auto-Bandwidth features on the LSPs in a active stateful PCE model using PCEP where the bandwidth to be adjusted is calculated on the PCC (head-end of the LSP).

2. Conventions Used in This Document

2.1. 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].

2.2. 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], [I-D.ietf-pce-pce-initiated-lsp], and [I-D.ietf-pce-stateful-pce].

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

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

Bandwidth-Sample: The bandwidth sample of the measured traffic rate 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.

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

Up-Adjustment-Threshold: This parameter is used to decide when the

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 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 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 is greater than or equal to the Overflow-Threshold value.

Overflow-Threshold: This parameter is used to decide when the 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 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 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 is greater than or equal to the Underflow-Threshold value.

Underflow-Threshold: This parameter is used to decide when the 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 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.

3. Requirements for PCEP Extensions

The PCEP extensions required for auto-bandwidth are summarized in the following table.

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

Table 1: 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:

- It is required to identify and inform the PCC, the LSP that 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.

- 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

Auto-Bandwidth feature allows automatic and dynamic adjustment of the reserved bandwidth of an LSP over time, i.e. without network operator

intervention. The new bandwidth reservation is determined by measuring the actual traffic flowing through the LSP. If the traffic flowing through the LSP is lower than the configured or current 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 bandwidth of the LSP, it can potentially cause congestion or packet loss in the network. With Auto-Bandwidth feature, the LSP bandwidth is set to some arbitrary value (including zero) during initial setup time, which will be periodically adjusted based on the actual bandwidth demand. The bandwidth adjustment uses the make-before-break (MBB) signaling method so that there is no disruption to the traffic flow.

4.2. Auto-bandwidth Theory of Operation

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 which is the head-end node of the LSP. The traffic rate samples are accumulated over the Adjustment-Interval period (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, will 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).

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.

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, the two representations help an operator to handle conditions when the bandwidth usage becomes too large or too small. In any case, the auto-bandwidth adjusted bandwidth is only reported to a PCE when MaxAvgBw crosses a threshold.

4.3. Scaling Considerations

It should be noted that any bandwidth change requires re-signaling of an LSP in a make-before-break fashion, 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

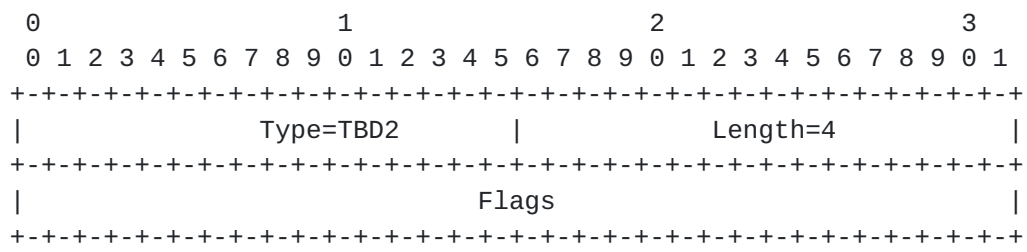
During PCEP Initialization Phase, PCEP Speakers (PCE or PCC) advertise their support of Automatic Bandwidth adjustment feature. A PCEP Speaker includes the "Auto-Bandwidth Capability" TLV, in the OPEN Object to advertise its support for PCEP Auto-Bandwidth extensions. The presence of the "Auto-Bandwidth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

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. If the PCEP speaker that supports the extensions of this document but did not advertise this capability, then upon receipt of AUTO-BANDWIDTH-ATTRIBUTES TLV in LSPA object, it SHOULD generate a PCErr with error-type 19 (Invalid Operation), error-value TBD4 (Auto-Bandwidth capability was not advertised) and terminate the PCEP session.

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.

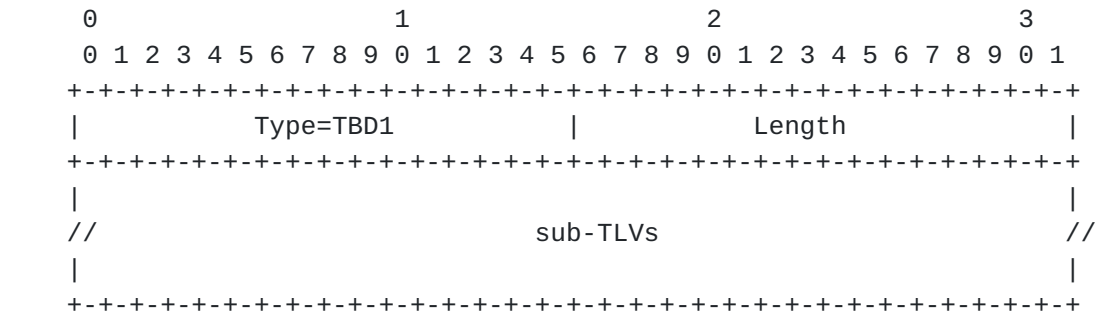
5.2. AUTO-BANDWIDTH-ATTRIBUTES TLV

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 [I-D.ietf-pce-pce-initiated-lsp], 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 indicate the PCEP speaker wish to disable the feature.

The format of the AUTO-BANDWIDTH-ATTRIBUTES TLV is shown in the following figure:



AUTO-BANDWIDTH-ATTRIBUTES TLV format

Type: TBD1

Length: The Length field defines the length of the value portion in octets as per [RFC5440].

Value: This comprises one or more sub-TLVs.

Following sub-TLVs are defined in this document:

Type	Len	Name
1	4	Sample-Interval sub-TLV
2	4	Adjustment-Interval sub-TLV
3	4	Down-Adjustment-Interval sub-TLV
4	4	Adjustment-Threshold sub-TLV
5	4	Adjustment-Threshold-Percentage sub-TLV
6	4	Down-Adjustment-Threshold sub-TLV
7	4	Down-Adjustment-Threshold-Percentage sub-TLV
8	4	Minimum-Bandwidth sub-TLV
9	4	Maximum-Bandwidth sub-TLV
10	8	Overflow-Threshold sub-TLV
11	4	Overflow-Threshold-Percentage sub-TLV
12	8	Underflow-Threshold sub-TLV
13	4	Underflow-Threshold-Percentage sub-TLV

Future specification can define additional sub-TLVs.

The sub-TLVs are encoded to inform the PCEP peer the various sampling and adjustment parameters. If sub-TLVs are not present, the default

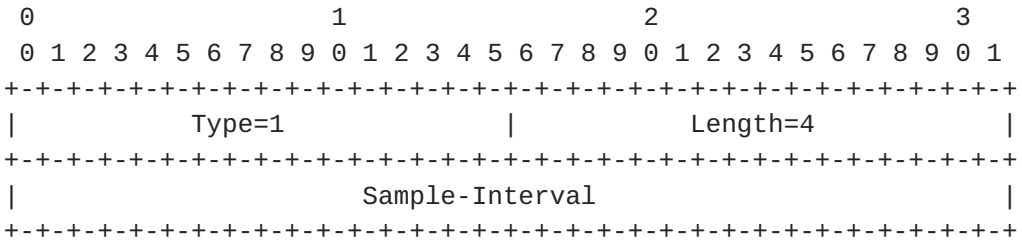
values as specified in this document are used or otherwise based on the local policy are assumed.

All sub-TLVs are optional and any unrecognized sub-TLV MUST be silently ignored. If a sub-TLV of same type appears more than once, only the first occurrence is processed and all others MUST be ignored.

The following sub-sections describe the sub-TLVs which are currently defined to be carried within the AUTO-BANDWIDTH-ATTRIBUTES TLV.

5.2.1. Sample-Interval sub-TLV

The Sample-Interval sub-TLV specifies a time interval in seconds at which traffic samples are collected at the PCC.



Sample-Interval sub-TLV format

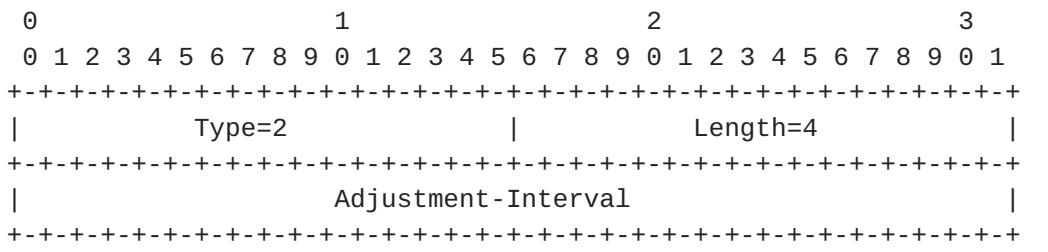
The Type is 1, Length is 4, and the value comprises of 4-octet time interval, the valid range is from 1 to 604800, in seconds. The default value is 300 seconds. The sample-interval parameter MUST NOT be greater than the (down) adjustment-interval.

5.2.2. Adjustment Intervals

The sub-TLVs in this section are encoded to inform the PCEP peer the adjustment interval parameters. An implementation MAY require to set different adjustment interval values for when the bandwidth usage trend is moving upwards or downwards. The Adjustment-Interval sub-TLV specify the time interval for both upward and downward trend (Up-Adjustment-Interval and Down-Adjustment-Interval). If the operator would like to use a different adjustment interval during the downward trend, the Down-Adjustment-Interval sub-TLV is included.

5.2.2.1. Adjustment-Interval sub-TLV

The Adjustment-Interval sub-TLV specifies a time interval in seconds at which bandwidth adjustment should be made when MaxAvgBw is greater than or less than the current bandwidth reservation of the LSP.

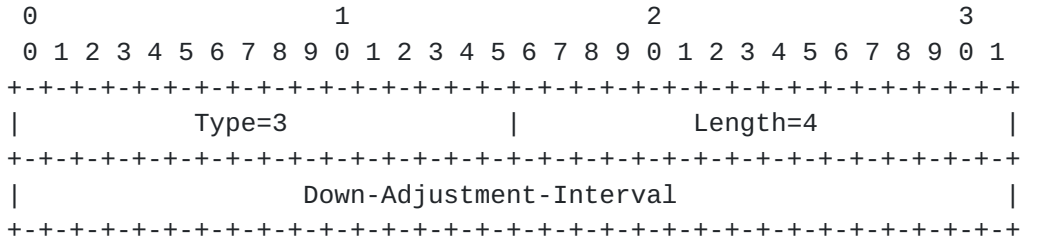


Adjustment-Interval sub-TLV format

The Type is 2, Length is 4, and the value comprises of 4-octet time interval, the valid range is from 1 to 604800, in seconds. The default value is 86400 seconds. The adjustment-interval parameter MUST NOT be less than the sample-interval.

5.2.2.2. Down-Adjustment-Interval sub-TLV

The Down-Adjustment-Interval sub-TLV specifies a time interval in seconds at which bandwidth adjustment should be made when MaxAvgBw is less than the current bandwidth reservation of the LSP. This parameter overwrites the Adjustment-Interval for the downward trend.



Down-Adjustment-Interval sub-TLV format

The Type is 3, Length is 4, and the value comprises of 4-octet time interval, 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.

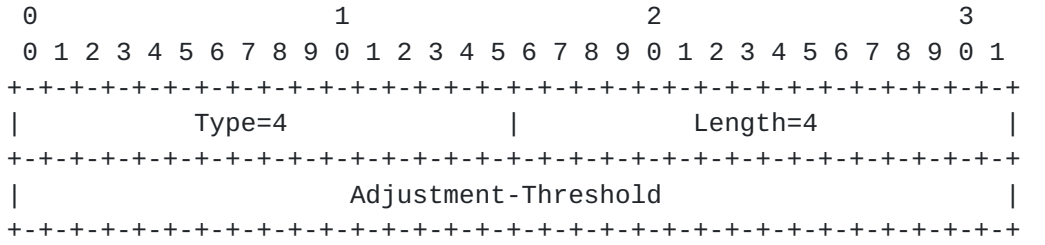
5.2.3. Adjustment Thresholds

The sub-TLVs in this section are encoded to inform the PCEP peer 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 specify 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 specify 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.

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.



Adjustment-Threshold sub-TLV format

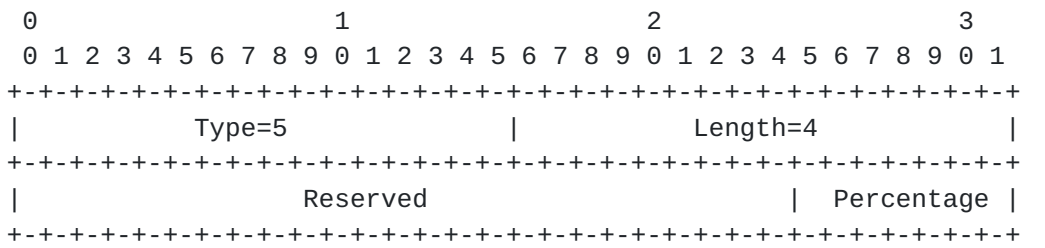
The Type is 4, Length is 4, and the value comprises of -

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

If the difference between the current MaxAvgBw and the current bandwidth reservation is greater than or less than or equal to the threshold value, the LSP bandwidth is adjusted to the current bandwidth demand (MaxAvgBw). The default adjustment-threshold value is not set.

5.2.3.2. Adjustment-Threshold-Percentage sub-TLV

The Adjustment-Threshold-Percentage 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.



Adjustment-Threshold-Percentage sub-TLV format

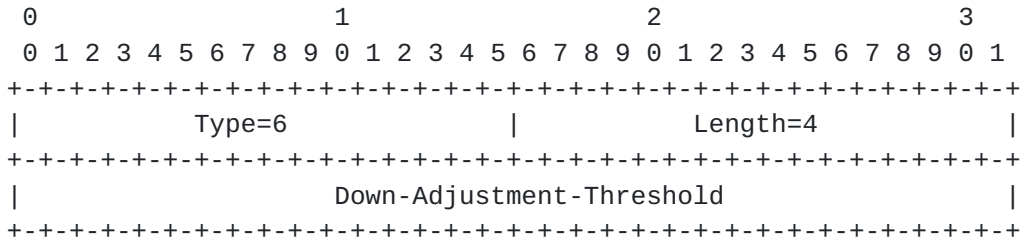
The Type is 5, Length is 4, and the value comprises of -

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

- Percentage: The Adjustment-Threshold value, encoded in percentage (an integer from 0 to 100). If the percentage difference between the current MaxAvgBw and the current bandwidth reservation is greater than or less than or equal to the threshold percentage, the LSP bandwidth is adjusted to the current bandwidth demand (MaxAvgBw). The default value is 5 percent.

5.2.3.3. Down-Adjustment-Threshold sub-TLV

The Down-Adjustment-Threshold 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 for the downward trend.



Down-Adjustment-Threshold sub-TLV format

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

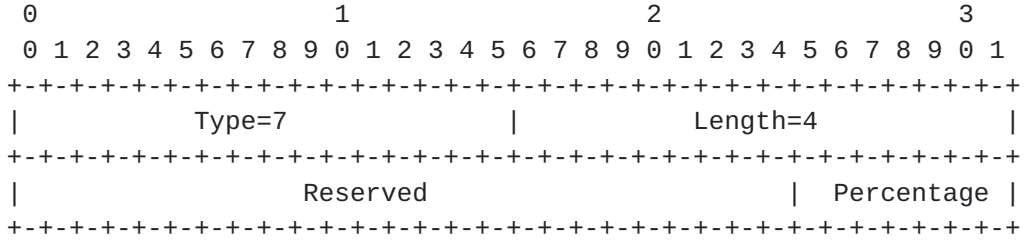
- Down-Adjustment Threshold: The absolute Down-Adjustment-Threshold bandwidth value, encoded in IEEE floating point format (see [IEEE.754.1985]), expressed in bytes per second. Refer to Section 3.1.2 of [RFC3471] for a table of commonly used values.

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). The default value equals the adjustment-threshold.

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.



Down-Adjustment-Threshold-Percentage sub-TLV format

The Type is 7, Length is 4, and the value comprises of -

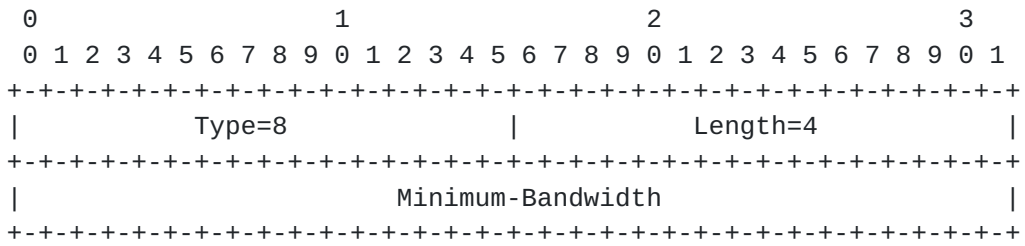
- Reserved: SHOULD be set to zero on transmission and MUST be ignored on receipt.

- Percentage: The Down-Adjustment-Threshold value, encoded in percentage (an integer from 0 to 100). If the percentage difference between the current bandwidth reservation and the current MaxAvgBw is greater than or equal to the threshold percentage, the LSP bandwidth is adjusted to the current bandwidth demand (MaxAvgBw). The default value equals the adjustment-threshold-percentage.

5.2.4. Minimum and Maximum Bandwidth Values

5.2.4.1. Minimum-Bandwidth sub-TLV

The Minimum-Bandwidth sub-TLV specify the minimum bandwidth allowed for the LSP, and is expressed in bytes per second. The LSP bandwidth cannot be adjusted below the minimum bandwidth value.

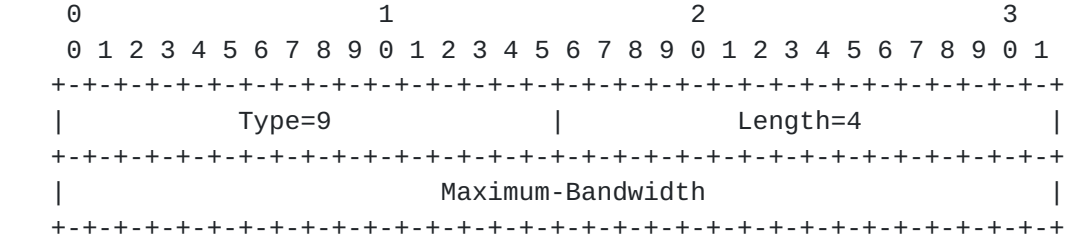


Minimum-Bandwidth sub-TLV format

The Type is 8, Length is 4, and the value comprises of 4-octet bandwidth value encoded in IEEE floating point format (see [IEEE.754.1985]), expressed in bytes per second. The default minimum-bandwidth value is set to 0. Refer to Section 3.1.2 of [RFC3471] for a table of commonly used values.

5.2.4.2. Maximum-Bandwidth sub-TLV

The Maximum-Bandwidth sub-TLV specify the maximum bandwidth allowed for the LSP, and is expressed in bytes per second. The LSP bandwidth cannot be adjusted above the maximum bandwidth value.



Maximum-Bandwidth sub-TLV format

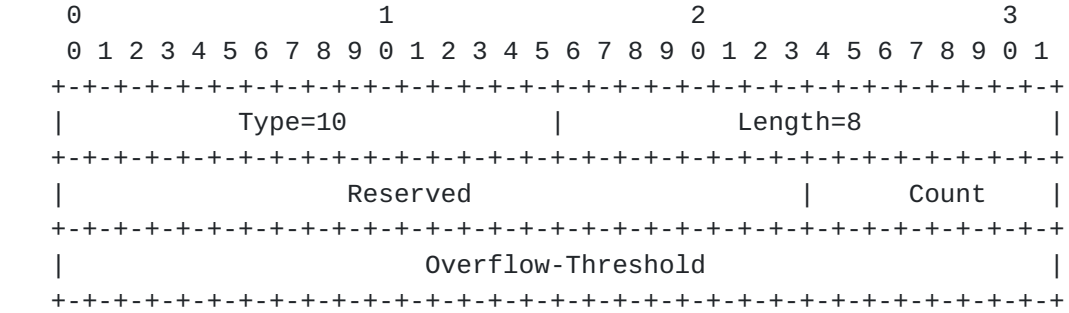
The Type is 9, Length is 4, and the value comprises of 4-octet bandwidth value encoded in IEEE floating point format (see [IEEE.754.1985]), expressed in bytes per second. The default maximum-bandwidth value is not set. Refer to Section 3.1.2 of [RFC3471] for a table of commonly used values.

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. The default 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 bandwidth should be adjusted immediately.



Overflow-Threshold sub-TLV format

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

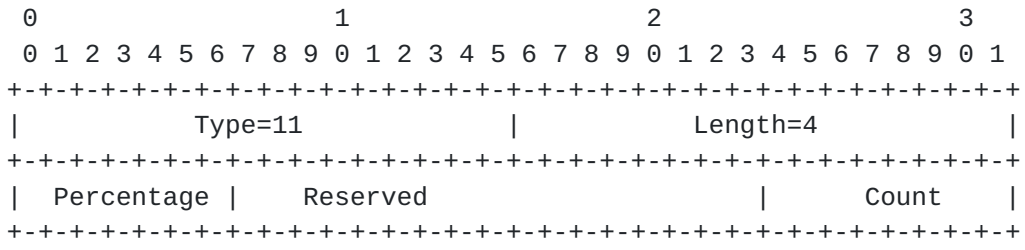
- Reserved: SHOULD be set to zero on transmission and MUST be ignored on receipt.

- Count: The Overflow-Count value, 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.

- Overflow-Threshold: The absolute Overflow-Threshold bandwidth value, encoded in IEEE floating point format (see [IEEE.754.1985]), expressed in bytes per second. Refer to Section 3.1.2 of [RFC3471] for a table of commonly used values. If the increase of the current MaxAvgBw from the current bandwidth reservation is greater than or equal to the threshold value, the overflow condition is met.

5.2.5.2. Overflow-Threshold-Percentage sub-TLV

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



Overflow-Threshold-Percentage sub-TLV format

The Type is 11, Length is 4, and the value comprises of -

- Percentage: The Overflow-Threshold value, encoded in percentage (an integer from 0 to 100). 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.

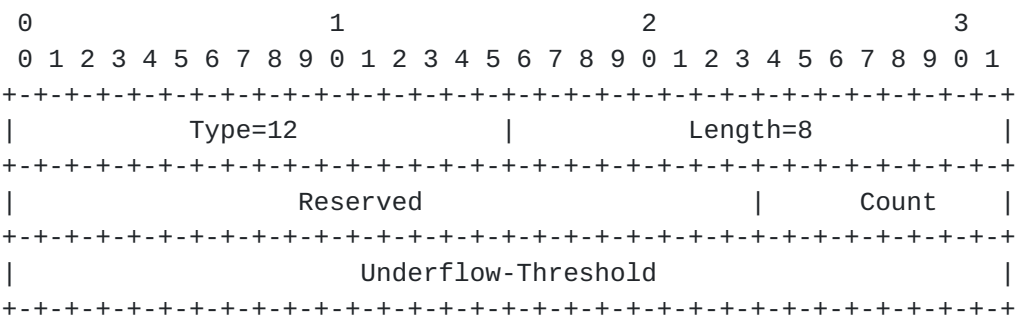
- Reserved: SHOULD be set to zero on transmission and MUST be ignored on receipt.

- Count: The Overflow-Count value, 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.

5.2.5.3. Underflow-Threshold sub-TLV

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



Underflow-Threshold sub-TLV format

The Type is 12, Length is 8, and the value comprises of -

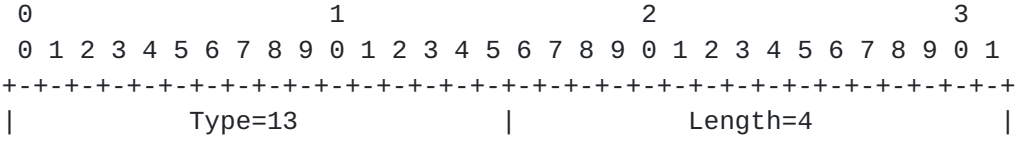
- Reserved: SHOULD be set to zero on transmission and MUST be ignored on receipt.

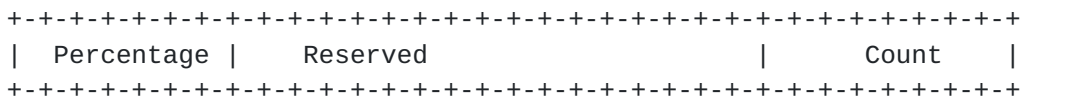
- Count: The Underflow-Count value, encoded in integer. The value 0 is considered to be invalid. The number of consecutive samples for which the underflow condition MUST be met for the LSP bandwidth to be immediately adjusted to the current bandwidth demand, bypassing the down-adjustment-interval.

- Underflow-Threshold: The absolute Underflow-Threshold bandwidth value, encoded in IEEE floating point format (see [IEEE.754.1985]), expressed in bytes per second. Refer to Section 3.1.2 of [RFC3471] for a table of commonly used values. If the decrease of the current MaxAvgBw from the current bandwidth reservation is greater than or equal to the threshold value, the underflow condition is met.

5.2.5.4. Underflow-Threshold-Percentage sub-TLV

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





Underflow-Threshold-Percentage sub-TLV format

The Type is 13, Length is 4, and the value comprises of -

- Percentage: The Underflow-Threshold value, encoded in percentage (an integer from 0 to 100). 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.

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

- Count: The Underflow-Count value, 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.

5.3. BANDWIDTH Object

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

- Requested Bandwidth: BANDWIDTH Object-Type value is 1.

- 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 [I.D.ietf-pce-pce-initiated-lsp].

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 definition (RBNF) of the PCInitiate message [I.D.ietf-pce-pce-initiated-lsp] 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 [I-D.ietf-pce-stateful-pce].

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 definition (RBNF) of the PCUpd message [I-D.ietf-pce-stateful-pce] is unchanged by this document.

5.6. The PCRpt Message

The PCRpt message [I-D.ietf-pce-stateful-pce] is a PCEP message sent by a PCC to a PCE to report the status of one or more LSPs.

For PCE-Initiated LSPs [I.D.ietf-pce-pce-initiated-lsp], 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 definition (RBNF) of the PCRpt message [I-D.ietf-pce-stateful-pce] 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 silently ignored.

- 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.

- 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 TLV which do not add any new security concerns beyond those discussed in [RFC5440] and [I-D.ietf-pce-stateful-pce] in itself. Some deployments may find the auto-bandwidth information as extra sensitive as it could be used to influence LSP path computation and LSP setup with adverse effect. Additionally, snooping of PCEP messages with such data or using PCEP messages for network reconnaissance, may give an attacker sensitive information about the operations of the network. Thus, such deployment should employ suitable PCEP security mechanisms like TCP Authentication Option (TCP-AO) [RFC5925] or [I-D.ietf-pce-pceps].

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 allocated only by an IETF Review action [RFC5226]. Each bit should be tracked with the following qualities:

- Bit number (counting from bit 0 as the most significant bit)
- Capability description
- 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 in the "PCEP TLV Type Indicators" for the sub-TLVs carried in the AUTO-BANDWIDTH-ATTRIBUTES TLV. New sub-TLV are allocated only by an IETF Review action [RFC5226].

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]
14-65535	Unassigned	[This document]

8.4. Error Object

This document defines a new Error-Value for PCErr message of Error-Type 19 (Invalid Operation) [I-D.ietf-pce-stateful-pce]; 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 Types 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.

[RFC5226] Narten, T. and H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 5226, May 2008.

[RFC5440] Vasseur, JP. and JL. Le Roux, "Path Computation Element (PCE) Communication Protocol (PCEP)", RFC 5440, March 2009.

[I-D.ietf-pce-stateful-pce] Crabbe, E., Minei, I., Medved, J., and R. Varga, "PCEP Extensions for Stateful PCE", draft-ietf-pce-stateful-pce (work in progress).

[I-D.ietf-pce-pce-initiated-lsp] Crabbe, E., Minei, I., Sivabalan, S., and R. Varga, "PCEP Extensions for PCE-initiated LSP Setup in a Stateful PCE Model", draft-ietf-pce-pce-initiated-lsp (work in progress).

9.2. Informative References

[RFC3471] Berger, L., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Functional Description", RFC 3471, January 2003.

[RFC5925] Touch, J., Mankin, A., and R. Bonica, "The TCP Authentication Option", RFC 5925, June 2010.

[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.

[RFC8051] Zhang, X. and I. Minei, "Applicability of a Stateful Path Computation Element (PCE)", RFC 8051, January 2017.

[I-D.ietf-pce-pceps] Lopez, D., Dios, O., Wu, W., and D. Dhody, "Secure Transport for PCEP", draft-ietf-pce-pceps (work in progress).

[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).

[I-D.gandhi-pce-pm] Gandhi, R., Wen, B., Barth, C., and D. Dhody, "PCEP Extensions for Reporting MPLS-TE LSP Performance Measurements", draft-gandhi-pce-pm (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.

Email: rgandhi@cisco.com

Luyuan Fang
eBay
USA

Email: lufang@ebay.com