

IDR
Internet-Draft
Intended status: Standards Track
Expires: February 2, 2017

S. Shah
K. Patel
Cisco Systems
S. Bajaj
Juniper Network
L. Tomotaki
Verizon
M. Boucadair
Orange
August 1, 2016

**Inter-domain SLA Exchange Attribute
draft-ietf-idr-sla-exchange-09.txt**

Abstract

Network administrators typically enforce Quality of Service (QoS) policies according to Service Level Agreement (SLA) with their providers. The enforcement of such policies often relies upon vendor-specific configuration language. Both learning of SLA, either thru SLA documents or via some other out-of-band method, and translating them to vendor specific configuration language is a complex, often manual, process and prone to errors.

This document specifies an optional transitive attribute to signal SLA parameters in-band, across administrative boundaries (considered as Autonomous Systems (AS)), thus simplifying and facilitating some of the complex provisioning tasks in situations where BGP is available as a routing protocol.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of [BCP 78](#) and [BCP 79](#).

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on February 2, 2017.

Copyright Notice

Copyright (c) 2016 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to [BCP 78](#) 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.	Terminology	4
3.	QoS Attribute Definition	5
3.1.	QoS Attribute SubType	6
3.2.	SLA SubType	7
3.3.	SLA Content for ADVERTISE SLA Event	9
3.3.1.	Supported IPFIX identifiers for Traffic Class Elements	13
3.3.2.	Traffic Class Service types and respective TLVs . . .	14
4.	Originating SLA Notification	21
4.1.	SLA Contexts	21
4.1.1.	SLA Advertisement for Point-to-Point Connection . . .	21
4.1.2.	SLA Advertisement for Destination AS Multiple Hops Away	22
5.	QoS Attribute Handling at Forwarding Nodes	22
5.1.	BGP Node Capable of Processing QoS Attribute	23
5.2.	QoS Attribute Handling at Receiver	23
6.	Error Handling	23
7.	Deployment Considerations	24
8.	Acknowledgements	25
9.	IANA Considerations	25
10.	Security Considerations	27
11.	References	28
11.1.	Normative References	28
11.2.	Informative References	29
	Authors' Addresses	30

1. Introduction

Typically there is a contractual Service Level Agreement (SLA) for QoS established between a customer and a provider or between providers [[RFC7297](#)]. This QoS SLA defines the nature of the various traffic classes and services needed within each traffic class. The contract may include full line-rate or sub line-rate without additional traffic classes, or it may contain additional traffic classes and service definitions for those traffic classes. Finer granular traffic classes may be based on some standard code points (e.g., based on DSCP (Differentiated Services Code Point)) or specific set of prefixes.

Once the contractual QoS SLA is established, QoS SLA parameters are enforced in some or all participating devices by deriving those parameters into configuration information on respective devices. The network administrator translates the QoS SLA to QoS policies using router (vendor) specific provisioning language. In a multi-vendor network, translating SLAs into technology-specific and vendor-specific configuration requires the network administrator to consider specific configuration of each vendor. There does not exist any standard protocol to translate SLA agreements into technical clauses and configurations and thus both the steps of out of band learning of negotiated SLA and provisioning them in a vendor specific language can be complex and error-prone.

SLA parameters may have to be exchanged through organizational boundaries, thru SLA documents or via some other off-band method, to an administrator provisioning actual devices. For example, to provide voice services, the provider may negotiate QoS parameters (like min/max rates) for such traffic classified under the EF (Expedited Forwarding) codepoint in Diffserv-enabled [[RFC2475](#)] networks. The Administrator at the CE (Customer Edge) not only will have to know that provider's service for voice traffic is EF-based but will also have to know how to implement DSCP EF classification rule along with Low Latency Service, and possibly min/max rate enforcement for the optimal use of bandwidth, as per vendor specific provisioning language.

The Inter-domain exchange of QoS SLA policy described in this document does not require any specific method for the provider in establishing SLAs. It only requires that the provider wishes to send the QoS SLA policy via BGP UPDATE [[RFC4271](#)] messages from the provider to a set of receivers (BGP peers). In reaction to, a receiving router may translate that to relevant QoS policy definition on the device. The SLA negotiation and assurance is outside the scope of this document.

This document defines a new optional BGP transitive attribute, referred as QoS Attribute, which has as one of its sub-types the SLA policy. The BGP node of the originating AS sends this QoS Attribute, for prefixes this QoS SLA Policy applies to, in a BGP UPDATE message that will be distributed to a list of destination ASes. The QoS SLA policy can be for inbound traffic to the advertising AS or outbound traffic from the advertising AS, or both.

Protocols and data models are being created to standardize setting routing configuration parameters within networks. YANG data models [[RFC6020](#)] are being developed so that NETCONF ([RFC6241](#)) or RESTConf ([I-D.ietf-netconf-restconf](#)) can set these standardize in configuration mechanisms. For ephemeral state, the I2RS protocol is being developed to set ephemeral state. While these protocols provide valid configuration within a domain or across domains, some providers desire to exchange QoS parameters in-band utilizing BGP peering relationships. This is similar to the distribution of Flow Specification information via BGP peering relationships (see [[RFC5575](#)] and [[RFC7674](#)]).

2. Terminology

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](#)].

BGP Speaker: A functional component on a BGP capable device that functions as per BGP specification.

BGP Peers: BGP Speakers adjacent to each other.

QoS Attribute Speaker: A functional component on a BGP capable device that produces and/or processes content of the QoS Attribute. A device that is QoS Attribute Speaker is also always a BGP Speaker. However, a BGP Speaker not necessarily always a QoS Attribute Speaker.

QoS Attribute content is produced and processed outside the function of the BGP Speaker and thus content of the QoS Attribute is completely opaque to the BGP Speaker. At BGP capable device where QoS Attribute content is produced, length and value of the QoS Attribute is passed from QoS Attribute Speaker to the BGP Speaker where BGP Speaker inserts the attribute into the BGP UPDATE message with appropriate attribute flags, attribute type, and length and value passed from the QoS Attribute Speaker. Similarly, a BGP capable device when receives QoS Attribute in the BGP UPDATE message, BGP Speaker extracts QoS Attribute value from the message and passes it to the QoS Attribute Speaker where QoS Attribute Speaker processes

the content from that passed down value. How the content of the QoS Attribute is passed from the QoS Attribute Speaker to the BGP Speaker and vice versa is implementation specific.

In the context of use of QoS Attribute for SLA parameters exchange, following roles are defined further within the scope of the QoS Attribute Speaker.

SLA Producer: This is a QoS Attribute Speaker that produces QoS Attribute for the SLA SubType.

SLA Consumer: This is a QoS Attribute Speaker that is intended receiver of QoS Attribute with the SLA SubType.

3. QoS Attribute Definition

The QoS Attribute is an optional transitive attribute (TBD - attribute code to be assigned by IANA) which is applicable to the Source AS and NLRIs advertised in the BGP UPDATE message this attribute is included in. The format of the QoS Attribute is shown in Figure 1.

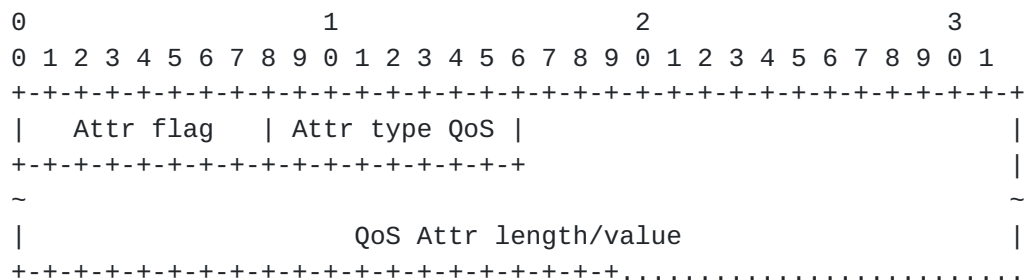


Figure 1

Attribute flags - 8-bits field

highest order bit (bit 0) - MUST be set to 1, since this is an optional attribute

2nd higher order bit (bit 1) - MUST be set to 1, since this is a transitive attribute

The content of the QoS Attribute is further specified with flags, applicable to QoS Attribute content, and a SubType in a TLV form.

3.1. QoS Attribute SubType

The Value field of the QoS Attribute contains the following:

QoS Attribute flags

Tuple (SubType of the QoS Attribute, SubType length, SubType value)

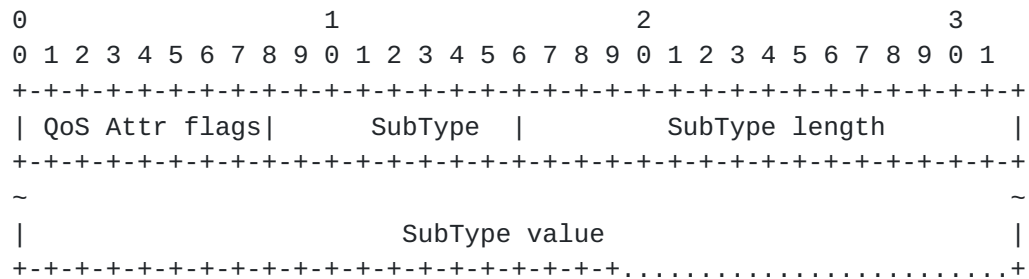


Figure 2 - Format of QoS Attribute

QoS Attr flags - 8-bits field

All bits of this field are currently un-used. The space is provided for future use. All bits MUST be set to zero when sent. The values (0x01-0xFF) are reserved, and MUST be ignored when received.

SubType - 8-bits field with following values:

0x00 = reserved

0x01 = SLA

0x02 - 0xf0 = reserved for future use (Standards Action)

0xf1 - 0xff = Private use

The only SubType of the QoS Attribute defined in this specification is the SLA SubType.

SubType length - 16-bits field that specifies length of the SubType value in number of octets.

SubType value - variable length field, as expressed in SubType length, that contains information about a specified SubType. For the SLA SubType the information is about sender and receiver(s), and SLA parameters as described in [Section 3.2](#).

3.2. SLA SubType

Format of the SLA SubType Value field is shown in Figure 3.

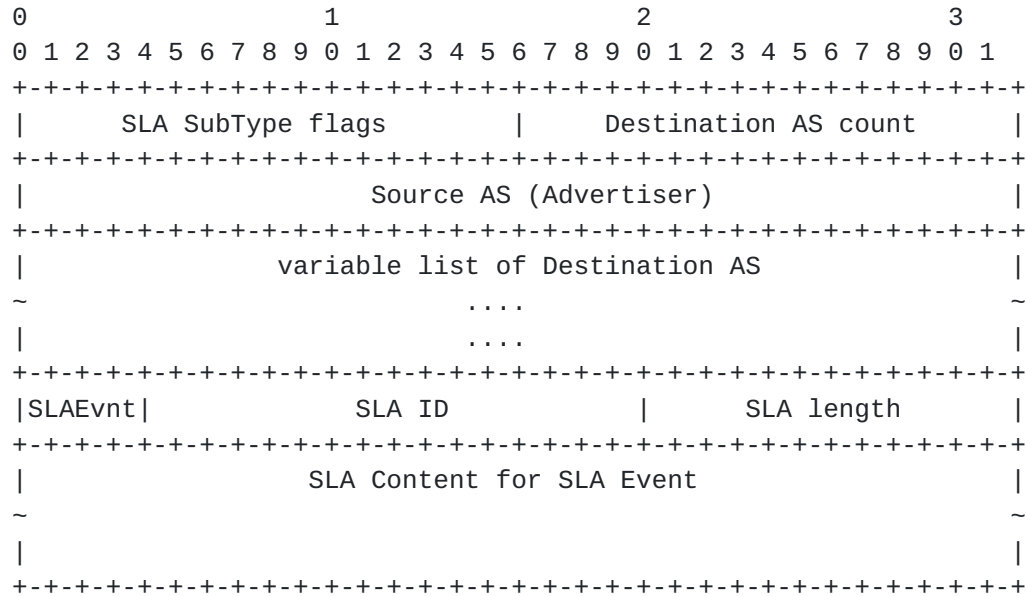


Figure 3 - Format of the SLA SubType of the QoS attribute

SLA SubType flags - 16-bits field

highest order bit (bit 0) -

If set to 1 indicates that SLA ID and SLA Content, specified in this SLA SubType, is from the source AS to the list of Destination AS specified in the same SLA SubType.

If set to 0 indicates to ignore Source AS and list of Destination AS specified in this SLA SubType field. SLA ID and SLA Content, specified in this SLA SubType, are intended for the peer receiver of the BGP UPDATE message. In such a case, on reception of such a message, QoS Attribute Speaker SHOULD drop the QoS Attribute from the BGP UPDATE message and rest of the BGP UPDATE message should be processed by BGP Speaker as per BGP specification.

Rest all other bits are currently un-used.

Destination AS count - 16-bits field that specifies count of destination ASNs present in the Destination AS list.

This count has no functional value when highest order bit in the SLA SubType flags field is set to 0. When highest order bit is set to 1 and if this count is 0 then that is an error condition which should be handled as described in [Section 6](#).

Source AS - 32-bits field (AS number space as defined in [RFC6793](#))

This is the AS where SLA Content is originated from. The Source AS MUST be of the same AS that is originating SLA ID and SLA Content.

When highest order bit in the SLA SubType flags field is set to 0, this Source AS value MUST be ignored. In such a case SLA ID and SLA Content of this SLA SubType is intended for the peer receiver of the BGP UPDATE message.

When highest order bit in the SLA SubType flags field is set to 1, the Source AS value of 0 is illegal and thus should be considered an error which should be handled as described in [Section 6](#).

Destination AS list - variable length field that holds as many ASN identifiers, each is 32-bits (AS number space is defined in [RFC6793](#)), as specified in the Destination AS count field.

List of ASNs for which the SLA is relevant to, each of which is a 32-bit number. If Destination AS count is set to 0 then this field MUST NOT be included.

SLA Event - 4-bits field with following values:

0x0 = reserved

0x1 = ADVERTISE

0x2 to 0xf = Reserved for future use

The only SLA Event defined in this specification is ADVERTISE.

SLA ID - 16-bits field that specifies identifier which is unique in the scope of Source AS.

The significance of an SLA ID is in the context of the source that is advertising SLA Content. The SLA ID is not globally unique but it MUST be unique within the source AS.

The SLA ID applies to aggregate traffic to prefixes for a given AFI/SAFI that share the same Source AS and SLA ID.

If an advertised SLA ID is different from earlier advertised one, for the same prefix and from the same Source AS, indicates Source AS is advertising new SLA Content to replace the previous one advertised with the same SLA ID.

SLA Length - 12-bits field that specifies the length of the SLA Content. The length is expressed in octets. The SLA Content is optional for an advertised SLA ID. If the SLA Content need not be there, the SLA length field **MUST** be set to zero in such a case.

SLA Content - A variable length field (optional field)

The SLA Content field contains SLA parameters relevant to specified SLA SubType. Since the only defined SLA SubType is **ADVVERTISE**, this specification describes SLA Content only for the **ADVVERTISE** SLA Event.

If SLA Content field exists in a BGP UPDATE message that contains the QoS Attribute with an SLA SubType for SLA Event **ADVVERTISE**, format of the SLA Content is as described in [Section 3.3](#).

If the SLA Content field does not exist, then the advertised message refers to SLA Content advertised in the previous message for the same SLA ID. If there does not exist any prior SLA Content to relate to the advertised SLA ID, then receiver, SLA Consumer, can ignore the SLA advertisement and it may simply update Destination AS count and Destination AS list.

The lack of a valid prior SLA Content field does not make this attribute invalid, so the QoS Attribute **MUST** be forwarded as a valid BGP optional transitive attribute.

[3.3](#). SLA Content for **ADVVERTISE SLA Event**

The only SLA Event described in this specification is **ADVVERTISE**. The format of SLA Content for the **ADVVERTISE** Event is shown in Figure 4.

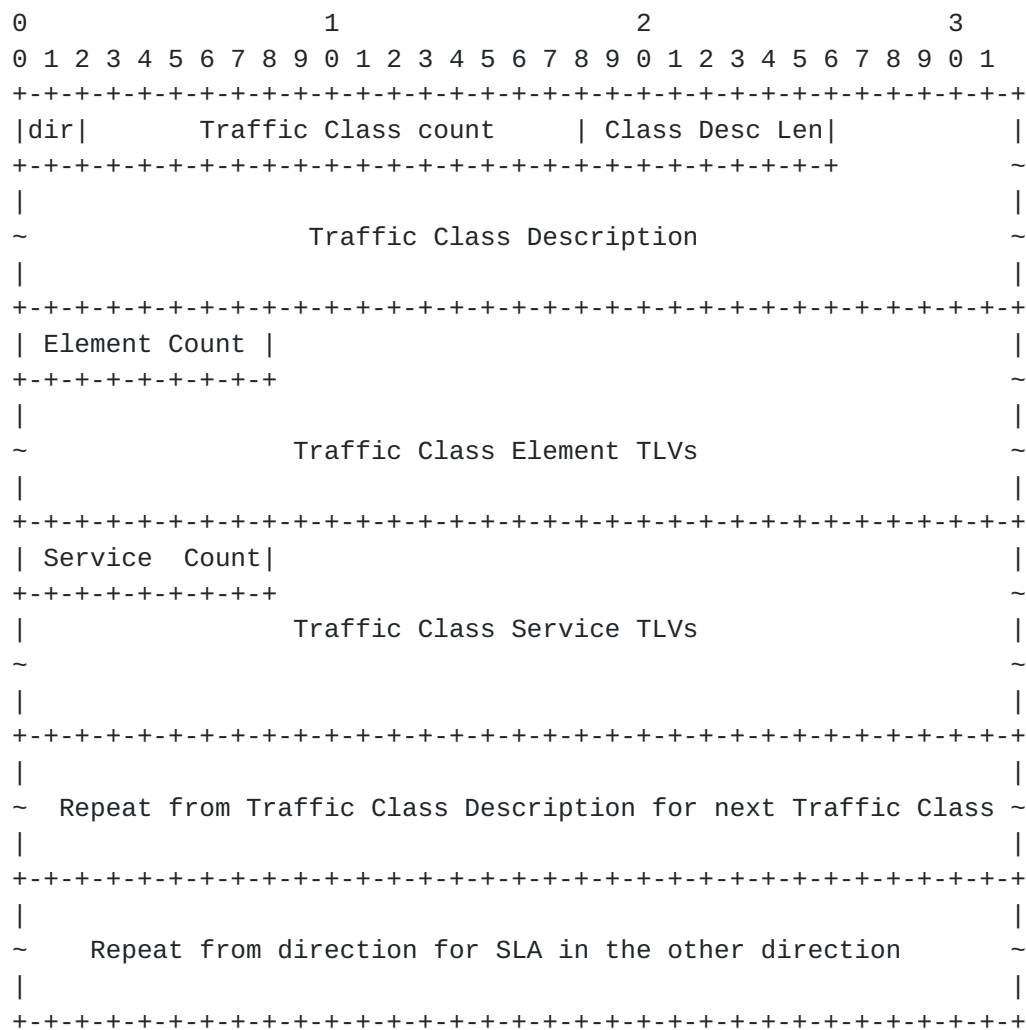


Figure 4 - SLA-Content for ADVERTISE SLA Event

SLA Content contains set of Traffic Class Elements (Classifiers) and Traffic Class Service TLVs for a list of Traffic Classes. This list of Traffic Classes MUST be specified for one direction first and then optionally followed by the specification for the other direction.

dir (Direction) - 2-bits field that specifies Direction of the traffic SLA is applicable to. The following values are defined:

0x0 = reserved

0x1 = incoming, traffic to source AS from destination AS

0x2 = outgoing, traffic from source AS towards destination AS

0x3 = for future use

Traffic Class (Classifier Group) count - 16 bits field that specifies number of Traffic Classes.

The value of zero (0x00) in this field is a special value which means no SLA for the traffic in a specified direction. When Traffic Class count is 0, for a specific direction, the rest of the SLA Content fields MUST NOT be encoded, for that specific direction.

Traffic Class Description Len - 8-bits field that specifies the length of the Traffic Class Description field. The length is expressed in octets.

The value of zero in this field indicates that no Traffic Class Description field follows.

Traffic Class Description - variable length field, as expressed in The Traffic Class Description Len field, MUST carry UTF-8 encoded [\[RFC3629\]](#) description.

Traffic Class Elements (Classifier) Count - 8-bits field that specifies the count of Traffic Class Elements.

The value zero (0x00) means there are no Traffic Class Elements in the traffic class, and thus the Traffic Class is to classify rest of the traffic not captured otherwise by other Traffic Classes in the set for a specified direction.

Traffic Class that has 0 elements MUST be presented last in the advertised list of Traffic Classes for a specific Direction. Otherwise it is considered an error condition which should be handled as described in [Section 6](#).

The QoS Attribute advertised from a specific source MUST NOT have more than one such Traffic Classes (Traffic Class with 0 elements count). If there are more than one such Traffic Classes present then it is an error condition which should be handled as described in [Section 6](#).

Traffic Class Element TLVs - (optional) variable length field holding as many TLVs specified by the Traffic Class Elements Count field. Each TLV has the following format:

IPFIX Element Identifier - 8-bits field that specifies IPFIX Identifiers listed in Table 1.

Length of Value field - 8-bits field that specifies the length, expressed in octets, of the value field.

Value - A variable field that specifies a value appropriate for the IPFIX Element Identifier. It is an error, if the value field does not contain the appropriate format, which should be handled as described in [Section 6](#). Only the IPFIX elements shown in Table 1 are supported.

Any Traffic Class Element advertised in the QoS Attribute only applies to the advertised AFI/SAFI NLRI within the BGP UPDATE message the QoS Attribute is contained in. If a receiver, SLA Consumer, receives a BGP UPDATE message with QoS Attribute for an unsupported AFI/SAFI then SLA Consumer MAY ignore advertised SLA. SLA Consumer MAY update only Destination AS count and Destination AS list, and then QoS Attribute and rest of the BGP UPDATE message MUST be forwarded as per QoS Attribute and BGP protocol specification.

Traffic Class Service Count - 8-bits field that specifies count of Traffic Class Service TLVs.

A value of zero is a special value indicating "no bounded service" (a.k.a., Best Effort (BE)).

Traffic Class Service TLVs - (optional) variable length field with the following format for the TLVs

Traffic Class Service type - 16-bits field that specifies a service type. Each service type is detailed in [Section 3.3.2](#). The list of available service types are,

0x00 = reserved

0x01 = TRAFFIC_CLASS_TSPEC

0x02 = L2_OVERHEAD

0x03 = MINRATE_IN_PROFILE_MARKING

0x04 = MINRATE_OUT_PROFILE_MARKING

0x05 = MAXRATE_IN_PROFILE_MARKING

0x06 = MAXRATE_OUT_PROFILE_MARKING

0x07 = DROP_THRESHOLD

0x08 = RELATIVE_PRIORITY

0x09 = SUB_TRAFFIC_CLASSES

Length of Value field - 08-bits field that specifies the length of the value field. The length of the value is expressed in octets.

Value - a variable length field that specifies the value appropriate for each of the Service Types. It is an error, if this field does not contain the appropriate format, which should be handled as described in [Section 6](#). The format of the value for each of the service types is described in [Section 3.3.2](#)

3.3.1. Supported IPFIX identifiers for Traffic Class Elements

IPFIX [[RFC7012](#)] has well defined identifier set for a large number of packet attributes; an IPFIX IANA registry maintains values for packet classifier attributes (<https://www.ietf.org/assignments/ipfix/ipfix.xml#ipfix-information-elements>). Only the IPFIX attributes listed in Table 1 are supported. Any new attribute to be supported by SLA SubType MUST be a Standards Action as described in IANA section.

+-----+-----+-----+-----+			
ID	Name		
+-----+-----+-----+-----+			
195	ipDiffServCodePoint		
203	mplsTopLabelExp		
244	dot1qPriority		
8	sourceIPv4Address		
27	sourceIPv6Address		
9	sourceIPv4PrefixLength		
29	sourceIPv6PrefixLength		
44	sourceIPv4Prefix		
170	sourceIPv6Prefix		
12	destinationIPv4Address		
28	destinationIPv6Address		
13	destinationIPv4PrefixLength		
30	destinationIPv6PrefixLength		
45	destinationIPv4Prefix		
169	destinationIPv6Prefix		
4	protocolIdentifier		
7	sourceTransportPort		
11	destinationTransportPort		
+-----+-----+-----+-----+			

Table 1

3.3.2. Traffic Class Service types and respective TLVs

3.3.2.1. TRAFFIC_CLASS_TSPEC

The TRAFFIC_CLASS_TSPEC TLV definition:

Type - 0x01

Length - 8-bits field that specifies length, expressed in octets, of the value field. The length of the value field MUST be specified to be 12 octets to hold the value defined as per format below.

Value - TRAFFIC_CLASS_TSPEC value consists of the (r), (b), (p) parameters as described in Invocation Information section of [RFC2212] and shown in Figure 5. Note that inheriting the definition of TSPEC (Traffic SPECification) here does not enable RFC2212 functionality. Only the format of the Traffic Specification is used in this specification.

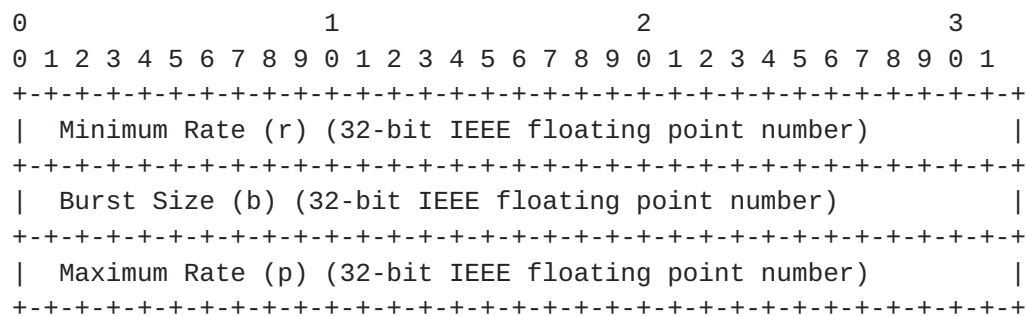


Figure 5 - Traffic Class TSPEC

Format of Parameters (r), (b) and (p): are 32-bit IEEE floating point numbers. Positive infinity is represented as an IEEE single precision floating-point number with an exponent of all ones and a sign mantissa of all zeros. The format of IEEE floating-point numbers is further summarized in [RFC4506].

Parameter (r): indicates min-rate of the traffic class. This rate indicates the minimum rate, measured in octets of Layer 2 (L2) datagrams per second (a.k.a, bytes per second), that the service advertiser is providing for a given class of traffic on advertiser's hop. Note that it does not necessarily translate to a minimum rate service to the receiver of an SLA unless the traffic class definition clearly represents a sole receiver of an SLA. If there is no SLA for min-rate, the value of (r) MUST be set to 0.

Parameter (b): indicates maximum burst size, measured in octets of L2 datagram size. Since queuing delay can be considered a function of burst size (b) and min-rate (r), in presence of non-zero parameter (r), parameter (b) represents bounded delay for the Traffic Class. This delay is a single hop queuing delay when SLA is to be implemented at the resource constrained bottleneck. In other words this burst size can be considered as a buffer size. Value of 0 for parameter (b) means the advertiser does not mandate specific bounded delay.

Parameter (p): indicates max-rate of the traffic class. Just like min-rate, max-rate, measured in octets of L2 datagrams per second (a.k.a., bytes per second), field here also indicates service provided by advertiser. If advertiser does not have any specific value to set for a given class of traffic, it MAY be set to physical interface line rate or any other indirect limit that may affect this class' maximum rate. In absence of any such known value, it MUST be set to positive infinity. Value 0 is considered an error which should be handled as described in [Section 6](#).

3.3.2.2. L2_OVERHEAD

The L2_OVERHEAD TLV definition:

Type - 0x02

Length - 8-bits field that specifies length, expressed in octets, of the value field.

Value - Layer 2 overhead in octets

L2_OVERHEAD defines Layer 2 (L2) specific data in octets, on top of IP datagram size, in a layer 2 frame. By default the rate and burst advertised in the TRAFFIC_CLASS_TSPEC TLV are applicable to the packet size including L2 packet overhead. For the SLA Consumer directly connected to the SLA Producer, this overhead is the L2 overhead of the local link where advertised SLA is received.

However, in cases where advertised SLA is for a SLA Consumer multiple hops away, L2 overhead from the source, SLA Producer, perspective may be different from the local link L2 overhead at the receiver, SLA Consumer. In such cases, explicit notification of size of L2 overhead from the SLA Producer suggests what per packet L2 overhead is applicable to the rate and burst advertised in the TRAFFIC_CLASS_TSPEC TLV. L2_OVERHEAD TLV SHOULD BE ignored by the SLA Consumer if there does not exist TRAFFIC_CLASS_TSPEC TLV for the specified direction.

Advertised L2 value of 0 means SLA advertised is for IP packet size.

3.3.2.3. MINRATE_IN_PROFILE_MARKING

This Traffic Class Service Type defines action performed, by the SLA Producer, on packets that are compliant to the min-rate specified in the TRAFFIC_CLASS_TSPEC TLV. If min-rate specified in the TRAFFIC_CLASS_TSPEC TLV is 0 then TLV for this Traffic Class Service Type SHOULD NOT be advertised. MINRATE_IN_PROFILE_MARKING TLV SHOULD BE ignored by the SLA Consumer if there does not exist TRAFFIC_CLASS_TSPEC TLV for the specified direction, or min-rate specified in the TRAFFIC_CLASS_TSPEC TLV is 0.

The MINRATE_IN_PROFILE_MARKING TLV definition:

Type - 0x03

Length - 8-bits field that specifies length, expressed in octets, of the value field. The length of the value field MUST be specified to be 2 octets to hold the value defined as per format below.

Value - contains the Marking code-point type and value

Marking code-point type - 8-bits IPFIX Element Identifier.

Marking code-point value - 8-bits code-point number.

The marking code-point type of 0x00 is a drop identifier. When marking code-point type value is 0x00 (that is drop), the marking code-point value in this case has no meaning and thus the value in this field should be ignored.

The following table lists the supported IPFIX Identifiers. Any value other than 0 or identifier from the following table is an error condition which should be handled as described in [Section 6](#).

ID	Name
195	ipDiffServCodePoint
203	mplsTopLabelExp
244	dot1qPriority

Table 2

3.3.2.4. MINRATE_OUT_PROFILE_MARKING

This Traffic Class Service Type defines action performed, at the SLA Producer, on packets that are not compliant to the min-rate specified in the TRAFFIC_CLASS_TSPEC TLV. If min-rate specified in the TRAFFIC_CLASS_TSPEC TLV is 0 then TLV for this Traffic Class Service Type SHOULD NOT be advertised. MINRATE_OUT_PROFILE_MARKING TLV SHOULD BE ignored by the SLA Consumer if there does not exist TRAFFIC_CLASS_TSPEC TLV for the specified direction, or min-rate specified in the TRAFFIC_CLASS_TSPEC TLV is 0.

The MINRATE_OUT_PROFILE_MARKING TLV definition:

Type - 0x04

Length - 8-bits field that specifies length, expressed in octets, of the value field. The length of the value field MUST be specified to be 2 octets to hold the value defined as per format below.

Value - contains the Marking code-point type and value

Marking code-point type - 8-bits IPFIX Element Identifier

Marking code-point value - 8-bits code-point number

The marking code-point type of 0x00 is a drop identifier. When marking code-point type value is 0x00 (that is drop), the marking code-point value in this case has no meaning and thus the value in this field should be ignored.

Table 2 lists the supported IPFIX Identifiers. Any value other than 0 or identifier from the Table 2 is an error condition which should be handled as described in [Section 6](#).

3.3.2.5. MAXRATE_IN_PROFILE_MARKING

This Traffic Class Service Type defines action performed, at the SLA Producer, on packets that are compliant to the max-rate specified in the TRAFFIC_CLASS_TSPEC TLV. MAXRATE_IN_PROFILE_MARKING TLV SHOULD BE ignored by the SLA Consumer if there does not exist TRAFFIC_CLASS_TSPEC TLV for the specified direction.

The MAXRATE_IN_PROFILE_MARKING TLV definition:

Type - 0x05

Length - 8-bits field that specifies length, expressed in octets, of the value field. The length of the value field MUST be specified to be 2 octets to hold the value defined as per format below.

Value - contains the Marking code-point type and value

Marking code-point type - 8-bits IPFIX Element Identifier

Marking code-point value - 8-bits code-point number

The marking code-point type of 0x00 is a drop identifier. When marking code-point type value is 0x00 (that is drop), the marking code-point value in this case has no meaning and thus the value in this field should be ignored.

Table 2 lists the supported IPFIX Identifiers. Any value other than 0 or identifier from the Table 2 is an error condition which should be handled as described in [Section 6](#).

3.3.2.6. MAXRATE_OUT_PROFILE_MARKING

This Traffic Class Service Type defines action performed, at the SLA Producer, on packets that are not compliant to the max-rate specified in the TRAFFIC_CLASS_TSPEC TLV. MAXRATE_OUT_PROFILE_MARKING TLV SHOULD BE ignored by the SLA Consumer if there does not exist TRAFFIC_CLASS_TSPEC TLV for the specified direction.

The MAXRATE_OUT_PROFILE_MARKING TLV definition:

Type - 0x06

Length - 8-bits field that specifies length, expressed in octets, of the value field. The length of the value field MUST be specified to be 2 octets to hold the value defined as per format below.

Value - contains the Marking code-point type and value

Marking code-point type - 8-bits IPFIX Element Identifier

Marking code-point value - 8-bits code-point number

The marking code-point type of 0x00 is a drop identifier. When marking code-point type value is 0x00 (that is drop), the marking code-point value in this case has no meaning and thus the value in this field should be ignored.

Table 2 lists the supported IPFIX Identifiers. Any value other than 0 or identifier from the Table 2 is an error condition which should be handled as described in [Section 6](#).

3.3.2.7. Precedence between MINRATE and MAXRATE

The precedence between MINRATE_IN_PROFILE_MARKING, MINRATE_OUT_PROFILE_MARKING, MAXRATE_IN_PROFILE_MARKING, and MAXRATE_OUT_PROFILE_MARKING when all four are advertised is:

- MINRATE_IN_PROFILE_MARKING takes highest precedence (that is over MAXRATE_IN_PROFILE_MARKING),
- MAXRATE_IN_PROFILE_MARKING takes precedence over MINRATE_OUT_PROFILE_MARKING, and
- MAXRATE_OUT_PROFILE_MARKING takes precedence over MINRATE_OUT_PROFILE_MARKING

3.3.2.8. DROP_THRESHOLD

The DROP_THRESHOLD TLV definition:

Type - 0x07

Length - 8-bits field that specifies length, expressed in octets, of the value field.

Value - Count of drop thresholds, followed by content for each drop threshold in the form of (code-point type, count of code-points, list of code-points, threshold value).

Count of drop thresholds - 8-bits field that specifies number of drop thresholds specified in this TLV. Content of each drop threshold is to follow following format

Code-point type - 8-bits IPFIX Element Identifier from the list shown in Table 6.

Count of code-points - 8-bits field that specifies number of code-point values to follow for a specified code-point type.

List of code-points - each code-point value is specified in size of 8 bits and thus total size for this field is 8 bits multiplied by as many number of code-points specified.

Burst value - This is a fixed size 32-bits IEEE floating point number that specifies burst value in unit of bytes.

+-----+-----+-----+-----+-----+	
ID	Name
+-----+-----+-----+-----+-----+	
195	ipDiffServCodePoint
203	mplsTopLabelExp
244	dot1qPriority
+-----+-----+-----+-----+-----+	

Table 3

[3.3.2.9.](#) **RELATIVE_PRIORITY**

The RELATIVE_PRIORITY TLV definition:

Type - 0x08

Length - 8-bits field that specifies length, expressed in octets, of the value field. Given supported range of priority values in this specification, the length of the value field MUST be limited to and thus MUST be specified exactly as 1 octet.

Value - A value from range of 0 - 255. Lower the value means higher the priority

Relative priority indicates scheduling priority of this traffic class. Voice traffic, for example, which requires lowest latency compared to any other traffic, may have lowest value advertised in relative priority. For two different traffic classification groups where one application group may be considered more important than the other but from a scheduling perspective does not require to be distinguished with a different priority, relative priority for those classification groups may be advertised with the same value.

[3.3.2.10.](#) **SUB_TRAFFIC_CLASSES**

The SUB_TRAFFIC_CLASSES TLV definition:

Type - 0x09

Length - 16-bits field that specifies total length, expressed in octets, of a subset of Traffic Class TLVs encoded in the value field

Value - A subset of Traffic Class TLVs

For SLAs where a specific Traffic Class may further be defined by a subset of more granular Traffic Classes, each with its own set of

Traffic Class Elements and Service types definitions,
SUB_TRAFFIC_CLASSES service type SHOULD be used to specify them.

4. Originating SLA Notification

The QoS Attribute for the SLA SubType MUST only be added to the BGP UPDATE message at the node that is SLA Producer. Any QoS Attribute Speaker, in the path to the SLA Consumer MUST NOT modify content of that attribute except modification of the Destination AS list.

QoS Attribute with the SLA SubType SHOULD NOT be advertised periodically just for the purpose of KEEPALIVE between SLA Producer and SLA Consumer. Some sort of SLA policy change, at the SLA Producer, may be considered as a trigger for the advertisement.

For any modified SLA policy at the SLA Producer, SLA Producer MUST re-advertise the entire set of SLA parameters. There is no provision to advertise partial set of SLA parameters. If modified SLA policy is to mean no SLA between SLA Producer and SLA Consumer, then SLA Content MUST be sent with the same SLA ID with the same AS Source and NLRI prefix, as were used to advertise earlier SLA parameters, and the Traffic Class count set to 0.

4.1. SLA Contexts

4.1.1. SLA Advertisement for Point-to-Point Connection

In certain cases, the advertisement of an SLA is intended to relate to aggregate traffic over a point-to-point connection between a specific destination and a specific source. A point-to-point connection may be a physical link or a virtual link (e.g. a tunnel). In such cases, a BGP UPDATE message with source AS number and NLRI prefix as an IP address of an SLA Producer can uniquely identify physical/virtual link in order to establish the context for the advertised SLA for that point to point link.

In the simplest case where Provider (e.g., PE) and Customer (e.g., CE) devices are directly connected via a physical link and have only a single link between them, the CE can uniquely identify the forwarding link to the PE with the following:

- o AS number of the PE,
- o NLRI prefix being an IP address of the PE, that is the next hop address from CE to PE.

The SLA advertised in the QoS Attribute in the BGP UPDATE message sent from the PE to a CE, along with the PE's AS number and PE's IP

address, establishes SLA context for the aggregate traffic through CE-to-PE link.

The SLA advertised in the QoS Attribute in the BGP UPDATE message from PE to CE, with PE's AS number and any other prefix, means SLA for that specific prefix based traffic, a subset of traffic through CE-to-PE link.

Even though this example is in the context of IP prefixes, QoS Attribute's SLA exchange does not have to be limited to the IP address family (IPv4 and IPv6). SLA advertisement is generic to all forms of NLRI types that are supported by the BGP specification (like IPv4, IPv6, VPN-IPv4, VPN-IPv6).

When BGP UPDATE message with the QoS Attribute, containing SLA SubType, is triggered for a point-to-point connection (physical or logical), the Source AS number in the SLA SubType SHOULD BE set to SLA Producer's AS number and destination AS number SHOULD BE set to AS number of BGP peer's that is targeted SLA Consumer. Alternatively, highest order bit in the SLA SubType flags MAY BE set to ignore Source AS and destination AS values from the SLA SubType content since SLA advertised is meant specifically for the BGP peer.

4.1.2. SLA Advertisement for Destination AS Multiple Hops Away

When advertised SLA is not for the BGP peer of an SLA Producer, the Source AS field, in the SLA SubType, MUST be set. The list of destination AS(es) also MUST be set, in the SLA SubType, to avoid flooding of the QoS Attribute data in the network beyond those destinations. Destination AS(es) is a list of SLA Consumers the advertised SLA is intended for.

If a new prefix is learned and traffic with this new prefix is subject to SLA parameters that have already been advertised before for other existing prefixes, then the BGP UPDATE for this new prefix MAY include QoS Attribute containing just an SLA ID that was advertised earlier. This BGP UPDATE message does not require to have the whole SLA Content. The SLA ID is sufficient to relate SLA parameters to new advertised prefixes.

5. QoS Attribute Handling at Forwarding Nodes

The propagation of the QoS Attribute in the BGP UPDATE messages depends on the rules detailed in the following sub-sections.

5.1. BGP Node Capable of Processing QoS Attribute

If a BGP peer is also a QoS Attribute Speaker, it MAY process the QoS Attribute. If BGP UPDATE message has a QoS Attribute with a list of destination ASes, QoS Attribute Speaker MAY trim the list and adjust the count of the destination AS to exclude ones that are not required in further announcement of BGP UPDATE messages.

A QoS Attribute Speaker MUST drop SLA SubType from the QoS Attribute, if there are no more ASes left in the QoS Attribute's destination list. The rest of the QoS Attribute contents may be forwarded if there exist other SubTypes of QoS Attribute and forwarding rules meet other SubTypes requirements. If there is no other SubTypes in that QoS Attribute content then QoS Attribute Speaker MUST drop the entire QoS Attribute all together. BGP Speaker MAY announce further other attributes and NLRI information, if they meet rules defined by other attributes and BGP specification.

Except extracting the entire SLA SubType of the QoS Attribute and trimming the list of Destination AS list, all other content MUST NOT be modified by any QoS Attribute Speaker or BGP Speaker in the path of a BGP UPDATE message.

5.2. QoS Attribute Handling at Receiver

Once QoS Attribute with the SLA SubType is received at intended receiver (SLA Consumer) , processing of advertised SLA Content is optional for the SLA Consumer. SLA Consumer MAY just trim the Destination AS list as per rules described in this specification, without processing any other content of the Attribute. If intended receiver is not a QoS Attribute Speaker than BGP Speaker MUST forward this attribute without any change if rest of the BGP UPDATE message also meets forwarding rules as per BGP specification.

When BGP UPDATE messages are triggered only as a result of SLA policy change, propagating BGP UPDATE message beyond intended SLA Consumers is not necessary. If the SLA Consumer device implementations are capable of policy based filtering, it may implement a policy to filter such BGP UPDATE messages based on prefixes and QoS Attribute containing SLA SubType.

6. Error Handling

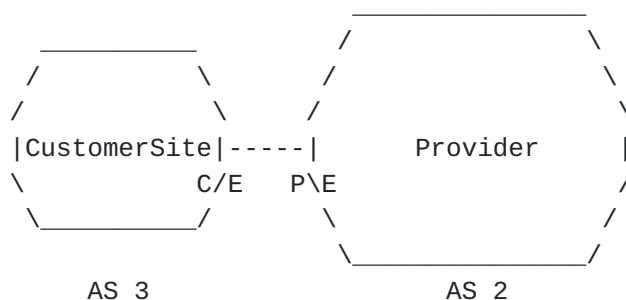
Error conditions, while processing of the QoS Attribute content, MUST be handled with the approach of attribute discard as described in [\[RFC7606\]](#). Processing of QoS Attribute content is done by QoS Attribute Speaker and thus in case of errors, resulting in attribute discard, QoS Attribute Speaker SHOULD convey such indication to the

BGP Speaker and rest of the BGP message SHOULD BE processed by the BGP Speaker as per BGP specification.

7. Deployment Considerations

One of the use cases is for a provider to advertise contracted SLA parameters to a Customer Edge (CE) in cases where eBGP is deployed between PE and CE. The SLA parameters may already be provisioned by the provider on the PE device (facing CE). This provisioned SLA parameters are then advertised thru proposed QoS Attribute to the CE device. The CE device may read the QoS Attribute and SLA SubType content to implement the QoS policy on the device.

Contracted SLA from PE to CE may be full line-rate or sub line-rate or finer granular controlled services. The advertised SLA can be useful when contracted service is sub-rate of a link and/or when for finer granular traffic classes that are controlled (e.g. voice, video services may be capped to certain rate).

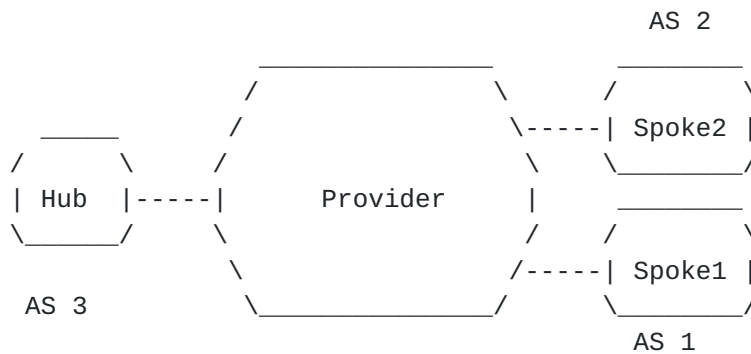


SLA_ADVERTISE: AS2 to AS3

NLRI = PE ip address

Figure 6 - Example 1

Another use case can be to advertise SLAs among different network sites within one Enterprise network. In Hub and Spoke deployments, Administrator may define SLAs at spoke and advertise QoS SLA parameters to the Hub thru BGP updates. In Figure 7, each spoke (AS1 and AS2) are connected to Hub (AS3) via a VPN tunnel. As shown in Figure 7, AS2 can advertise SLA to AS3 in the context of that tunnel ip address.



SLA_ADVERTISE: AS2 to AS3
 NLRI = AS2 tunnel address

SLA_ADVERTISE: AS1 to AS3
 NLRI = AS1 tunnel address

Figure 7 - Example 2

Deployment options are not limited to involving CEs, PE-to-CE or CE-to-CE, only. For any contract between two providers, SLA parameters may be advertised from one to the other.

8. Acknowledgements

Thanks to Fred Baker, David Black, Sue Hares, Benoit Claise and Alvaro Retana for their suggestions and to Christian Jacquenet, Ken Briley, Rahul Patel, Fred Yip, Lou Berger, Brian Carpenter, Bertrand Duvivier, Bruno Decraene for the review.

9. IANA Considerations

This document defines a new BGP optional transitive path attribute, called QoS Attribute. IANA action is required to allocate a new code-point in the BGP path Attributes registry.

IANA is requested to create a registry for QoS Attribute SubTypes. This is a registry of 1 octet value, divided into two pools. One pool of numbers to be assigned on a standards action/early allocation basis. The initial assignments are as shown below. The other pool is for the private use, available range for which is as shown below.

QoS Attribute SubTypes	
=====	
Reserved	0x00
SLA	0x01
Reserved	0x02-0xf0 (Standards Action)
Private use	0xf1-0xff

IANA is requested to create a registry for QoS Attribute SLA SubType flags. This is registry for 8-bits. The initial assignments are as shown below.

QoS Attribute SLA SubType Flags

=====

Highest order bit (bit 0) - to indicate source and destination AS context
Reserved - bits 1 to 15 (Standards Action)

IANA is requested to create a registry for QoS Attribute SLA Event Types. This is a registry of 4-bits value, divided into two pools. One pool of numbers to be assigned on a standards action/early allocation basis. One pool of numbers to be assigned on a standards action/early allocation basis. The initial assignments are as shown below. The other pool is for the private use, available range for which is as shown below.

QoS Attribute SLA Event Types

=====

Reserved	0x0
ADVERTISE	0x1
Reserved	0x2 - 0xc (Standards Action)
Private use	0xd - 0xf

IANA is requested to create a registry to define QoS Attribute SLA Direction. This is the direction in forwarding path, advertised QoS SLA is applicable to. This is a 2-bit registry. Values for QoS Attribute SLA direction are:

QoS Attribute SLA Direction

=====

Reserved	0x0
To source AS from destination AS	0x1
From source AS to destination AS	0x2
Reserved (Standards Action)	0x3

QoS Attribute SLA Traffic Class Element Types will be referring to existing IPFIX IANA types as listed in Table 1. While IPFIX registry is maintained by IANA out of scope of this specification, the use of IPFIX identifiers for this specification are limited to what is described in Table 1. Any new addition of IPFIX identifiers to this table should be a Standards Action.

IANA is requested to create a registry for QoS Attribute SLA Traffic Class Service Types. This is a registry of 2 octet values, to be assigned on a standards action/early allocation basis. The initial assignments are:

Traffic Class Service Type	Value
=====	=====
Reserved	0x00
TRAFFIC_CLASS_TSPEC	0x01
L2_OVERHEAD	0x02
MINRATE_IN_PROFILE_MARKING	0x03
MINRATE_OUT_PROFILE_MARKING	0x04
MAXRATE_IN_PROFILE_MARKING	0x05
MAXRATE_OUT_PROFILE_MARKING	0x06
DROP_THRESHOLD	0x07
RELATIVE_PRIORITY	0x08
SUB_TRAFFIC_CLASSES	0x09
Standards Action	0x0A - 0x3FFF
FCFS	0x4000 - 0x4FF0

10. Security Considerations

BGP security vulnerabilities analysis is documented in [[RFC4272](#)] while BGP-related security considerations are discussed in [[RFC4271](#)]. Also, the reader may refer to [[RFC7132](#)] for more details about BGP path threat model. Rest of the content in this section discusses additional privacy and security considerations that are applicable to the attribute defined in this document.

The information conveyed in the QoS Attribute SLA SubType reveals sensitive data that should not be exposed publicly to non-authorized parties. Deployment considerations mainly target use of QoS Attribute and SLA SubType in managed networks and those where a trust relationship is in place (Customer to Provider, or Provider to Provider). It is NOT RECOMMENDED to enable this attribute at the scale of the Internet unless if means to prevent leaking sensitive information are enforced.

The attribute may be advertised by a misbehaving node to communicate SLA parameters that are not aligned with the SLA agreements. Though the enforcement of SLA parameters is outside the scope of this document, it is RECOMMENDED that the SLA Consumer to enforce a set of validation checks before translating the SLA parameters conveyed in the QoS attributes into provisioning actions. Such validations MAY rely on SLA parameters like the origin AS or SLA ID, like generating SLA ID using pseudo-random schemes [[RFC4086](#)].

Means to prevent route hijacking SHOULD BE considered. Such means include RPKI based origin validation [[RFC7115](#)] and BGP Path validation (e.g., [[I-D.ietf-sidr-bgpsec-protocol](#)]).

11. References

11.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", [BCP 14](#), [RFC 2119](#), DOI 10.17487/RFC2119, March 1997, <<http://www.rfc-editor.org/info/rfc2119>>.
- [RFC2212] Shenker, S., Partridge, C., and R. Guerin, "Specification of Guaranteed Quality of Service", [RFC 2212](#), DOI 10.17487/RFC2212, September 1997, <<http://www.rfc-editor.org/info/rfc2212>>.
- [RFC2434] Narten, T. and H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", [RFC 2434](#), DOI 10.17487/RFC2434, October 1998, <<http://www.rfc-editor.org/info/rfc2434>>.
- [RFC3629] Yergeau, F., "UTF-8, a transformation format of ISO 10646", STD 63, [RFC 3629](#), DOI 10.17487/RFC3629, November 2003, <<http://www.rfc-editor.org/info/rfc3629>>.
- [RFC4271] Rekhter, Y., Ed., Li, T., Ed., and S. Hares, Ed., "A Border Gateway Protocol 4 (BGP-4)", [RFC 4271](#), DOI 10.17487/RFC4271, January 2006, <<http://www.rfc-editor.org/info/rfc4271>>.
- [RFC4272] Murphy, S., "BGP Security Vulnerabilities Analysis", [RFC 4272](#), DOI 10.17487/RFC4272, January 2006, <<http://www.rfc-editor.org/info/rfc4272>>.
- [RFC4506] Eisler, M., Ed., "XDR: External Data Representation Standard", STD 67, [RFC 4506](#), DOI 10.17487/RFC4506, May 2006, <<http://www.rfc-editor.org/info/rfc4506>>.
- [RFC6793] Vohra, Q. and E. Chen, "BGP Support for Four-Octet Autonomous System (AS) Number Space", [RFC 6793](#), DOI 10.17487/RFC6793, December 2012, <<http://www.rfc-editor.org/info/rfc6793>>.
- [RFC7012] Claise, B., Ed. and B. Trammell, Ed., "Information Model for IP Flow Information Export (IPFIX)", [RFC 7012](#), DOI 10.17487/RFC7012, September 2013, <<http://www.rfc-editor.org/info/rfc7012>>.

- [RFC7115] Bush, R., "Origin Validation Operation Based on the Resource Public Key Infrastructure (RPKI)", [BCP 185](#), [RFC 7115](#), DOI 10.17487/RFC7115, January 2014, <<http://www.rfc-editor.org/info/rfc7115>>.
- [RFC7132] Kent, S. and A. Chi, "Threat Model for BGP Path Security", [RFC 7132](#), DOI 10.17487/RFC7132, February 2014, <<http://www.rfc-editor.org/info/rfc7132>>.
- [RFC7606] Chen, E., Ed., Scudder, J., Ed., Mohapatra, P., and K. Patel, "Revised Error Handling for BGP UPDATE Messages", [RFC 7606](#), DOI 10.17487/RFC7606, August 2015, <<http://www.rfc-editor.org/info/rfc7606>>.

11.2. Informative References

- [I-D.ietf-netconf-restconf]
Bierman, A., Bjorklund, M., and K. Watsen, "RESTCONF Protocol", [draft-ietf-netconf-restconf-15](#) (work in progress), July 2016.
- [I-D.ietf-sidr-bgpsec-protocol]
Lepinski, M. and K. Sriram, "BGPsec Protocol Specification", [draft-ietf-sidr-bgpsec-protocol-17](#) (work in progress), June 2016.
- [RFC2475] Blake, S., Black, D., Carlson, M., Davies, E., Wang, Z., and W. Weiss, "An Architecture for Differentiated Services", [RFC 2475](#), DOI 10.17487/RFC2475, December 1998, <<http://www.rfc-editor.org/info/rfc2475>>.
- [RFC4086] Eastlake 3rd, D., Schiller, J., and S. Crocker, "Randomness Requirements for Security", [BCP 106](#), [RFC 4086](#), DOI 10.17487/RFC4086, June 2005, <<http://www.rfc-editor.org/info/rfc4086>>.
- [RFC5575] Marques, P., Sheth, N., Raszuk, R., Greene, B., Mauch, J., and D. McPherson, "Dissemination of Flow Specification Rules", [RFC 5575](#), DOI 10.17487/RFC5575, August 2009, <<http://www.rfc-editor.org/info/rfc5575>>.
- [RFC6020] Bjorklund, M., Ed., "YANG - A Data Modeling Language for the Network Configuration Protocol (NETCONF)", [RFC 6020](#), DOI 10.17487/RFC6020, October 2010, <<http://www.rfc-editor.org/info/rfc6020>>.

- [RFC6241] Enns, R., Ed., Bjorklund, M., Ed., Schoenwaelder, J., Ed., and A. Bierman, Ed., "Network Configuration Protocol (NETCONF)", [RFC 6241](#), DOI 10.17487/RFC6241, June 2011, <<http://www.rfc-editor.org/info/rfc6241>>.
- [RFC7297] Boucadair, M., Jacquenet, C., and N. Wang, "IP Connectivity Provisioning Profile (CPP)", [RFC 7297](#), DOI 10.17487/RFC7297, July 2014, <<http://www.rfc-editor.org/info/rfc7297>>.
- [RFC7674] Haas, J., Ed., "Clarification of the Flowspec Redirect Extended Community", [RFC 7674](#), DOI 10.17487/RFC7674, October 2015, <<http://www.rfc-editor.org/info/rfc7674>>.

Authors' Addresses

Shitanshu Shah
Cisco Systems
170 W. Tasman Drive
San Jose, CA 95134
US

Email: svshah@cisco.com

Keyur Patel
Cisco Systems
170 W. Tasman Drive
San Jose, CA 95134
US

Email: keyupate@cisco.com

Sandeep Bajaj
Juniper Network
1194 N. Mathilda Avenue
Sunnyvale, CA 94089
US

Email: sbajaj@juniper.net

Luis Tomotaki
Verizon
400 International
Richardson, TX 75081
US

Email: luis.tomotaki@verizon.com

Mohamed Boucadair
Orange
Rennes
35000
France

Email: mohamed.boucadair@orange.com

