Stream: Internet Engineering Task Force (IETF)
RFC: 9843
Category: Standards Track
Published: August 2025
ISSN: 2070-1721
Authors:
S. Hegde *Juniper Networks Inc.*
W. Britto *Juniper Networks Inc.*
R. Shetty *Juniper Networks Inc.*
B. Decraene *Orange*
P. Psenak *Cisco Systems*
T. Li *Juniper Networks Inc.*

RFC 9843
IGP Flexible Algorithms: Bandwidth, Delay, Metrics, and Constraints

Abstract

Many networks configure the IGP link metric relative to the link capacity, and high bandwidth traffic gets routed per the link capacity. Flexible Algorithms provide mechanisms to create constraint-based paths in an IGP. This specification documents a generic metric type and a set of bandwidth-related constraints to be used in Flexible Algorithms.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

High bandwidth traffic such as residential Internet traffic and machine-to-machine elephant flows benefit from using high capacity links. Accordingly, many network operators define a link's metric relative to its capacity to help direct traffic to higher bandwidth links, but this is no guarantee that lower bandwidth links will be avoided, especially in failure scenarios. To ensure that elephant flows are only placed on high capacity links, it would be useful to explicitly exclude the high throughput traffic from utilizing links below a certain capacity. Flex-Algorithm [RFC9350] provides a mechanism to create constrained paths by defining a set of parameters consisting of calculation-type, metric-type, and a set of constraints. In this document, we define further extensions to the Flexible Algorithm Definition (FAD) that will allow operators additional control over their traffic flows, especially with respect to bandwidth constraints.

Historically, IGPs have done path computation by minimizing the sum of the link metrics along the path from source to destination. While the metric has been administratively defined, implementations have defaulted to a metric that is inversely proportional to link bandwidth. This has driven traffic to higher bandwidth links and has required manual metric manipulation to achieve the desired loading of the network.

Over time, with the addition of different traffic types, the need for alternate types of metrics has evolved. Flex-Algorithm already supports using the minimum link delay and the administratively assigned traffic-engineering metrics in path computation. However, it is clear that additional metrics may be of interest in different situations. A network operator may seek to minimize their operational costs and thus may want a metric that reflects the actual fiscal costs of using a link. Other traffic may require low jitter, leading to an entirely different set of metrics. With Flex-Algorithm, all of these different metrics, and more, could be used concurrently on the same network.

In some circumstances, path computation constraints, such as administrative groups, can be used to ensure that traffic avoids particular portions of the network. These strict constraints are appropriate when there is an absolute requirement to avoid parts of the topology, even in failure conditions. However, if the requirement is less strict, then using a high metric in a portion of the topology may be more appropriate.

This document defines a family of generic metrics that can advertise various types of administratively assigned metrics. This document proposes standard metric-types that have specific semantics and require to be standardized. This document also specifies user-defined metric-types where specifics are not defined so that administrators are free to assign semantics as they see fit.

In Section 4, this document specifies a new bandwidth-based metric type to be used with Flex-Algorithm and other applications. Section 3 defines additional FAD [RFC9350] constraints that allow the network administrator to preclude the use of low bandwidth links or high delay links.

Section 4.1 defines mechanisms to automatically calculate link metrics based on the parameters defined in the FAD and the advertised Maximum Link Bandwidth of each link. This is advantageous because administrators can change their criteria for metric assignment centrally, without individual modification of each link metric throughout the network. The procedures described in this document are intended to assign a metric to a link based on the total link capacity, and they are not intended to update the metric based on actual traffic flow. Thus, the procedures described in this document are not a replacement to the capability of a PCE [RFC4655], which has a dynamic view of the network and provides real-time bandwidth management or a distributed bandwidth management protocol.

1.1. Requirements Language

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

2. Generic Metric Advertisement

IS-IS [RFC1195] and OSPF [RFC2328] advertise a metric for each link in their respective link state advertisements. Multiple metric types are already supported. Administratively assigned metrics are described in the original OSPF and IS-IS specifications. The Traffic Engineering Default Metric is defined in [RFC5305] and [RFC3630], and the Min Unidirectional delay metric is defined in [RFC8570] and [RFC7471]. Other metrics, such as jitter, reliability, and fiscal cost may be helpful, depending on the traffic class. Rather than attempt to enumerate all possible metrics of interest, this document specifies a generic mechanism for advertising metrics.

Each generic metric advertisement is on a per-link and per-metric-type basis. The metric advertisement consists of a metric type field and a value for the metric. The metric type field has been assigned in the "IGP Metric-Type" IANA registry. Metric types 0-127 are standard metric types as assigned by IANA. This document further specifies a user-defined metric type space of metric types 128-255. These are user defined and can be assigned by an operator for local use.

Implementations **MUST** support sending and receiving generic metric sub-TLV in Application-Specific Link Attributes (ASLA) encodings as well as in the TLV 22/extended link LSA/TE-LSAs. The usage of a generic metric by an individual application is subject to the same rules that apply to other link attributes as defined in [RFC3630], [RFC5305], [RFC9479], [RFC9492], and [RFC9350].

2.1. IS-IS Generic Metric Sub-TLV

The IS-IS Generic Metric sub-TLV specifies the link metric for a given metric type. Typically, this metric is assigned by a network administrator. The Generic Metric sub-TLV is advertised in the TLVs/sub-TLVs below:

a. TLV 22 (Extended IS reachability) [RFC5305]
b. TLV 222 (MT-ISN) [RFC5120]
c. TLV 23 (IS Neighbor Attribute) [RFC5311]
d. TLV 223 (MT IS Neighbor Attribute) [RFC5311]
e. TLV 141 (Inter-AS Reachability Information) [RFC9346]
f. sub-TLV 16 (Application-Specific Link Attributes (ASLA)) of TLV 22/222/23/223/141 [RFC9479]
g. TLV 25 (L2 Bundle Member Attributes) [RFC8668]. Marked as "y(s)" (shareable among bundle members).

The Generic Metric sub-TLV, type 17, is 6 octets in length.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Type        |     Length    |  metric-type  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                         Value                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 1: IS-IS Generic Metric Sub-TLV

where:

Type (1 octet):
An 8-bit field assigned by IANA (17). This value uniquely identifies the Generic Metric TLV.

Length (1 octet):
An 8-bit field indicating the total length, in octets, of the subsequent fields. For this TLV, the Length is set to 4.

Metric-Type (1 octet):
An 8-bit field specifying the type of metric. The value is taken from the "IGP Metric-Type" registry maintained by IANA. The metric type may be any value that is indicated as allowed in the Generic Metric sub-TLV by the "IGP Metric-Type" registry.

Value (3 octets):
A 24-bit unsigned integer representing the metric value. The valid range is from 0 to 16,777,215 (0xFFFFFF).

The Generic Metric sub-TLV **MAY** be advertised multiple times. For a particular metric type, the Generic Metric sub-TLV **MUST** be advertised only once for a link when advertised in TLV 22, 222, 23, 223, and 141. When the Generic Metric sub-TLV is advertised in ASLA, each metric type **MUST** be advertised only once per-application for a link. If there are multiple Generic Metric sub-TLVs advertised for a link for the same metric type (and the same application in case of ASLA) in one or more received Link State Protocol Data Units (LSPDUs), advertisement in the lowest-numbered fragment **MUST** be used, and the subsequent instances **MUST** be ignored.

For a link, if the metric type corresponds to a metric type for which legacy advertisement mechanisms exist (e.g., the IGP Metric, the Min Unidirectional Link Delay, or the Traffic Engineering Default Metric), the legacy metric types **MUST** be utilized from the existing TLV or sub-TLVs. If a Generic Metric advertises a legacy metric, it **MUST** be ignored.

A metric value of 0xFFFFFF is considered a maximum link metric, and a link having this metric value **MUST** be used during Flex-Algorithm calculations as a last resort link as described in Section 15.3 of [RFC9350]. A link can be made unusable by Flex-Algorithm by leaving out Generic Metric advertisement of the particular metric-type that the Flex-Algorithm uses, as described in [RFC9350].

During the router maintenance activity, the Generic Metric for all the links on the node **MAY** be set to a maximum value of 16,777,215 (0xFFFFFF), as it is the maximum usable link metric for the Flex-Algorithm calculations.

2.2. OSPF Generic Metric Sub-TLV

The OSPF Generic Metric sub-TLV specifies the link metric for a given metric type. Typically, this metric is assigned by a network administrator. The Generic Metric sub-TLV is advertised in the TLVs below:

a. sub-TLV of TE Link TLV (2) of OSPF TE LSA [RFC3630].
b. sub-TLV of TE Link TLV (2) of OSPFv2 Inter-AS-TE-v2 LSA [RFC5392].
c. sub-TLV of TE Link TLV (2) of OSPFv3 Intra-Area-TE-LSA [RFC5329].

d. sub-TLV of TE Link TLV (2) of OSPFv3 Inter-AS-TE-v3 LSA [RFC5392].
e. sub-TLV of Application-Specific Link Attributes (ASLA) sub-TLV [RFC9492] of the OSPFv2 Extended Link TLV [RFC7684].
f. sub-TLV of Application-Specific Link Attributes (ASLA) sub-TLV [RFC9492] of the OSPFv3 Router-Link TLV [RFC8362].
g. sub-TLV of the OSPFv2 L2 Bundle Member Attributes sub-TLV [RFC9356].
h. sub-TLV of the OSPFv3 L2 Bundle Member Attributes sub-TLV [RFC9356].

The Generic Metric sub-TLV, types 25/36/34, is eight octets in length.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|              Type             |             Length            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| metric-type   |          Reserved (MBZ)                       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                            Value                              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 2: OSPF Generic Metric Sub-TLV

where:

Type (2 octets):
A 16-bit field assigned by IANA (25/36/34). This value uniquely identifies the Generic Metric TLV.

Length (2 octets):
A 16-bit field indicating the total length, in octets, of the subsequent fields. For this TLV, the Length is set to 8.

Metric-Type (1 octet):
An 8-bit field specifying the type of metric. The value is taken from the "IGP Metric-Type" registry maintained by IANA. The metric type may be any value that is indicated as allowed in the Generic Metric sub-TLV by the "IGP Metric-Type" registry.

Reserved (3 octets):
Must set to zero by the sender and must be ignored by the receiver.

Value (4 octets):
A 32-bit unsigned integer representing the metric value. The valid range is from 0 to 4,294,967,295 (0xFFFFFFFF).

The Generic Metric sub-TLV **MAY** be advertised multiple times. For a particular metric type, the Generic Metric sub-TLV **MUST** be advertised only once for a link when advertised as (a) through (d) above. When the Generic Metric sub-TLV is advertised as a sub-TLV of ASLA, it **MUST** be advertised only once per application for a link. If there are multiple Generic Metric sub-TLVs advertised for a link for the same metric type (and the same application in case of ASLA) in one or more received LSAs, advertisement in the lowest-numbered LSA **MUST** be used, and the subsequent instances **MUST** be ignored.

For a link, if the metric type corresponds to a metric type for which legacy advertisement mechanisms exist (e.g., the IGP Metric, the Min Unidirectional Link Delay, or the Traffic Engineering Default Metric), the legacy metric types **MUST** be utilized from the existing TLV or sub-TLVs. If a Generic Metric advertises a legacy metric, it **MUST** be ignored.

A metric value of 0xFFFFFFFF is considered a maximum link metric, and a link having this metric value **MUST** be used during Flex-Algorithm calculations as a last resort link, as described in Section 15.3 of [RFC9350].

A link can be made unusable by Flex-Algorithm by leaving out Generic Metric advertisement of the particular metric-type that the Flex-Algorithm uses, as described in [RFC9350].

During the router maintenance activity, the Generic Metric for all the links on the node **MAY** be set to a maximum value of 4,294,967,295 (0xFFFFFFFF), as it is the maximum usable link metric for the Flex-Algorithm calculations.

2.3. Generic Metric Applicability to Flexible Algorithm Multi-Domain/Multi-Area Networks

Generic Metric can be used by Flex-Algorithm by specifying the metric type in the Flexible Algorithm Definitions. When Flex-Algorithm is used in a multi-area network, [RFC9350] defines the Flexible Algorithm Prefix Metric (FAPM) sub-TLV that carries the Flexible-Algorithm-specific metric. Metrics carried in FAPM will be equal to the metric to reach the prefix for that Flex-Algorithm in its source area or domain (source area from the Area Border Router (ABR) perspective). When Flex-Algorithm uses Generic Metric, the same procedures as described in Section 13 of [RFC9350] are used to send and process the FAPM sub-TLV.

3. FAD Constraint Sub-TLVs

Large high throughput flows are referred to as "elephant flows". Directing an elephant flow down a low-bandwidth link might congest the link and cause other critical application traffic flowing on the link to drop. Thus, in the context of Flex-Algorithm, it would be useful to be able to constrain the topology to only those links capable of supporting a minimum amount of bandwidth.

If the capacity of a link is constant, this can already be achieved through the use of administrative groups. However, when a Layer 3 link is actually a collection of Layer 2 links (Link Aggregation Group (LAG) / Layer 2 Bundle), the link bandwidth will vary based on the set of active constituent links. This could be automated by having an implementation vary the advertised administrative groups based on bandwidth, but this seems unnecessarily complex and expressing this requirement as a direct constraint on the topology seems simpler. This is also advantageous if the minimum required bandwidth changes, as this constraint would provide a single centralized, coordinated point of control.

To satisfy this requirement, this document defines an Exclude Minimum Bandwidth constraint. When this constraint is advertised in a FAD, a link will be pruned from the Flex-Algorithm topology if the link's advertised Maximum Link Bandwidth is below the advertised Minimum Bandwidth value.

Similarly, this document defines an Exclude Maximum Link Delay constraint. Applications, such as High-Frequency Trading are sensitive to link delays and may perform poorly in networks prone to delay variability, such as those with transparent Layer 2 link recovery mechanisms or satellite links. Mechanisms already exist to measure the link delay dynamically and advertise it in the IGP. Networks that employ dynamic link-delay measurement, may want to exclude links that have a delay over a given threshold.

3.1. IS-IS FAD Constraint Sub-TLVs

3.1.1. IS-IS Exclude Minimum Bandwidth Sub-TLV

IS-IS Flex-Algorithm Exclude Minimum Bandwidth (FAEMB) sub-TLV is a sub-TLV of the IS-IS FAD sub-TLV. It has the following format:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|      Type     |    Length     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      Min Bandwidth                            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 3: IS-IS FAEMB Sub-TLV

where:

Type (1 octet):
: An 8-bit field assigned by IANA (6). This value uniquely identifies the FAEMB sub-TLV.

Length (1 octet):
: An 8-bit field indicating the total length, in octets, of the subsequent fields. For this sub-TLV, the Length is set to 4.

Min Bandwidth (4 octets):
: A 32-bit field specifying the link bandwidth encoded in IEEE floating point format (32 bits) [IEEE754-2019]. The units are bytes per second.

The FAEMB sub-TLV **MUST** appear once at most in the FAD sub-TLV. If it appears more than once, the IS-IS FAD sub-TLV **MUST** be ignored by the receiver.

The Minimum bandwidth advertised in the FAEMB sub-TLV **MUST** be compared with Maximum Link Bandwidth advertised in sub-sub-TLV 9 of the ASLA sub-TLV [RFC9479]. If the L-flag is set in the ASLA sub-TLV, the Minimum bandwidth advertised in the FAEMB sub-TLV **MUST** be compared with the Maximum Link Bandwidth as advertised in the sub-TLV 9 of the TLV 22/222/23/223/141 [RFC5305], as defined in Section 4.2 of [RFC9479].

If the Maximum Link Bandwidth is lower than the Minimum Link Bandwidth advertised in the FAEMB sub-TLV, the link **MUST** be excluded from the Flex-Algorithm topology. If a link does not have the Maximum Link Bandwidth advertised but the FAD contains this sub-TLV, then that link **MUST NOT** be excluded from the topology based on the Minimum Bandwidth constraint.

3.1.2. IS-IS Exclude Maximum Delay Sub-TLV

IS-IS Flex-Algorithm Exclude Maximum Delay (FAEMD) sub-TLV is a sub-TLV of the IS-IS FAD sub-TLV. It has the following format:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|      Type     |    Length     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Max Link Delay            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 4: IS-IS FAEMD Sub-TLV

where:

Type (1 octet):
: An 8-bit field assigned by IANA (7). This value uniquely identifies the FAEMD sub-TLV.

Length (1 octet):
: An 8-bit field indicating the total length, in octets, of the subsequent fields. For this sub-TLV, the Length is set to 3.

Max Link Delay (3 octets):
: A 24-bit field specifying the Maximum link delay in microseconds.

The FAEMD sub-TLV **MUST** appear only once in the FAD sub-TLV. If it appears more than once, the IS-IS FAD sub-TLV **MUST** be ignored by the receiver.

The Maximum link delay advertised in the FAEMD sub-TLV **MUST** be compared with Min Unidirectional Link Delay advertised in sub-sub-TLV 34 of the ASLA sub-TLV [RFC9479]. If the L-flag is set in the ASLA sub-TLV, the Maximum link delay advertised in the FAEMD sub-TLV **MUST** be compared with Min Unidirectional Link Delay as advertised by the sub-TLV 34 of the TLV 22/222/23/223/141 [RFC8570], as defined in Section 4.2 of [RFC9479].

If the Min Unidirectional Link Delay value is higher than the Maximum link delay advertised in the FAEMD sub-TLV, the link **MUST** be excluded from the Flex-Algorithm topology. If a link does not have the Min Unidirectional Link Delay advertised but the FAD contains this sub-TLV, then that link **MUST NOT** be excluded from the topology based on the Maximum Delay constraint.

3.2. OSPF FAD Constraint Sub-TLVs

3.2.1. OSPF Exclude Minimum Bandwidth Sub-TLV

OSPF Flex-Algorithm Exclude Minimum Bandwidth (FAEMB) sub-TLV is a sub-TLV of the OSPF FAD TLV. It has the following format:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|      Type                     |    Length                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      Min Bandwidth                            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 5: OSPF FAEMB Sub-TLV

where:

Type (2 octets):
: A 16-bit field assigned by IANA (6). This value uniquely identifies the OSPF FAEMB sub-TLV.

Length (2 octets):
: A 16-bit field indicating the total length, in octets, of the subsequent fields. For this sub-TLV, the Length is set to 4.

Min Bandwidth (4 octets):
: A 32-bit field specifying the link bandwidth encoded in IEEE floating point format (32 bits) [IEEE754-2019]. The units are bytes per second.

The FAEMB sub-TLV **MUST** only appear once in the FAD sub-TLV. If it appears more than once, the OSPF FAD TLV **MUST** be ignored by the receiver.

The Maximum Link Bandwidth as advertised in the Extended Link TLV in the Extended Link Opaque LSA in OSPFv2 [RFC7684] or as a sub-TLV of the Router-Link TLV of the E-Router-LSA Router-Link TLV in OSPFv3 [RFC8362] **MUST** be compared against the Minimum bandwidth advertised in the FAEMB sub-TLV. If the link bandwidth is lower than the Minimum bandwidth advertised in the FAEMB sub-TLV, the link **MUST** be excluded from the Flex-Algorithm topology.

If a link does not have the Maximum Link Bandwidth advertised but the FAD contains this sub-TLV, then that link **MUST** be included in the topology and proceed to apply further pruning rules for the link.

3.2.2. OSPF Exclude Maximum Delay Sub-TLV

The OSPF Flex-Algorithm Exclude Maximum Delay (FAEMD) sub-TLV is a sub-TLV of the OSPF FAD TLV. It has the following format.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|      Type                     |    Length                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  RESERVED     |                     Max Link Delay            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 6: OSPF FAEMD Sub-TLV

where:

Type (2 octets):
: A 16-bit field assigned by IANA (7). This value uniquely identifies the OSPF FAEMD sub-TLV.

Length (2 octets):
: A 16-bit field indicating the total length, in octets, of the subsequent fields. For this sub-TLV, the Length is set to 4.

Reserved (1 octet):
: Must be set to zero by the sender and must be ignored by the receiver.

Max Link Delay (3 octets):
: A 24-bit field specifying the Maximum link delay in microseconds.

The FAEMD sub-TLV **MUST** only appear once in the OSPF FAD TLV. If it appears more than once, the OSPF FAD TLV **MUST** be ignored by the receiver.

The Min Delay value advertised via the Min/Max Unidirectional Link Delay of the ASLA sub-TLV [RFC9492] **MUST** be compared against the Maximum delay advertised in the FAEMD sub-TLV. If the Min Unidirectional Link Delay is higher than the Maximum delay advertised in the FAEMD sub-TLV, the link **MUST** be excluded from the Flex-Algorithm topology. If the Min/Max Unidirectional Link Delay is not advertised for a link but the FAD contains this sub-TLV, then that link **MUST NOT** be excluded from the topology based on the Maximum Delay constraint.

4. Bandwidth Metric Advertisement

Historically, IGP implementations have made default metric assignments based on link bandwidth. This has proven to be useful but has suffered from having different defaults across implementations and from the rapid growth of link bandwidths. With Flex-Algorithm, the network administrator can define a function that will produce a metric for each link and have each node automatically compute each link's metric based on its bandwidth.

This document defines a standard metric type for this purpose called the "Bandwidth Metric". The Bandwidth Metric **MAY** be advertised in the Generic Metric sub-TLV with the metric type set to "Bandwidth Metric". IS-IS and OSPF will advertise this type of metric in their link advertisements. The Bandwidth Metric is a link attribute and for the advertisement and processing of this attribute for Flex-Algorithm, **MUST** follow Section 12 of [RFC9350].

Flex-Algorithm uses this metric type by specifying the bandwidth metric as the metric type in a FAD TLV. A FAD TLV may also specify an automatic computation of the bandwidth metric based on a link's advertised bandwidth. An explicit advertisement of a link's bandwidth metric using the Generic Metric sub-TLV overrides this automatic computation. The automatic bandwidth metric calculation sub-TLVs are advertised in the FAD TLV, and these parameters are applicable to applications such as Flex-Algorithm that make use of the FAD TLV.

4.1. Automatic Metric Calculation

Networks that are designed to be highly regular and that follow uniform metric assignment may want to simplify their operations by automatically calculating the bandwidth metric. When a FAD advertises the metric type as Bandwidth Metric and the link does not have the Bandwidth Metric advertised, automatic metric derivation can be used with additional FAD constraint advertisement as described in this section.

If a link's bandwidth changes, then the delay in learning about the change may create the possibility of micro-loops in the topology. This is no different from the IGP's susceptibility to micro-loops during a metric change. The micro-loop avoidance procedures described in [SR-LOOP-AVOID] or any other mechanism as described in the framework [RFC5715] can be used to avoid micro-loops when the automatic metric calculation is deployed.

Computing the metric between adjacent systems based on bandwidth becomes more complex in the case of parallel adjacencies. If there are parallel adjacencies between systems, then the bandwidth between the systems is the sum of the bandwidth of the parallel links. This is somewhat more complex to deal with, so there is an optional mode for computing the aggregate bandwidth.

4.1.1. Automatic Metric Calculation Modes

4.1.1.1. Simple Mode

In simple mode, the Maximum Link Bandwidth of a single Layer 3 link is used to derive the metric. This mode is suitable for deployments that do not use parallel Layer 3 links. In this case, the computation of the metric is straightforward. If a Layer 3 link is composed of a Layer 2 bundle, then the link bandwidth is the sum of the bandwidths of the working components and may vary with Layer 2 link failures.

4.1.1.2. Interface Group Mode

The simple mode of metric calculation may not work well when there are multiple parallel Layer 3 interfaces between two nodes. Ideally, the metric between two systems should be the same given the same bandwidth, whether the bandwidth is provided by parallel Layer 2 links or parallel Layer 3 links. To address this, in Interface Group Mode, nodes **MUST** compute the aggregate bandwidth of all parallel adjacencies, **MUST** derive the metric based on the aggregate bandwidth, and **MUST** apply the resulting metric to each of the parallel adjacencies. Note that a single elephant flow is normally pinned to a single Layer 3 interface. If the single Layer 3 link bandwidth is not sufficient for any single elephant flow, the mechanisms to solve this issue are outside the scope of this document.

```
        A------B====C====F====D
               |              |
                ------E-------
```

Figure 7: Parallel Interfaces

For example, in the above diagram, there are two parallel links between B->C, C->F, F->D. Let us assume the link bandwidth is uniform 10 Gbps on all links. When bandwidth is used to derive the metric for the links, the metric for each link will be the same. Traffic from B to D will be forwarded as B->E->D because the metric will be lower. Since the bandwidth is higher on the B->C->F->D path, the metric for that path should be lower than the B->E->D path to attract the traffic on the B->C->F->D path. Interface Group Mode should be preferred in cases where there are parallel Layer 3 links.

In the Interface Group Mode, every node **MUST** identify the set of parallel links between a pair of nodes based on IGP link advertisements and **MUST** consider cumulative bandwidth of the parallel links while arriving at the metric of each link.

The parallel Layer 3 links between two nodes may not have the same bandwidth. In such cases, the method described in Interface Group Mode will result in the same metric being used for all the parallel links, which may cause undesired load balancing on the links. In such cases, a device may locally apply a load-balancing factor relative to the link bandwidth on the ECMP next hops. The load-balancing mechanisms are outside the scope of this document.

4.1.2. Automatic Metric Calculation Methods

In automatic metric calculation for simple and Interface Group Mode, Maximum Link Bandwidth of the links is used to derive the metric. There are two types of automatic metric derivation methods.

1. Reference bandwidth method
2. Bandwidth thresholds method

4.1.2.1. Reference Bandwidth Method

In many networks, the metric is inversely proportional to the link bandwidth. The administrator or implementation selects a reference bandwidth and the metric is derived by dividing the reference bandwidth by the advertised Maximum Link Bandwidth. Advertising the reference bandwidth in the FAD constraints allows the metric computation to be done on every node for each link. The metric is computed using reference bandwidth and the advertised link bandwidth. Centralized control of this reference bandwidth simplifies management in the case where the reference bandwidth changes. In order to ensure that small bandwidth changes do not change the link metric, it is useful to define the granularity of the bandwidth that is of interest. The link bandwidth will be truncated to this granularity before deriving the metric.

For example,

reference bandwidth = 1000G

Granularity = 20G

The derived metric is 10 for link bandwidth in the range 100G to 119G

4.1.2.2. Bandwidth Thresholds Method

The reference bandwidth approach described above provides a uniform metric value for a range of link bandwidths. In certain cases, there may be a need to define non-proportional metric values for the varying ranges of link bandwidth. For example, bandwidths from 10G to 30G are assigned metric value 100, bandwidth from 30G to 70G are assigned a metric value of 50, and bandwidths greater than 70G have a metric of 10. In order to support this, a staircase mapping based on bandwidth thresholds is supported in the FAD. This advertisement contains a set of threshold values and associated metrics.

4.1.3. IS-IS FAD Constraint Sub-TLVs for Automatic Metric Calculation

4.1.3.1. Reference Bandwidth Sub-TLV

This section provides FAD constraint advertisement details for the reference bandwidth method of metric calculation, as described in Section 4.1.2.1. The Flexible Algorithm Definition Reference Bandwidth (FADRB) sub-TLV is a sub-TLV of the IS-IS FAD sub-TLV. It has the following format:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|      Type     |    Length     |     Flags     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      Reference Bandwidth                      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      Granularity Bandwidth                    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 8: IS-IS FADRB Sub-TLV

where:

Type (1 octet):
An 8-bit field assigned by IANA (8). This value uniquely identifies the IS-IS FADRB sub-TLV.

Length (1 octet):
An 8-bit field indicating the total length, in octets, of the subsequent fields. For this sub-TLV, the Length is set to 9.

Flags (1 octet):
An 8-bit field containing flags.

```
 0 1 2 3 4 5 6 7
+-+-+-+-+-+-+-+-+
|G|             |
+-+-+-+-+-+-+-+-+
```

G-flag:
When set, Interface Group Mode **MUST** be used to derive total link bandwidth. Unassigned bits **MUST** be set to zero.

Reference Bandwidth (4 octets):
A 32-bit field with Bandwidth encoded in IEEE floating point format [IEEE754-2019]. The units are bytes per second.

Granularity Bandwidth (4 octets):
A 32-bit field with Bandwidth encoded in IEEE floating point format [IEEE754-2019]. The units are bytes per second.

When granularity_bw is less than or equal to Total_link_bandwidth, then:

Metric calculation: (Reference_bandwidth) / (Total_link_bandwidth - (Modulus of(Total_link_bandwidth, granularity_bw)))

When granularity_bw is greater than Total_link_bandwidth, then:

Metric calculation: Reference_bandwidth / Total_link_bandwidth

The division used here is integer division. Modulus of operation is defined as a remainder value when two numbers are divided.

The Granularity Bandwidth value ensures that the metric does not change when there is a small change in the link bandwidth. The IS-IS FADRB sub-TLV **MUST NOT** appear more than once in an IS-IS FAD sub-TLV. If it appears more than once, the IS-IS FAD sub-TLV **MUST** be ignored by the receiver. The value advertised in the Reference Bandwidth field **MUST** be non-zero. If a zero value is advertised in the Reference Bandwidth field in the IS-IS FADRB sub-TLV, the sub-TLV **MUST** be ignored.

If a Generic Metric sub-TLV with a Bandwidth metric type is advertised for a link, the Flex-Algorithm calculation **MUST** use the advertised Bandwidth Metric and **MUST NOT** use the automatically derived metric for that link. In case of Interface Group Mode, if all the parallel links have been advertised with the Bandwidth Metric, the individual link Bandwidth Metric **MUST** be used. If only some links among the parallel links have the Bandwidth Metric advertisement, the Bandwidth Metric for such links **MUST** be ignored, and automatic Metric calculation **MUST** be used to derive link metric.

If the calculated metric evaluates to zero, a metric of 1 **MUST** be used.

If the calculated metric evaluates to a number greater than 0xFFFFFF, it is set to 0xFFFFFF.

4.1.3.2. Bandwidth Threshold Sub-TLV

This section provides FAD constraint advertisement details for the Bandwidth Thresholds method of metric calculation as described in Section 4.1.2.2. The Flexible Algorithm Definition Bandwidth Threshold (FADBT) sub-TLV is a sub-TLV of the IS-IS FAD sub-TLV. It has the following format:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     Type      |    Length     |     Flags     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Bandwidth Threshold 1                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Threshold Metric 1        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Bandwidth Threshold 2                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Threshold Metric 2        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Bandwidth Threshold 3                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
                               .....
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Threshold Metric N-1      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Bandwidth Threshold N                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Threshold Metric N        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 9: IS-IS FADBT Sub-TLV

where:

Type (1 octet):
: An 8-bit field assigned by IANA (9). This value uniquely identifies the IS-IS FADBT sub-TLV.

Length (1 octet):
: An 8-bit field indicating the total length, in octets, of the subsequent fields. For this sub-TLV, the Length is calculated as (1+n*7). Here, N is equal to the number of Threshold Metrics specified. n **MUST** be greater than or equal to 1.

Flags (1 octet):

```
 0 1 2 3 4 5 6 7
+-+-+-+-+-+-+-+-+
|G| | |         |
+-+-+-+-+-+-+-+-+
```

G-flag: When set, Interface Group Mode **MUST** be used to derive total link bandwidth.

Unassigned bits **MUST** be set to zero.

Following is the staircase bandwidth threshold and associated metric values.

Bandwidth Threshold 1 (4 octets):
Minimum Link Bandwidth is encoded in IEEE floating point format (32 bits)[IEEE754-2019]. The units are bytes per second.

Threshold Metric 1 (3 octets):
Metric value range (1 - 16,777,215 (0xFFFFFF))

Bandwidth Threshold n (4 octets):
Maximum Link Bandwidth is encoded in IEEE floating point format (32 bits)[IEEE754-2019]. The units are bytes per second.

Threshold Metric n (3 octets):
Metric value range (1 - 16,777,215 (0xFFFFFF))

When the G-flag is set, the cumulative bandwidth of the parallel links is computed as described in Section 4.1.1.2. If the G-flag is not set, the advertised Maximum Link Bandwidth is used.

The assignment of the Bandwidth Metric based on computed link bandwidth is described below.

The Bandwidth Metric for a link during the Flex-Algorithm Shortest Path First (SPF) calculation **MUST** be assigned according to the following rules:

1. When the computed link bandwidth is less than Bandwidth Threshold 1:

   The Bandwidth Metric **MUST** be set to the maximum metric value of 4,261,412,864.

2. When the computed link bandwidth is greater than or equal to Bandwidth Threshold 1 and less than Bandwidth Threshold 2:

   The Bandwidth Metric **MUST** be set to Threshold Metric 1.

3. When the computed link bandwidth is greater than or equal to Bandwidth Threshold 2 and less than Bandwidth Threshold 3:

   The Bandwidth Metric **MUST** be set to Threshold Metric 2.

4. In general, for all integer values of X such that $1 \le X < N$:

   When the computed link bandwidth is greater than or equal to Bandwidth Threshold X and less than Bandwidth Threshold X+1:

   The Bandwidth Metric **MUST** be set to Threshold Metric X.

5. When the computed link bandwidth is greater than or equal to Bandwidth Threshold N:

   The Bandwidth Metric **MUST** be set to Threshold Metric N.

Notes:

- The term "Bandwidth Threshold X" refers to a predefined threshold value corresponding to the index X.
- The term "Threshold Metric X" refers to the metric value associated with Bandwidth Threshold X.
- N represents the total number of bandwidth thresholds defined in the system.

Implementations **MUST** ensure that these rules are consistently applied to maintain interoperability and optimal path computation within the network.

The IS-IS FADBT sub-TLV **MUST NOT** appear more than once in an IS-IS FAD sub-TLV. If it appears more than once, the IS-IS FAD sub-TLV **MUST** be ignored by the receiver.

A FAD **MUST NOT** contain both the FADBT sub-TLV and the FADRB sub-TLV. If both of these sub-TLVs are advertised in the same FAD for a Flexible Algorithm, the FAD **MUST** be ignored by the receiver.

If a Generic Metric sub-TLV with Bandwidth metric type is advertised for a link, the Flex-Algorithm calculation **MUST** use the Bandwidth Metric advertised on the link and **MUST NOT** use the automatically derived metric for that link.

In case of Interface Group Mode, if all the parallel links have been advertised with the Bandwidth Metric, the individual link Bandwidth Metric **MUST** be used. If only some links among the parallel links have advertised the Bandwidth Metric, the Bandwidth Metric for such links **MUST** be ignored and automatic Metric calculation **MUST** be used to derive the link metric.

4.1.4. OSPF FAD Constraint Sub-TLVs for Automatic Metric Calculation

4.1.4.1. Reference Bandwidth Sub-TLV

The Flexible Algorithm Definition Reference Bandwidth (FADRB) sub-TLV is a sub-TLV of the OSPF FAD TLV. It has the following format:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|      Type                     |    Length                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    Flags   |     Reserved                                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      Reference Bandwidth                      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      Granularity Bandwidth                    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 10: OSPF FADRB Sub-TLV

where:

Type (2 octets):
A 16-bit field assigned by IANA (8). This value uniquely identifies the OSPF FADRB sub-TLV.

Length (2 octets):
A 16-bit field indicating the total length, in octets, of the subsequent fields. For this sub-TLV, Length is set to 14.

Flags (1 octet):

```
 0 1 2 3 4 5 6 7
+-+-+-+-+-+-+-+-+
|G| | |         |
+-+-+-+-+-+-+-+-+
```

G-flag: When set, Interface Group Mode **MUST** be used to derive total link bandwidth.

Unassigned bits **MUST** be set to zero.

Reference Bandwidth (4 octets):
Bandwidth encoded in 32 bits in IEEE floating point format [IEEE754-2019]. The units are in bytes per second.

Granularity Bandwidth (4 octets):
Bandwidth encoded in 32 bits in IEEE floating point format [IEEE754-2019]. The units are in bytes per second.

When granularity_bw is less than or equal to Total_link_bandwidth, then:

Metric calculation:
(Reference_bandwidth) / (Total_link_bandwidth - (Modulus of(Total_link_bandwidth, granularity_bw)))

When granularity_bw is greater than Total_link_bandwidth, then:

Metric calculation:
Reference_bandwidth/ Total_link_bandwidth

The division used here is integer division.

Modulus of operation is defined as a remainder value when two numbers are divided.

The Granularity Bandwidth value is used to ensure that the metric does not change when there is a small change in the link bandwidth. The OSPF FADRB sub-TLV **MUST NOT** appear more than once in an OSPF FAD TLV. If it appears more than once, the OSPF FAD TLV **MUST** be ignored by the receiver. The value advertised in the Reference Bandwidth field **MUST** be non-zero. If a zero value is advertised in the Reference Bandwidth field in the OSPF FADRB sub-TLV, the sub-TLV **MUST** be ignored. If a Generic Metric sub-TLV with Bandwidth metric type is advertised for a link, the Flex-Algorithm calculation **MUST** use the advertised Bandwidth Metric on the link and **MUST NOT** use the automatically derived metric for that link. In the case of Interface Group Mode, the following procedures apply:

- When all parallel links have advertised the Bandwidth Metric: The individual link Bandwidth Metric **MUST** be used for each link.

- When only a subset of the parallel links have advertised the Bandwidth Metric: The Bandwidth Metric advertisements for those links **MUST** be ignored. In this scenario, automatic metric calculation **MUST** be used to derive the link metrics for all parallel links.

If the calculated metric evaluates to zero, a metric of 1 **MUST** be used.

If the calculated metric evaluates to a number greater than 0xFFFFFFFF, it is set to 0xFFFFFFFF.

4.1.4.2. Bandwidth Threshold Sub-TLV

The Flexible Algorithm Definition Bandwidth Threshold (FADBT) sub-TLV is a sub-TLV of the OSPF FAD TLV. It has the following format:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|      Type                     |    Length                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    Flags   | Reserved                                         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      Bandwidth Threshold 1                    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      Threshold Metric 1                       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      Bandwidth Threshold 2                    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      Threshold Metric 2                       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      Bandwidth Threshold 3                    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
                              .....
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      Threshold Metric N-1                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      Bandwidth Threshold N                    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      Threshold Metric N                       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 11: OSPF FADBT Sub-TLV

where:

Type (2 octets):
: A 16-bit field assigned by IANA (9). This value uniquely identifies the OSPF FADBT sub-TLV.

Length (2 octets):
: A 16-bit field indicating the total length, in octets, of the subsequent fields. For this sub-TLV, Length is set to 2 + N*8 octets. Here, N is equal to the number of Threshold Metrics specified. N **MUST** be greater than or equal to 1.

Flags (1 octet):

```
 0 1 2 3 4 5 6 7
+-+-+-+-+-+-+-+-+
|G|             |
+-+-+-+-+-+-+-+-+
```

G-flag: When set, Interface Group Mode **MUST** be used to derive total link bandwidth.

Unassigned bits **MUST** be set to zero.

Following is the staircase bandwidth threshold and associated metric values.

Bandwidth Threshold 1 (4 octets):
Minimum Link Bandwidth is encoded in IEEE floating point format (32 bits)[IEEE754-2019]. The units are bytes per second.

Threshold Metric 1 (4 octets):
Metric value range (1 - 4,294,967,296 (0xFFFFFFFF))

Bandwidth Threshold N (4 octets):
Maximum Link Bandwidth is encoded in IEEE floating point format (32 bits)[IEEE754-2019]. The units are bytes per second.

Threshold Metric N (4 octets):
Metric value range (1 - 4,294,967,296 (0xFFFFFFFF))

When the G-flag is set, the cumulative bandwidth of the parallel links is computed as described in Section 4.1.1.2. If the G-flag is not set, the advertised Maximum Link Bandwidth is used.

The assignment of the Bandwidth Metric based on computed link bandwidth is described below.

During the Flex-Algorithm SPF calculation, the Bandwidth Metric for a link **MUST** be assigned according to the following rules:

1. When the computed link bandwidth is less than Bandwidth Threshold 1:

   The Bandwidth Metric **MUST** be set to the maximum metric value of 4,294,967,296.

2. When the computed link bandwidth is greater than or equal to Bandwidth Threshold 1 and less than Bandwidth Threshold 2:

   The Bandwidth Metric **MUST** be set to Threshold Metric 1.

3. When the computed link bandwidth is greater than or equal to Bandwidth Threshold 2 and less than Bandwidth Threshold 3:

   The Bandwidth Metric **MUST** be set to Threshold Metric 2.

4. In general, for all integer values of X where 1 ≤X < N:

When the computed link bandwidth is greater than or equal to Bandwidth Threshold X and less than Bandwidth Threshold X+1:

The Bandwidth Metric **MUST** be set to Threshold Metric X.

5. When the computed link bandwidth is greater than or equal to Bandwidth Threshold N:

   The Bandwidth Metric **MUST** be set to Threshold Metric N.

Notes:

- Bandwidth Threshold X refers to the predefined bandwidth threshold corresponding to index X.
- Threshold Metric X is the metric value associated with Bandwidth Threshold X.
- N represents the total number of bandwidth thresholds defined in the system.

Implementations **MUST** consistently apply these rules to ensure accurate path computations and maintain interoperability within the network.

The OSPF FADBT sub-TLV **MUST NOT** appear more than once in an OSPF FAD sub-TLV. If it appears more than once, the OSPF FAD sub-TLV **MUST** be ignored by the receiver.

A FAD **MUST NOT** contain both the FADBT sub-TLV and the FADRB sub-TLV. If both these sub-TLVs are advertised in the same FAD for a Flexible Algorithm, the FAD **MUST** be ignored by the receiver.

If a Generic Metric sub-TLV with Bandwidth metric type is advertised for a link, the Flex-Algorithm calculation **MUST** use the Bandwidth Metric advertised on the link and **MUST NOT** use the automatically derived metric for that link.

Metric Assignment in Interface Group Mode:

When a link does not have a configured Bandwidth Metric, the automatically derived metric **MUST** be used for that link.

In the context of Interface Group Mode, the following rules apply to parallel links:

- If all parallel links have advertised the Bandwidth Metric:

  The individual link Bandwidth Metrics **MUST** be used for each link during path computation.
- If only some of the parallel links have advertised the Bandwidth Metric:
  - The Bandwidth Metric advertisements for those links **MUST** be ignored.
  - Automatic metric calculation **MUST** be used to derive the link metrics for all parallel links.

This approach ensures consistent metric calculation and avoids discrepancies caused by partial Bandwidth Metric advertisements among parallel links.

5. Bandwidth Metric Considerations

This section specifies the rules of deriving the Bandwidth Metric if and only if the winning FAD for the Flex-Algorithm specifies the metric-type as "Bandwidth Metric".

1. If the Generic Metric sub-TLV with Bandwidth metric type is advertised for the link as described in Section 4, it **MUST** be used during the Flex-Algorithm calculation.
2. If the Generic Metric sub-TLV with Bandwidth metric type is not advertised for the link and the winning FAD for the Flex-Algorithm does not specify the automatic bandwidth metric calculation (as defined in Section 4.1), the link is treated as if the Bandwidth Metric is not available for the link.
3. If the Generic Metric sub-TLV with Bandwidth metric type is not advertised for the link and the winning FAD (Section 5.3 of [RFC9350]) for the Flex-Algorithm specifies the automatic bandwidth metric calculation (as defined in Section 4.1), the Bandwidth Metric **MUST** be automatically calculated per the procedures defined in Section 4.1. If the Link Bandwidth is not advertised for a link, the link **MUST** be pruned for the Flex-Algorithm calculations.
4. In ISIS, for Flex-Algorithm purposes, the Link Bandwidth is advertised as a sub-sub-TLV 9 of the Flex-Algorithm-specific ASLA sub-TLV. It is also possible to advertise the link bandwidth or Flex-Algorithm in sub-TLV 9 of TLV 22/222/23/223/141 [RFC5305] together with the L-flag set in the Flex-Algorithm-specific ASLA advertisement. In the absence of both of these advertisements, the bandwidth of the link is not available for Flex-Algorithm purposes.

6. Calculation of Flex-Algorithm Paths

The following two new additional rules are added to the existing rules in the Flex-Algorithm calculations specified in Section 13 of [RFC9350]:

6. Check if any exclude FAEMB rule is part of the Flex-Algorithm definition. If such exclude rule exists and the link has Maximum Link Bandwidth advertised, check if the link bandwidth satisfies the FAEMB rule. If the link does not satisfy the FAEMB rule, the link **MUST** be pruned from the Flex-Algorithm computation.
7. Check if any exclude FAEMD rule is part of the Flex-Algorithm definition. If such exclude rule exists and the link has Min Unidirectional link delay advertised, check if the link delay satisfies the FAEMD rule. If the link does not satisfy the FAEMD rule, the link **MUST** be pruned from the Flex-Algorithm computation.

7. Backward Compatibility

This extension brings no new backward-compatibility issues. This document defines new FAD constraints in Sections 3, 4.1.3, and 4.1.4. As described in [RFC9350], any node that does not understand sub-TLVs in a FAD TLV stops participation in the corresponding Flex-Algorithm. The new extensions can be deployed among the nodes that are upgraded to understand the new extensions without affecting the nodes that are not upgraded. This document also defines a new metric advertisement as described in Section 2. As per Section 13 of [RFC9350], when the links do not advertise the metric-type specified by the selected FAD, the link is pruned from Flex-Algorithm calculations. The new metric-types and Flex-Algorithms using the new metric-types can be deployed in the network without affecting existing deployment.

8. Security Considerations

This document inherits security considerations from [RFC9350].

9. Operational Considerations

Operational considerations defined in [RFC9350] generally apply to the extensions defined in this document as well. This document defines a metric-type range for user-defined metrics. When user-defined metrics are used in an inter-area or inter-level network, all the domains should assign same meaning to the particular metric-type. The YANG data models for Flex-Algorithm extensions are defined in documents [OSPF-AUGMENTATION] and [ISIS-AUGMENTATION].

Before the router goes into maintenance activity, the traffic needs to be diverted away from the router. This is achieved by setting the overload bit or setting link metrics on the router to a high value. In case of Generic Metric, the link metrics can be set to a Maximum usable metric for OSPF and IS-IS. The traffic will be diverted away from the router to a shorter available path. If there are no alternate paths available, traffic will stay on the router as the links are not removed from the Flex-Algorithm calculation when they are set to a maximum metric per [RFC9350].

10. IANA Considerations

10.1. IGP Metric-Type Registry

The "IGP Metric-Type" registry has been updated to include another column specifying whether the particular metric-type is allowed in the Generic Metric sub-TLV. The range 128-255 is redefined by this document as a user-defined range, and this range does not require Standards Action [RFC8126].

Type	Description	Reference	Allowed in Generic-Metric
0	IGP Metric	Section 5.1 of [RFC9350]	No
1	Min Unidirectional Link Delay as defined in [RFC8570], Section 4.2 and [RFC7471], Section 4.2	Section 5.1 of [RFC9350]	No

Type	Description	Reference	Allowed in Generic-Metric
2	Traffic Engineering Default Metric as defined in [RFC5305], Section 3.7 and Traffic Engineering Metric as defined in [RFC3630], Section 2.5.5	Section 5.1 of [RFC9350]	No
3	Bandwidth Metric	RFC 9843	Yes
128-255	Reserved for User-Defined Metric	RFC 9843	Yes

Table 1: IGP Metric-Type Registry

10.2. IS-IS Sub-Sub-TLVs for Flexible Algorithm Definition Sub-TLV

The "IS-IS Sub-Sub-TLVs for Flexible Algorithm Definition Sub-TLV" registry is part of the "IS-IS TLV Codepoints" registry group.

Type: 6
Description: IS-IS Exclude Minimum Bandwidth
Reference: RFC 9843, Section 3.1.1

Type: 7
Description: IS-IS Exclude Maximum Delay
Reference: RFC 9843, Section 3.1.2

Type: 8
Description: IS-IS Reference Bandwidth
Reference: RFC 9843, Section 4.1.3.1

Type: 9
Description: IS-IS Bandwidth Threshold
Reference: RFC 9843, Section 4.1.3.2

10.3. OSPF Sub-TLVs for Flexible Algorithm Definition Sub-TLV

The "OSPF Flexible Algorithm Definition TLV Sub-TLVs" registry is part of the "Open Shortest Path First (OSPF) Parameters" registry group.

Bit Number: 6
Description: OSPF Exclude Minimum Bandwidth
Reference: RFC 9843, Section 3.2.1

Bit Number: 7
Description: OSPF Exclude Maximum Delay
Reference: RFC 9843, Section 3.2.2

Bit Number: 8
Description: OSPF Reference Bandwidth
Reference: RFC 9843, Section 4.1.4.1

Bit Number: 9
Description: OSPF Bandwidth Threshold
Reference: RFC 9843, Section 4.1.4.2

10.4. IS-IS Sub-TLVs for TLVs Advertising Neighbor Information

The "IS-IS Sub-TLVs for TLVs Advertising Neighbor Information" registry is part of the "IS-IS TLV Codepoints" registry group.

Type: 17
Description: Generic Metric
TLVs set to "Y": 22, 23, 25, 141, 222, and 223
Reference: RFC 9843, Section 2.1

10.5. Sub-Sub-TLV Codepoints for Application-Specific Link Attributes

The "IS-IS Sub-Sub-TLV Codepoints for Application-Specific Link Attributes" registry is part of the "IS-IS TLV Codepoints" registry group.

Type: 17
Description: Generic Metric
Reference: RFC 9843, Section 2.1

10.6. OSPFv2 Extended Link TLV Sub-TLVs

The "OSPFv2 Extended Link TLV Sub-TLVs" registry is part of the "Open Shortest Path First v2 (OSPFv2) Parameters" registry group.

Value: 25
Description: Generic Metric
L2BM: Y
Reference: RFC 9843, Section 2.2

10.7. Types for Sub-TLVs of TE Link TLV (Value 2)

The "Types for sub-TLVs of TE Link TLV (Value 2)" registry is part of the "Open Shortest Path First (OSPF) Traffic Engineering TLVs" registry group.

Value: 36
Description: Generic Metric
Reference: RFC 9843, Section 2.2

10.8. OSPFv3 Extended-LSA Sub-TLVs

The "OSPFv3 Extended-LSA Sub-TLVs" registry is part of the "Open Shortest Path First v3 (OSPFv3) Parameters" registry group.

Value: 34
Description: Generic Metric
L2BM: Y
Reference: RFC 9843, Section 2.2

11. References

11.1. Normative References

[IEEE754-2019] IEEE, "IEEE Standard for Floating-Point Arithmetic", IEEE Std 754-2019, DOI 10.1109/ieeestd.2019.8766229, 22 July 2019, <https://doi.org/10.1109/ieeestd.2019.8766229>.

[RFC1195] Callon, R., "Use of OSI IS-IS for routing in TCP/IP and dual environments", RFC 1195, DOI 10.17487/RFC1195, December 1990, <https://www.rfc-editor.org/info/rfc1195>.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/info/rfc2119>.

[RFC2328] Moy, J., "OSPF Version 2", STD 54, RFC 2328, DOI 10.17487/RFC2328, April 1998, <https://www.rfc-editor.org/info/rfc2328>.

[RFC3630] Katz, D., Kompella, K., and D. Yeung, "Traffic Engineering (TE) Extensions to OSPF Version 2", RFC 3630, DOI 10.17487/RFC3630, September 2003, <https://www.rfc-editor.org/info/rfc3630>.

[RFC5305] Li, T. and H. Smit, "IS-IS Extensions for Traffic Engineering", RFC 5305, DOI 10.17487/RFC5305, October 2008, <https://www.rfc-editor.org/info/rfc5305>.

[RFC5329] Ishiguro, K., Manral, V., Davey, A., and A. Lindem, Ed., "Traffic Engineering Extensions to OSPF Version 3", RFC 5329, DOI 10.17487/RFC5329, September 2008, <https://www.rfc-editor.org/info/rfc5329>.

[RFC5392] Chen, M., Zhang, R., and X. Duan, "OSPF Extensions in Support of Inter-Autonomous System (AS) MPLS and GMPLS Traffic Engineering", RFC 5392, DOI 10.17487/RFC5392, January 2009, <https://www.rfc-editor.org/info/rfc5392>.

[RFC7684] Psenak, P., Gredler, H., Shakir, R., Henderickx, W., Tantsura, J., and A. Lindem, "OSPFv2 Prefix/Link Attribute Advertisement", RFC 7684, DOI 10.17487/RFC7684, November 2015, <https://www.rfc-editor.org/info/rfc7684>.

[RFC8126] Cotton, M., Leiba, B., and T. Narten, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 8126, DOI 10.17487/RFC8126, June 2017, <https://www.rfc-editor.org/info/rfc8126>.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <https://www.rfc-editor.org/info/rfc8174>.

[RFC8362] Lindem, A., Roy, A., Goethals, D., Reddy Vallem, V., and F. Baker, "OSPFv3 Link State Advertisement (LSA) Extensibility", RFC 8362, DOI 10.17487/RFC8362, April 2018, <https://www.rfc-editor.org/info/rfc8362>.

[RFC8668] Ginsberg, L., Ed., Bashandy, A., Filsfils, C., Nanduri, M., and E. Aries, "Advertising Layer 2 Bundle Member Link Attributes in IS-IS", RFC 8668, DOI 10.17487/RFC8668, December 2019, <https://www.rfc-editor.org/info/rfc8668>.

[RFC9350] Psenak, P., Ed., Hegde, S., Filsfils, C., Talaulikar, K., and A. Gulko, "IGP Flexible Algorithm", RFC 9350, DOI 10.17487/RFC9350, February 2023, <https://www.rfc-editor.org/info/rfc9350>.

[RFC9356] Talaulikar, K., Ed. and P. Psenak, "Advertising Layer 2 Bundle Member Link Attributes in OSPF", RFC 9356, DOI 10.17487/RFC9356, January 2023, <https://www.rfc-editor.org/info/rfc9356>.

[RFC9479] Ginsberg, L., Psenak, P., Previdi, S., Henderickx, W., and J. Drake, "IS-IS Application-Specific Link Attributes", RFC 9479, DOI 10.17487/RFC9479, October 2023, <https://www.rfc-editor.org/info/rfc9479>.

[RFC9492] Psenak, P., Ed., Ginsberg, L., Henderickx, W., Tantsura, J., and J. Drake, "OSPF Application-Specific Link Attributes", RFC 9492, DOI 10.17487/RFC9492, October 2023, <https://www.rfc-editor.org/info/rfc9492>.

11.2. Informative References

[ISIS-AUGMENTATION] Lindem, A., Qu, Y., and S. Litkowski, "IS-IS YANG Model Augmentations for Additional Features - Release 1", Work in Progress, Internet-Draft, draft-ietf-lsr-isis-yang-augmentation-v1-10, 6 July 2025, <https://datatracker.ietf.org/doc/html/draft-ietf-lsr-isis-yang-augmentation-v1-10>.

[OSPF-AUGMENTATION] Lindem, A. and Y. Qu, "OSPF YANG Model Augmentations for Additional Features - Release 1", Work in Progress, Internet-Draft, draft-ietf-lsr-ospf-yang-augmentation-v1-17, 6 July 2025, <https://datatracker.ietf.org/doc/html/draft-ietf-lsr-ospf-yang-augmentation-v1-17>.

[RFC4655] Farrel, A., Vasseur, J.-P., and J. Ash, "A Path Computation Element (PCE)-Based Architecture", RFC 4655, DOI 10.17487/RFC4655, August 2006, <https://www.rfc-editor.org/info/rfc4655>.

[RFC5120] Przygienda, T., Shen, N., and N. Sheth, "M-ISIS: Multi Topology (MT) Routing in Intermediate System to Intermediate Systems (IS-ISs)", RFC 5120, DOI 10.17487/RFC5120, February 2008, <https://www.rfc-editor.org/info/rfc5120>.

[RFC5311] McPherson, D., Ed., Ginsberg, L., Previdi, S., and M. Shand, "Simplified Extension of Link State PDU (LSP) Space for IS-IS", RFC 5311, DOI 10.17487/RFC5311, February 2009, <https://www.rfc-editor.org/info/rfc5311>.

[RFC5715] Shand, M. and S. Bryant, "A Framework for Loop-Free Convergence", RFC 5715, DOI 10.17487/RFC5715, January 2010, <https://www.rfc-editor.org/info/rfc5715>.

[RFC7471] Giacalone, S., Ward, D., Drake, J., Atlas, A., and S. Previdi, "OSPF Traffic Engineering (TE) Metric Extensions", RFC 7471, DOI 10.17487/RFC7471, March 2015, <https://www.rfc-editor.org/info/rfc7471>.

[RFC8570] Ginsberg, L., Ed., Previdi, S., Ed., Giacalone, S., Ward, D., Drake, J., and Q. Wu, "IS-IS Traffic Engineering (TE) Metric Extensions", RFC 8570, DOI 10.17487/RFC8570, March 2019, <https://www.rfc-editor.org/info/rfc8570>.

[RFC9346] Chen, M., Ginsberg, L., Previdi, S., and D. Xiaodong, "IS-IS Extensions in Support of Inter-Autonomous System (AS) MPLS and GMPLS Traffic Engineering", RFC 9346, DOI 10.17487/RFC9346, February 2023, <https://www.rfc-editor.org/info/rfc9346>.

[SR-LOOP-AVOID] Bashandy, A., Filsfils, C., Litkowski, S., Decraene, B., Francois, P., and P. Psenak, "Loop avoidance using Segment Routing", Work in Progress, Internet-Draft, draft-bashandy-rtgwg-segment-routing-uloop-17, 29 June 2024, <https://datatracker.ietf.org/doc/html/draft-bashandy-rtgwg-segment-routing-uloop-17>.

Appendix A. Updated List of Rules for Pruning Links in Flex-Algorithm Topology

This section lists the entire set of rules to prune links from Flex-Algorithm topology during Flexible Algorithm calculation. It includes the original rules defined in Section 13 of [RFC9350] as well as the new additions (rules 6 and 7) described in this document.

For all links in the topology:

1. Check if any exclude Administrative Group rule is part of the Flex-Algorithm Definition. If such exclude rule exists, check if any color that is part of the exclude rule is also set on the link. If such a color is set, the link **MUST** be pruned from the computation.
2. Check if any exclude SRLG rule is part of the Flex-Algorithm Definition. If such exclude rule exists, check if the link is part of any SRLG that is also part of the SRLG exclude rule. If the link is part of such SRLG, the link **MUST** be pruned from the computation.
3. Check if any include-any Administrative Group rule is part of the Flex-Algorithm Definition. If such include-any rule exists, check if any color that is part of the include-any rule is also set on the link. If no such color is set, the link **MUST** be pruned from the computation.
4. Check if any include-all Administrative Group rule is part of the Flex-Algorithm Definition. If such include-all rule exists, check if all colors that are part of the include-all rule are also set on the link. If all such colors are not set on the link, the link **MUST** be pruned from the computation.
5. If the Flex-Algorithm Definition uses something other than the IGP metric (Section 5 of [RFC9350]), and such metric is not advertised for the particular link in a topology for which the computation is done, such link **MUST** be pruned from the computation. A metric of value 0 **MUST NOT** be assumed in such a case.
6. Check if any exclude FAEMB rule is part of the Flex-Algorithm Definition. If such exclude rule exists and the link has Maximum Link Bandwidth advertised, check if the link bandwidth satisfies the FAEMB rule. If the link does not satisfy the FAEMB rule, the link **MUST** be pruned from the Flex-Algorithm computation.
7. Check if any exclude FAEMD rule is part of the Flex-Algorithm Definition. If such exclude rule exists and the link has Min Unidirectional Link Delay advertised, check if the link delay satisfies the FAEMD rule. If the link does not satisfy the FAEMD rule, the link **MUST** be pruned from the Flex-Algorithm computation.

Acknowledgements

Many thanks to Chris Bowers, Krzysztof Szarcowitz, Julian Lucek, Ram Santhanakrishnan, Ketan Talaulikar, and Acee Lindem for discussions and input.

Contributors

Salih K.A.
Juniper Networks
Email: salih@juniper.net

Authors' Addresses

Shraddha Hegde
Juniper Networks Inc.
Exora Business Park
Bangalore 560103
Karnataka
India
Email: shraddha@juniper.net

William Britto
Juniper Networks Inc.
Email: bwilliam@juniper.net

Rajesh Shetty
Juniper Networks Inc.
Email: mrajesh@juniper.net

Bruno Decraene
Orange
Email: bruno.decraene@orange.com

Peter Psenak
Cisco Systems
Email: ppsenak@cisco.com

Tony Li
Juniper Networks Inc.
Email: tony.li@tony.li