PCE Working Group
Internet-Draft
Intended status: Standards Track
Expires: August 2, 2017

D. Dhody
U. Palle
Huawei Technologies
R. Singh
Juniper Networks
R. Gandhi
Individual Contributor
L. Fang
eBay

January 29, 2017

PCEP Extensions for MPLS-TE LSP Automatic Bandwidth Adjustment with Stateful PCE
draft-ietf-pce-stateful-pce-auto-bandwidth-01

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 adjustment allows automatic and dynamic adjustment of the reserved bandwidth allocation of an TE LSP based on the volume of traffic flowing through it. 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. In this document, the focus is on Active stateful PCE where LSPs are configured at the PCC and control over them is delegated to the PCE. Further [I-D.ietf-pce-pce-initiated-lsp] describes the setup, maintenance and teardown of PCE-initiated LSPs for the stateful PCE model.

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. Ingress Label Switch Router (LSR) collects the traffic rate at each sample interval to determine the bandwidth demand of the LSP. This bandwidth information is then used to adjust the LSP bandwidth periodically. This feature is commonly

referred to as Auto-Bandwidth.

Enabling Auto-Bandwidth feature on an LSP results in the LSP automatically adjusting its bandwidth reservation based on the actual 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 LSP monitors the traffic flow and adjusts its bandwidth every adjustment-interval period. The bandwidth adjustment uses the make-before-break signaling method so that there is no interruption to the traffic flow. The Auto-Bandwidth is described in detail in Section 4.1. [RFC8051] describes the use-case for Auto-Bandwidth adjustment for passive and active stateful PCE.

The PCC (head-end of the LSP) monitors the traffic 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.

This document defines extensions needed to support Auto-Bandwidth feature on the LSPs in a active stateful PCE model using PCEP.

Note that, another document [I-D.gandhi-pce-pm], describes the PCEP extensions to report the performance measurements to the PCE, this includes the bandwidth usage information of a TE LSP and can be used at the PCE to calculate the new bandwidth to be adjusted.

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. Terminology

The following terminology is used in this document.

Active Stateful PCE: PCE that uses tunnel state information learned from PCCs to optimize path computations. Additionally, it actively updates tunnel parameters in those PCCs that delegated control over their tunnels to the PCE.

Delegation: An operation to grant a PCE temporary rights to modify a subset of tunnel parameters on one or more PCC's tunnels. Tunnels are delegated from a PCC to a PCE.

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

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

TE LSP: Traffic Engineering Label Switched Path.

Note the Auto-Bandwidth feature specific terms defined in Section 4.1.

3. Requirements for PCEP Extensions

The PCEP speaker supporting this document MUST have a mechanism to advertise the automatic bandwidth adjustment capability.

PCEP extensions required 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

Further Auto-Bandwidth deployment considerations are summarized below:

- It is required to identify and inform the PCEP peer, 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.

- Further for the LSP with Auto-Bandwidth feature enabled, an operator should be able to specify the adjustment parameters (i.e. configuration knobs) to control this feature (e.g. minimum/ maximum bandwidth range) and PCEP peer should be informed.

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 bandwidth adjustment uses the make-before-break signaling method so that there is no interruption to the traffic flow.

The new bandwidth reservation is determined by sampling 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 can be set to some arbitrary value (including zero) during initial setup time, and it will be periodically adjusted over time based on the actual bandwidth requirement.

Note the following definitions of the Auto-Bandwidth terms:

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 averaged traffic bandwidth rate in a given adjustment-interval.

Adjusted Bandwidth: This is the Auto-Bandwidth computed bandwidth that needs to be adjusted for the LSP.

Sample-Interval: The periodic time interval at which the traffic rate is collected as a sample.

Bandwidth-Sample (BwSample): The bandwidth sample collected at every sample interval to measure the traffic rate.

Adjustment-Interval: The periodic time interval at which the bandwidth adjustment should be made using the MaxAvgBw.

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

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

Adjustment-Threshold: This value 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 to the current bandwidth demand (Adjusted Bandwidth) at the adjustment-interval expiry.

Overflow-Count: This value 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 value 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 adjustment-interval if the overflow-condition is met consecutively for the Overflow-Count.

Underflow-Count: This value 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 value 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 adjustment-interval if the underflow-condition is met consecutively for the Underflow-Count.

4.2. Auto-bandwidth Theory of Operation

The traffic rate is periodically sampled at each sample-interval (which can be configured by the user and the default value as 5 minutes) by the PCC which is the head-end node of the LSP. The sampled traffic rates are accumulated over the adjustment-interval period (which can be configured by the user and the default value as 24 hours). The PCC is in-charge of calculating the bandwidth to be adjusted, will adjust the bandwidth of the LSP to the highest sampled traffic rate (MaxAvgBw) amongst the set of bandwidth samples collected over the adjustment-interval.

Note that the highest sampled traffic rate 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), the LSP bandwidth is adjusted 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 intervals for adjustment is based on the traffic pattern of the LSP.

In order to avoid frequent re-signaling, an operator may set a longer adjustment-interval value. 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.

All thresholds in this document could be represented in both absolute value and percentage, and could be used together.

4.3. Scaling Considerations

It should be noted that any bandwidth change would require 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 bandwidth to be adjusted (see

Section 5.6). Similarly if a PCC gets overwhelmed due to signaling churn, it can notify the PCE to temporarily suspend the LSP bandwidth adjustment.

5. Extensions to the PCEP

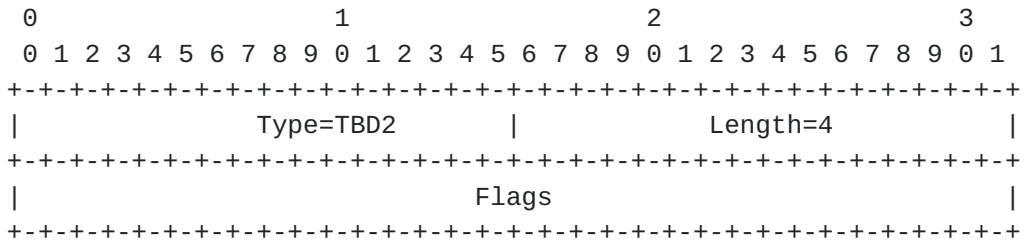
5.1. Capability Advertisement

During PCEP Initialization Phase, PCEP Speakers (PCE or PCC) advertise their support of Automatic Bandwidth Adjustment. 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 Adjustment 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 draft but did not advertise this capability, then upon receipt of AUTO-BANDWIDTH-ATTRIBUTE 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 it will 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). Currently no

flags are defined for this TLV.

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.

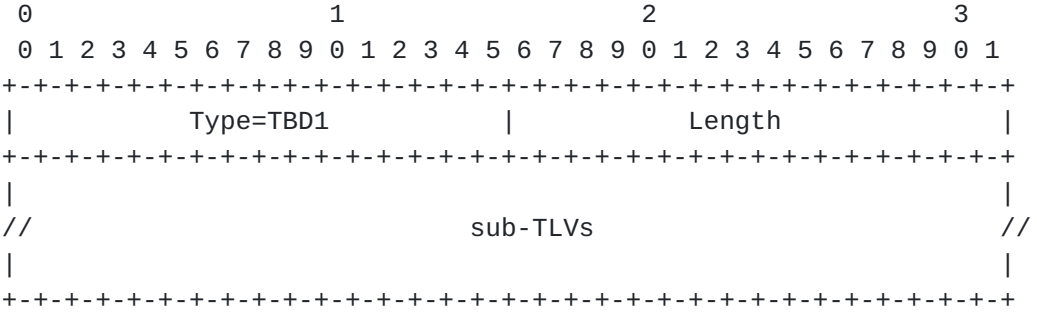
5.2. AUTO-BANDWIDTH-ATTRIBUTE TLV

The AUTO-BANDWIDTH-ATTRIBUTE 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 delegated LSPs, this TLV is carried in the PCRpt message in LSPA Object. This TLV is also included in the LSPA object with the PCUpd message to direct the PCE to use different parameters with the LSP.

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-ATTRIBUTE TLV is shown in the following figure:



AUTO-BANDWIDTH-ATTRIBUTE TLV format

Type: TBD1

Length: The Length field defines the length of the value portion in bytes 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	Adjustment-Threshold sub-TLV
4	4	Adjustment-Threshold-Percentage sub-TLV
5	4	Minimum-Bandwidth sub-TLV
6	4	Maximum-Bandwidth sub-TLV
7	8	Overflow-Threshold sub-TLV
8	4	Overflow-Threshold-Percentage sub-TLV
9	8	Underflow-Threshold sub-TLV
10	4	Underflow-Threshold-Percentage sub-TLV

Future specification can define additional sub-TLVs.

The presence of AUTO-BANDWIDTH-ATTRIBUTE TLV in LSPA Object means that the automatic bandwidth adjustment feature is enabled. 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 AUTO-BANDWIDTH-ATTRIBUTE TLV can also be carried in PCUpd message in LSPA Object in order to make updates to auto-bandwidth attributes such as Adjustment-Interval.

If sub-TLVs are not present, the default values based on the local policy are assumed.

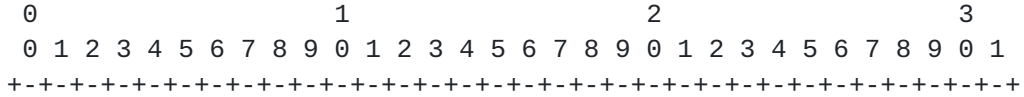
The sub-TLVs are encoded to inform the PCEP peer the various sampling and adjustment parameters.

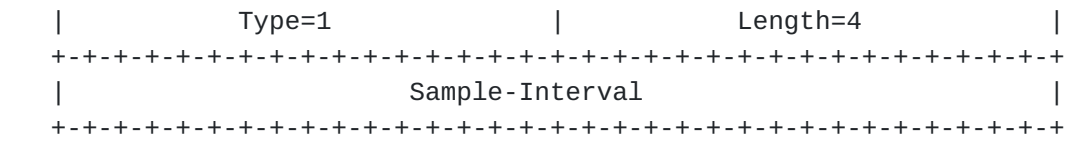
The following sub-sections describe the sub-TLVs which are currently defined to be carried within the AUTO-BANDWIDTH-ATTRIBUTE TLV.

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.

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.



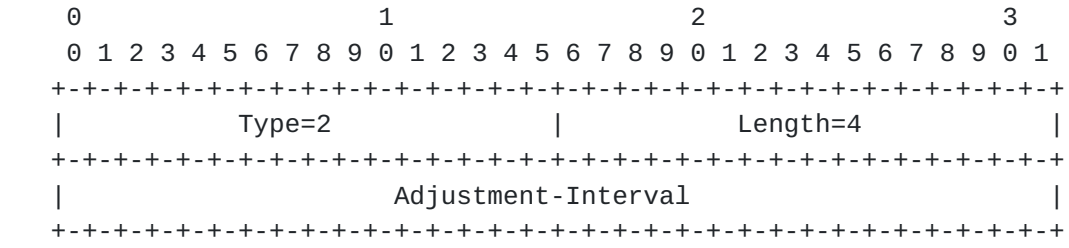


Sample-Interval sub-TLV format

5.2.2. Adjustment-Interval sub-TLV

The Adjustment-Interval sub-TLV specifies a time interval in seconds at which bandwidth adjustment should be made.

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 300 seconds.



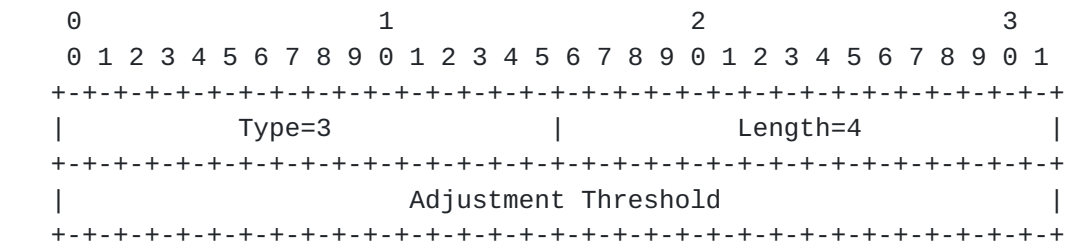
Adjustment-Interval sub-TLV format

5.2.3. Adjustment Threshold

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.

5.2.3.1. Adjustment-Threshold sub-TLV

The Adjustment-Threshold sub-TLV is used to decide when the LSP bandwidth should be adjusted.



Adjustment-Threshold sub-TLV format

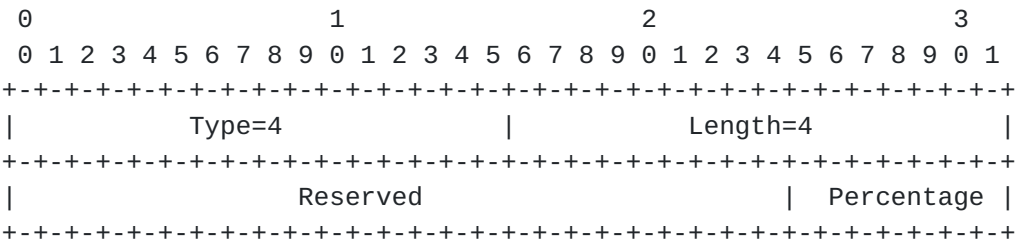
The Type is 3, Length is 4, and the value comprises of -

- 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 equal to the threshold value, the LSP bandwidth is adjusted to the current bandwidth demand.

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.



Adjustment-Threshold-Percentage sub-TLV format

The Type is 4, 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 equal to the threshold percentage, the LSP bandwidth is adjusted to the current bandwidth demand.

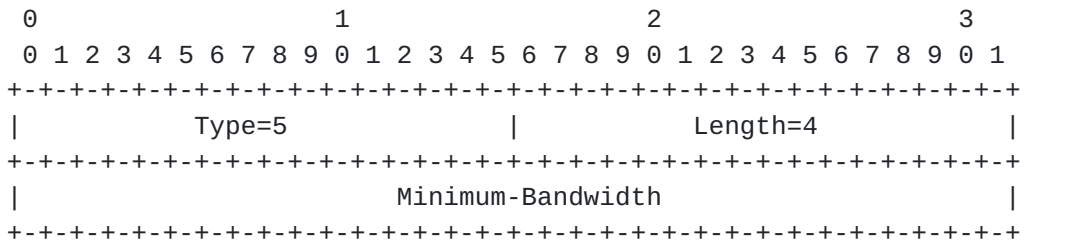
5.2.4. Minimum and Maximum Bandwidth Values

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.

The Type is 5, 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. Refer to Section 3.1.2 of [RFC3471] for a table of commonly used values.

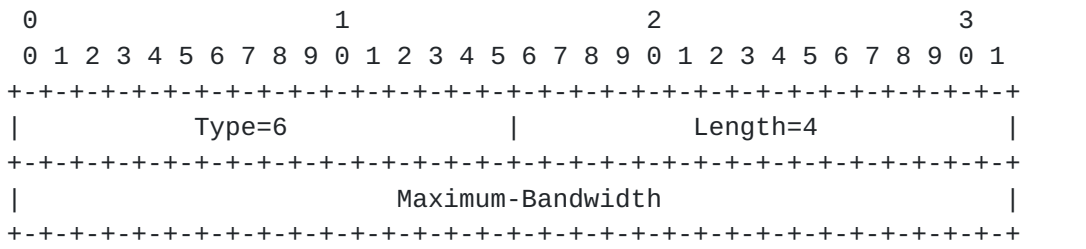


Minimum-Bandwidth sub-TLV format

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.

The Type is 6, 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. Refer to Section 3.1.2 of [RFC3471] for a table of commonly used values.



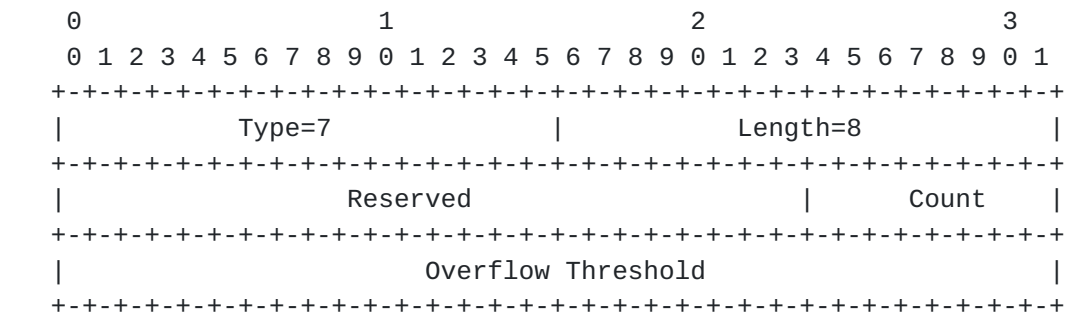
Maximum-Bandwidth sub-TLV format

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 the absolute value and the percentage for the threshold, in which case the bandwidth is immediately adjusted when either of the adjustment threshold conditions are met consecutively for the given count.

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 7, Length is 8, and the value comprises of -

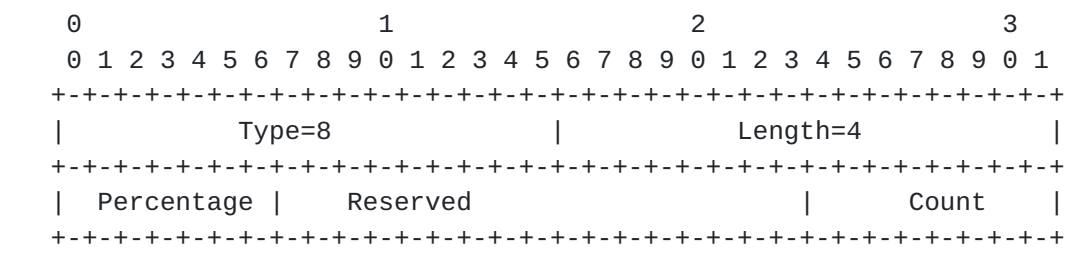
- Reserved: SHOULD be set to zero on transmission and MUST be ignored on receipt.

- 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 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 8, Length is 4, and the value comprises of -

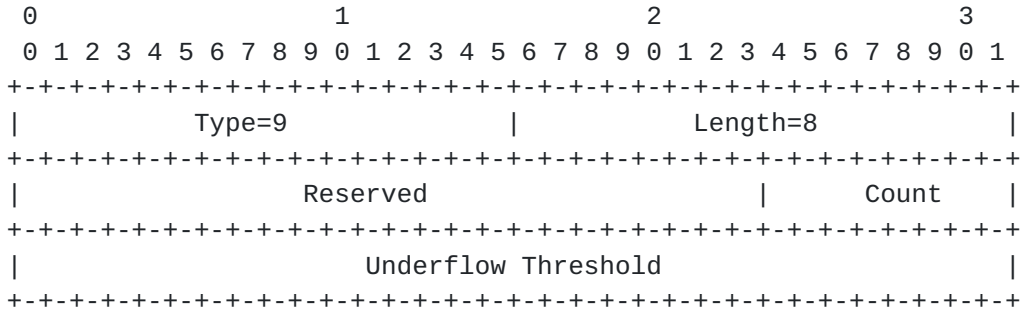
- Percentage: The Overflow-Threshold value, encoded in percentage (an integer from 0 to 100). If the percentage increase of the current MaxAvgBw from the current bandwidth reservation is greater than or equal to the threshold percentage, the overflow condition is met.

- 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 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 9, Length is 8, and the value comprises of -

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

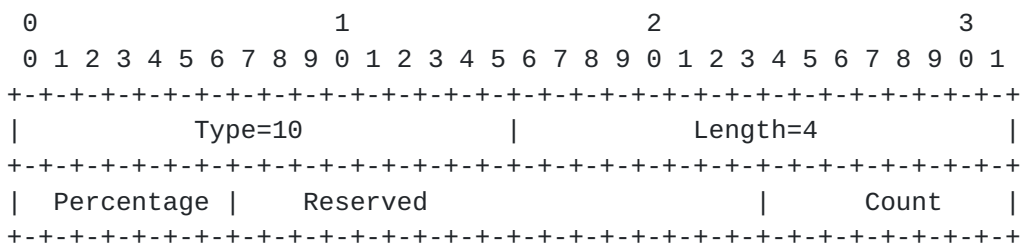
- Count: The Underflow-Count value, encoded in integer. The value 0 is considered to be invalid. The number of consecutive samples for which the underflow condition MUST be met for the LSP bandwidth to be immediately adjusted to the current bandwidth demand, bypassing the adjustment-interval.

- 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 10, 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 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.

PCC reports the calculated bandwidth to be adjusted (MaxAvgBw) to the

PCE using existing 'Requested Bandwidth with BANDWIDTH Object-Type as 1.

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 [I-D.ietf-pce-pce-initiated-lsp] with Auto-Bandwidth feature enabled, AUTO-BANDWIDTH-ATTRIBUTE TLV MUST be included in the LSPA object with the PCInitiate message. The rest of the processing remains unchanged.

5.5. The PCRpt Message

As specified in [I.D.ietf-pce-pce-initiated-lsp], the PCC creates the LSP using the attributes communicated by the PCE, and local values for the unspecified parameters. After the successful instantiation of the LSP, PCC automatically delegates the LSP to the PCE and generates an LSP State Report (PCRpt) for the LSP.

When LSP is delegated to a PCE for the very first time, BANDWIDTH object of type 1 is used to specify the requested bandwidth in the PCRpt message.

When the LSP is enabled with the Auto-Bandwidth feature, PCC SHOULD include the BANDWIDTH object of type 1 to specify the calculated bandwidth to be adjusted to the PCE in the PCRpt message.

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

5.6. 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. As described in Section 4.3 of this document, a PCEP speaker SHOULD notify its PCEP peer that it is overwhelmed, and on 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 adjustment is received, 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 adjustment, 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/re-routing, it can also notify the PCE to not adjust the LSP bandwidth while in overwhelm state.

5.7. 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 the LSPs with Auto-Bandwidth feature enabled, AUTO-BANDWIDTH-ATTRIBUTE TLV MUST be included in the LSPA object with the PCUpd message. The PCE can use this TLV to direct the PCC to change the auto bandwidth parameters. The rest of the processing remains unchanged.

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

6. Security Considerations

This document defines AUTO-BANDWIDTH-CAPABILITY TLV, AUTO-BANDWIDTH-ATTRIBUTE TLV which does 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 and could be used to influence path computation and 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 tunnel (at ingress (PCC) or PCE), the values for parameters like sample-interval, adjustment-interval, minimum-bandwidth, maximum-bandwidth, adjustment-threshold 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, the preferred mechanism for configuration is through a YANG model [I-D.ietf-pce-pcep-yang]. This 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

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

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

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, and MAY allow a limit to be placed on the rate of messages sent by a PCEP speaker and received from a peer related to auto-bandwidth. It MAY also allow sending a notification when the PCEP speaker is overwhelmed or a rate threshold is reached.

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 this registry. <http://www.iana.org/assignments/pcep/pcep.xhtml#pcep-tlv-type-indicators>.

Value	Name	Reference
TBD2	AUTO-BANDWIDTH-CAPABILITY	[This I.D.]
TBD1	AUTO-BANDWIDTH-ATTRIBUTE	[This I.D.]

8.2. AUTO-BANDWIDTH-CAPABILITY TLV Flag Field

IANA is requested to create a 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

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

8.3. AUTO-BANDWIDTH-ATTRIBUTE Sub-TLV

This document specifies the AUTO-BANDWIDTH-ATTRIBUTE Sub-TLVs. IANA is requested to create an "AUTO-BANDWIDTH-ATTRIBUTE Sub-TLV Types" sub-registry in the "PCEP TLV Type Indicators" for the sub-TLVs carried in the AUTO-BANDWIDTH-ATTRIBUTE 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 I.D.]
1	Sample-Interval sub-TLV	[This I.D.]
2	Adjustment-Interval sub-TLV	[This I.D.]

3	Adjustment-Threshold sub-TLV	[This I.D.]
4	Adjustment-Threshold-Percentage sub-TLV	[This I.D.]
5	Minimum-Bandwidth sub-TLV	[This I.D.]
6	Maximum-Bandwidth sub-TLV	[This I.D.]
7	Overflow-Threshold sub-TLV	[This I.D.]
8	Overflow-Threshold-Percentage sub-TLV	[This I.D.]
9	Underflow-Threshold sub-TLV	[This I.D.]
10	Underflow-Threshold-Percentage sub-TLV	[This I.D.]
11-65535	Unassigned	[This I.D.]

8.4. Error Object

This document defines a new Error-Value for PCErr message of type 19 (Invalid Operation) [I-D.ietf-pce-stateful-pce]); IANA is requested to make the following allocation from this registry. <http://www.iana.org/assignments/pcep/pcep.xhtml#pcep-error-object>

Error-Value	Meaning	Reference
TBD4	Auto-Bandwidth Capability was not Advertised	[This I.D.]

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 I.D.]
	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), October 2016.

[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), November 2016

[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: udayasree.palle@huawei.com

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

EMail: ravis@juniper.net

Rakesh Gandhi
Individual Contributor

EMail: rgandhi.ietf@gmail.com

Luyuan Fang
eBay
USA

EMail: luyuanf@gmail.com