Authors: W. Britto   S. Hegde   P. Kaneriya   R. Shetty
             Juniper Networks *Juniper Networks* *Juniper Networks* *Juniper Networks*

        R. Bonica   P. Psenak
      Juniper Networks   *Cisco Systems*

RFC 9502
IGP Flexible Algorithm in IP Networks

Abstract

This document extends IGP Flexible Algorithm so that it can be used with regular IPv4 and IPv6 forwarding.

Status of This Memo

This is an Internet Standards Track document.

This document is a product of the Internet Engineering Task Force (IETF). It represents the consensus of the IETF community. It has received public review and has been approved for publication by the Internet Engineering Steering Group (IESG). Further information on Internet Standards is available in Section 2 of RFC 7841.

Information about the current status of this document, any errata, and how to provide feedback on it may be obtained at https://www.rfc-editor.org/info/rfc9502.

Copyright Notice

Table of Contents

1. Introduction

An IGP Flexible Algorithm allows IGPs to compute constraint-based paths. The base IGP Flexible Algorithm specification describes how it is used with Segment Routing (SR) data planes: SR MPLS and SRv6.

An IGP Flexible Algorithm as specified in [RFC9350] computes a constraint-based path to:

- All Flexible-Algorithm-specific Prefix Segment Identifiers (SIDs) [RFC8402].
- All Flexible-Algorithm-specific SRv6 Locators [RFC8986].

Therefore, Flexible Algorithm cannot be deployed in the absence of SR or SRv6.

This document extends Flexible Algorithm, allowing it to compute paths to IPv4 and IPv6 prefixes.

2. 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.

3. Use Case Example

In this section, we illustrate one use case that motivates this specification: if a specific service can be identified by an IP address, traffic to it can use constraint-based paths computed according to this specification.

The System architecture for the 5G System [TS.23.501-3GPP] describes the N3 interface between gNodeB and UPF (User Plane Function).

Mobile networks are becoming more and more IP-centric. Each end-user session from a gNodeB can be destined to a specific UPF based on the session requirements. For example, some sessions require high bandwidth, while others need to be routed along the lowest latency path. Each UPF is assigned a unique IP address. As a result, traffic for different sessions is destined to a different destination IP address.

The IP address allocated to the UPF can be associated with an algorithm. The mobile user traffic is then forwarded along the path based on the algorithm-specific metric and constraints. As a result, traffic can be sent over a path that is optimized for minimal latency or highest bandwidth. This mechanism is used to achieve Service Level Agreement (SLA) appropriate for a user session.

4. Advertising Flexible Algorithm Definitions (FADs)

To guarantee loop-free forwarding, all routers that participate in a Flex-Algorithm **MUST** agree on the Flexible Algorithm Definition (FAD).

Selected nodes within the IGP domain **MUST** advertise FADs as described in Sections 5, 6, and 7 of [RFC9350].

5. Advertising IP Flexible Algorithm Participation

A node may use various algorithms when calculating paths to nodes and prefixes. Algorithm values are defined in the "IGP Algorithm Types" registry [IANA-ALG].

Only a node that is participating in a Flex-Algorithm is:

- Able to compute a path for such Flex-Algorithm
- Part of the topology for such Flex-Algorithm

Flexible Algorithm participation **MUST** be advertised for each Flexible Algorithm data plane independently, as specified in [RFC9350]. Using Flexible Algorithm for regular IPv4 and IPv6 prefixes represents an independent Flexible Algorithm data plane; as such, the Flexible Algorithm participation for the IP Flexible Algorithm data plane **MUST** be signaled independently of any other Flexible Algorithm data plane (e.g., SR).

All routers in an IGP domain participate in default algorithm 0. Advertisement of participation in IP Flexible Algorithm does not impact the router participation in default algorithm 0.

Advertisement of participation in IP Flexible Algorithm does not impact the router participation signaled for other data planes. For example, it is possible that a router participates in a particular Flex-Algorithm for the IP data plane but does not participate in the same Flex-Algorithm for the SR data plane.

The following sections describe how the IP Flexible Algorithm participation is advertised in IGP protocols.

5.1. The IS-IS IP Algorithm Sub-TLV

The IS-IS [ISO10589] IP Algorithm Sub-TLV is a sub-TLV of the IS-IS Router Capability TLV [RFC7981] and has the following format:
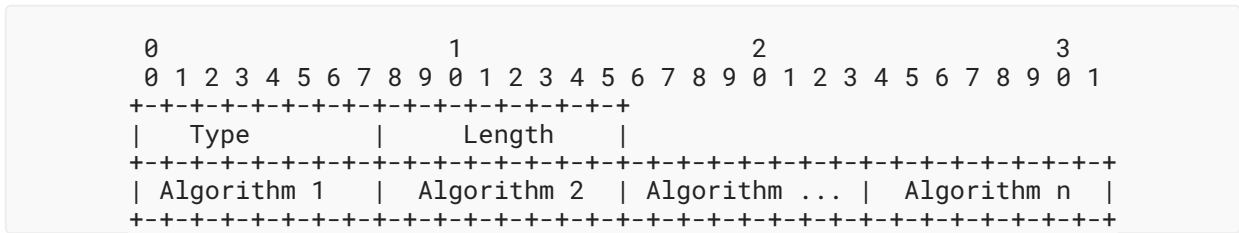
```
     0                   1                   2                   3
     0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |     Type      |    Length     |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    | Algorithm 1   | Algorithm 2   | Algorithm ... | Algorithm n   |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 1: IS-IS IP Algorithm Sub-TLV

Type (1 octet):   IP Algorithm Sub-TLV (Value 29)

Length (1 octet):   Variable

Algorithm (1 octet):   Value from 128 to 255

The IP Algorithm Sub-TLV **MUST** be propagated throughout the level and **MUST NOT** be advertised across level boundaries. Therefore, the S bit in the Router Capability TLV, in which the IP Algorithm Sub-TLV is advertised, **MUST NOT** be set.

The IP Algorithm Sub-TLV is optional. It **MUST NOT** be advertised more than once at a given level. A router receiving multiple IP Algorithm sub-TLVs from the same originator **MUST** select the first advertisement in the lowest-numbered Link State PDU (LSP), and subsequent instances of the IP Algorithm Sub-TLV **MUST** be ignored.

Algorithms outside the Flex-Algorithm range (128-255) **MUST** be ignored by the receiver. This situation **SHOULD** be logged as an error.

The IP Flex-Algorithm participation advertised in the IS-IS IP Algorithm Sub-TLV is topology independent. When a router advertises participation in the IS-IS IP Algorithm Sub-TLV, the participation applies to all topologies in which the advertising node participates.

5.2. The OSPF IP Algorithm TLV

The OSPF [RFC2328] IP Algorithm TLV is a top-level TLV of the Router Information Opaque Link State Advertisement (LSA) [RFC7770] and 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            |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |   Algorithm 1 | Algorithm...  |   Algorithm n |               |
    +-                                                             -+
    |                                                               |
    +                                                               +
```
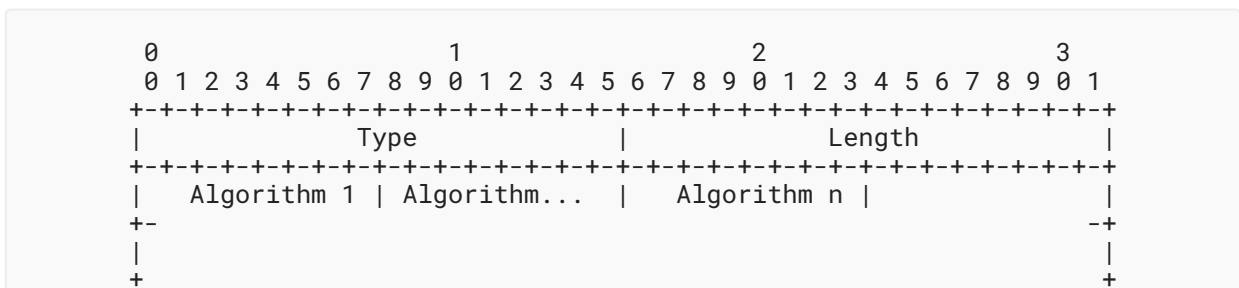
Figure 2: OSPF IP Algorithm TLV

Type (2 octets):   IP Algorithm TLV (21)

Length(2 octets):   Variable

Algorithm (1 octet):   Value from 128 to 255

The IP Algorithm TLV is optional. It **MUST** only be advertised once in the Router Information LSA.

Algorithms outside the Flex-Algorithm range (128-255) **MUST** be ignored by the receiver. This situation **SHOULD** be logged as an error.

When multiple IP Algorithm TLVs are received from a given router, the receiver **MUST** use the first occurrence of the TLV in the Router Information LSA. If the IP Algorithm TLV appears in multiple Router Information LSAs that have different flooding scopes, the IP Algorithm TLV in the Router Information LSA with the area-scoped flooding scope **MUST** be used. If the IP Algorithm TLV appears in multiple Router Information LSAs that have the same flooding scope, the IP Algorithm TLV in the Router Information LSA with the numerically smallest Instance ID (Opaque ID for OSPFv2 or Link State ID for OSPFv3) **MUST** be used, and subsequent instances of the IP Algorithm TLV **MUST** be ignored.

The Router Information LSA can be advertised at any of the defined flooding scopes (link, area, or Autonomous System (AS)). For the purpose of IP Algorithm TLV advertisement, area- or AS-scoped flooding is **REQUIRED**. The AS flooding scope **SHOULD NOT** be used unless local configuration policy on the originating router indicates domain-wide flooding.

The IP Flexible Algorithm participation advertised in the OSPF IP Algorithm TLV is topology independent. When a router advertises participation in OSPF IP Algorithm TLV, the participation applies to all topologies in which the advertising node participates.

6. Advertising IP Flexible Algorithm Reachability

To be able to associate the prefix with the Flex-Algorithm, the existing prefix reachability advertisements cannot be used, because they advertise the prefix reachability in default algorithm 0. Instead, new IP Flexible Algorithm reachability advertisements are defined in IS-IS and OSPF.

The M-flag in the FAD is not applicable to IP Algorithm Prefixes. Any IP Algorithm Prefix advertisement includes the Algorithm and Metric fields. When an IP Algorithm Prefix is advertised between areas or domains, the metric field in the IP Algorithm Prefix advertisement **MUST** be used irrespective of the M-flag in the FAD advertisement.

6.1. The IS-IS IPv4 Algorithm Prefix Reachability TLV

The IPv4 Algorithm Prefix Reachability top-level TLV is defined for advertising IPv4 Flexible Algorithm Prefix Reachability in IS-IS.

This new TLV shares the sub-TLV space defined for TLVs Advertising Prefix Reachability.

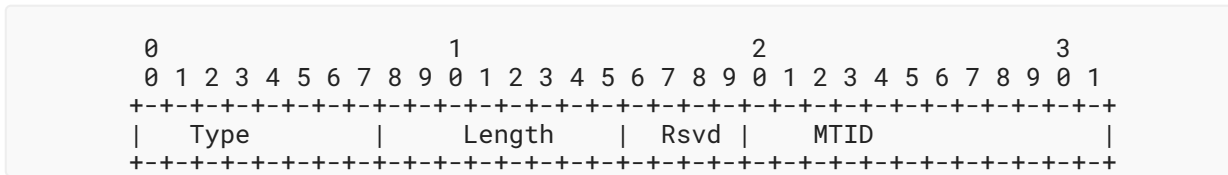The IS-IS IPv4 Algorithm Prefix Reachability TLV has the following format:

```
     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     | Rsvd |    MTID               |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 3: IS-IS IPv4 Algorithm Prefix Reachability TLV

Type (1 octet):   IPv4 Algorithm Prefix Reachability TLV (Value 126)

Length (1 octet):   Variable based on number of prefix entries encoded

Rsvd (4 bits):   Reserved for future use. They **MUST** be set to zero on transmission and **MUST** be ignored on receipt.

MTID (12 bits):   Multitopology Identifier as defined in [RFC5120]. Note that the value 0 is legal.
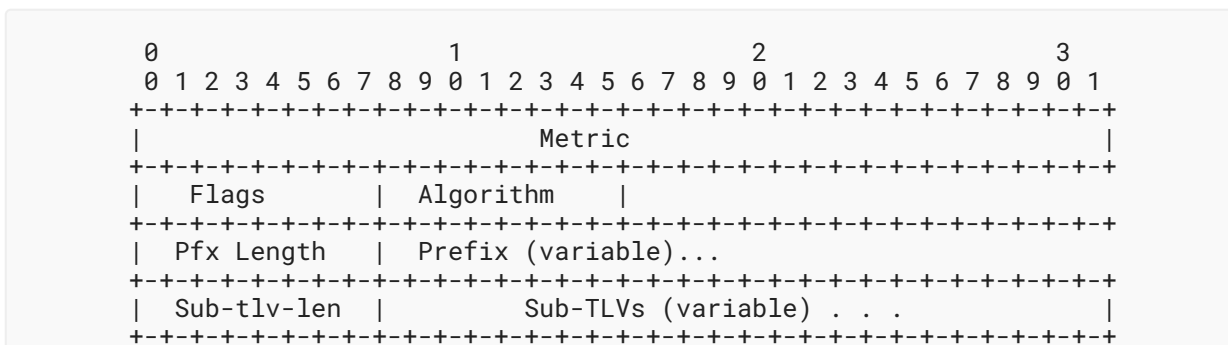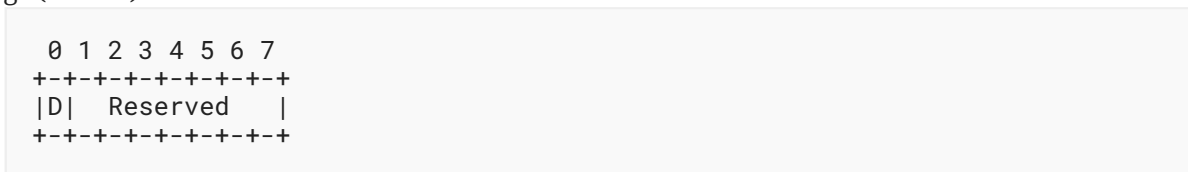
Followed by one or more prefix entries of the form:

```
     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
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                          Metric                              |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |    Flags      |   Algorithm   |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |  Pfx Length   |   Prefix (variable)...
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |  Sub-tlv-len  |       Sub-TLVs (variable) . . .              |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 4: IS-IS IPv4 Algorithm Prefix Reachability TLV

Metric (4 octets):   Metric information as defined in [RFC5305]

Flags (1 octet):

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

D-flag: The D-flag is described as the "up/down bit" in Section 4.1 of [RFC5305]. When the Prefix is leaked from level 2 to level 1, the D bit **MUST** be set. Otherwise, this bit **MUST** be clear. Prefixes with the D bit set **MUST NOT** be leaked from level 1 to level 2. This is to prevent looping.

The remaining bits: Are reserved for future use. They **MUST** be set to zero on transmission and **MUST** be ignored on receipt.

Algorithm (1 octet):   Associated Algorithm from 128 to 255

Prefix Len (1 octet):   Prefix length measured in bits

Prefix (variable length):   Prefix mapped to Flex-Algorithm

Optional Sub-TLV-length (1 octet):   Number of octets used by sub-TLVs

Optional sub-TLVs (variable length)

If the Algorithms in the IS-IS IPv4 Algorithm Prefix Reachability TLV are outside the Flex-Algorithm range (128-255), the IS-IS IPv4 Algorithm Prefix Reachability TLV **MUST** be ignored by the receiver. This situation **SHOULD** be logged as an error.

If a router receives multiple IPv4 Algorithm Prefix Reachability advertisements for the same prefix from the same originator, it **MUST** select the first advertisement in the lowest-numbered LSP and ignore any subsequent IPv4 Algorithm Prefix Reachability advertisements for the same prefix.

If a router receives multiple IPv4 Algorithm Prefix Reachability advertisements for the same prefix, from different originators, where all of them do not advertise the same algorithm, it **MUST** ignore all of them and **MUST NOT** install any forwarding entries based on these advertisements. This situation **SHOULD** be logged as an error.

In cases where a prefix advertisement is received in both an IPv4 Prefix Reachability TLV [RFC5305] [RFC5120] and an IPv4 Algorithm Prefix Reachability TLV, the IPv4 Prefix Reachability advertisement **MUST** be preferred when installing entries in the forwarding plane.

6.2.   The IS-IS IPv6 Algorithm Prefix Reachability TLV

The IS-IS IPv6 Algorithm Prefix Reachability TLV is identical to the IS-IS IPv4 Algorithm Prefix Reachability TLV, except that it has a distinct type. The type is 127.

If the Algorithms in the IS-IS IPv6 Algorithm Prefix Reachability TLV are outside the Flex-Algorithm range (128-255), the IS-IS IPv6 Algorithm Prefix Reachability TLV **MUST** be ignored by the receiver. This situation **SHOULD** be logged as an error.

If a router receives multiple IPv6 Algorithm Prefix Reachability advertisements for the same prefix from the same originator, it **MUST** select the first advertisement in the lowest-numbered LSP and ignore any subsequent IPv6 Algorithm Prefix Reachability advertisements for the same prefix.

If a router receives multiple IPv6 Algorithm Prefix Reachability advertisements for the same prefix, from different originators, where all of them do not advertise the same algorithm, it **MUST** ignore all of them and **MUST NOT** install any forwarding entries based on these advertisements. This situation **SHOULD** be logged as an error.

In cases where a prefix advertisement is received in both an IPv6 Prefix Reachability TLV [RFC5308] [RFC5120] and an IPv6 Algorithm Prefix Reachability TLV, the IPv6 Prefix Reachability advertisement **MUST** be preferred when installing entries in the forwarding plane.

In cases where a prefix advertisement is received in both an IS-IS SRv6 Locator TLV [RFC9352] and in IS-IS IPv6 Algorithm Prefix Reachability TLV, the receiver **MUST** ignore both of them and **MUST NOT** install any forwarding entries based on these advertisements. This situation **SHOULD** be logged as an error.

6.3. The OSPFv2 IP Algorithm Prefix Reachability Sub-TLV

A new sub-TLV of the OSPFv2 Extended Prefix TLV is defined for advertising IP Algorithm Prefix Reachability in OSPFv2, the OSPFv2 IP Algorithm Prefix Reachability Sub-TLV.

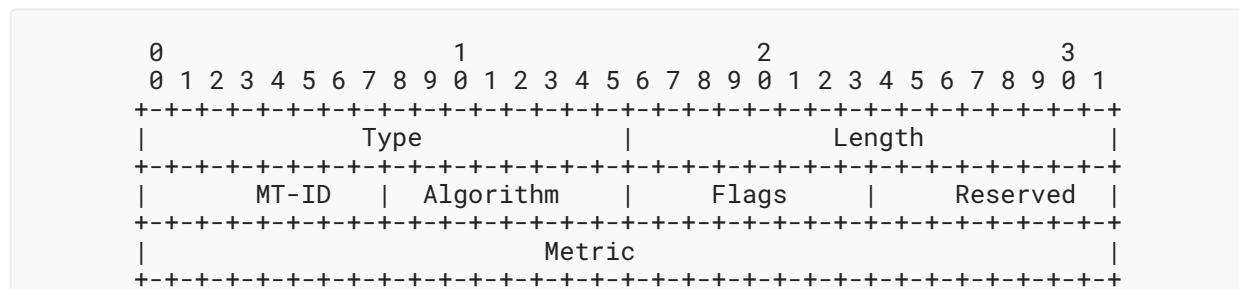The OSPFv2 IP Algorithm Prefix Reachability Sub-TLV has the following format:

```
    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            |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |     MT-ID     |   Algorithm   |      Flags     |    Reserved   |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                            Metric                             |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

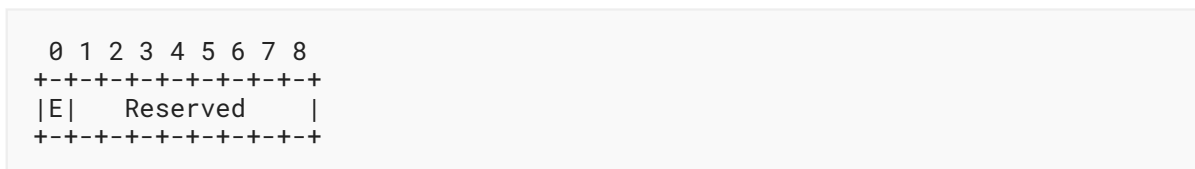*Figure 5: OSPFv2 IP Algorithm Prefix Reachability Sub-TLV*

Type (2 octets):   The value is 6

Length (2 octets):   8

MT-ID (1 octet):   Multi-Topology ID as defined in [RFC4915]

Algorithm (1 octet):   Associated Algorithm from 128 to 255

Flags (1 octet):   The following flags are defined:

```
   0 1 2 3 4 5 6 7 8
  +-+-+-+-+-+-+-+-+-+
  |E|   Reserved    |
  +-+-+-+-+-+-+-+-+-+
```

Where:

E bit: The same as the E bit defined in Appendix A.4.5 of [RFC2328].

The remaining bits: Are reserved for future use. They **MUST** be set to zero on transmission and **MUST** be ignored on receipt.

Reserved (1 octet):   **SHOULD** be set to 0 on transmission and **MUST** be ignored on reception.

Metric (4 octets):   The algorithm-specific metric value. The metric value of 0XFFFFFFFF **MUST** be considered unreachable.

If the Algorithms in the OSPFv2 IP Algorithm Prefix Reachability Sub-TLV are outside the Flex-Algorithm range (128-255), the OSPFv2 IP Algorithm Prefix Reachability Sub-TLV **MUST** be ignored by the receiver. This situation **SHOULD** be logged as an error.

An OSPFv2 router receiving multiple OSPFv2 IP Algorithm Prefix Reachability Sub-TLVs in the same OSPFv2 Extended Prefix TLV **MUST** select the first advertisement of this sub-TLV and **MUST** ignore all remaining occurrences of this sub-TLV in the OSPFv2 Extended Prefix TLV.

An OSPFv2 router receiving multiple OSPFv2 IP Algorithm Prefix Reachability TLVs for the same prefix from different originators where all of them do not advertise the same algorithm **MUST** ignore all of them and **MUST NOT** install any forwarding entries based on these advertisements. This situation **SHOULD** be logged as an error.

In cases where a prefix advertisement is received in any of the LSAs advertising the prefix reachability for algorithm 0 and in an OSPFv2 IP Algorithm Prefix Reachability Sub-TLV, only the prefix reachability advertisement for algorithm 0 **MUST** be used, and all occurrences of the OSPFv2 IP Algorithm Prefix Reachability Sub-TLV **MUST** be ignored.

When computing the IP Algorithm Prefix reachability in OSPFv2, only information present in the OSPFv2 Extended Prefix TLV **MUST** be used. There will not be any information advertised for the IP Algorithm Prefix in any of the OSPFv2 LSAs that advertise prefix reachability for algorithm 0. For the IP Algorithm Prefix, the OSPFv2 Extended Prefix TLV is used to advertise the prefix reachability, unlike for algorithm 0 prefixes, where the OSPFv2 Extended Prefix TLV is only used to advertise additional attributes -- but not the reachability itself.

6.3.1. The OSPFv2 IP Forwarding Address Sub-TLV

A new sub-TLV of the OSPFv2 Extended Prefix TLV is defined for advertising IP Forwarding Address, the OSPFv2 IP Forwarding Address Sub-TLV.

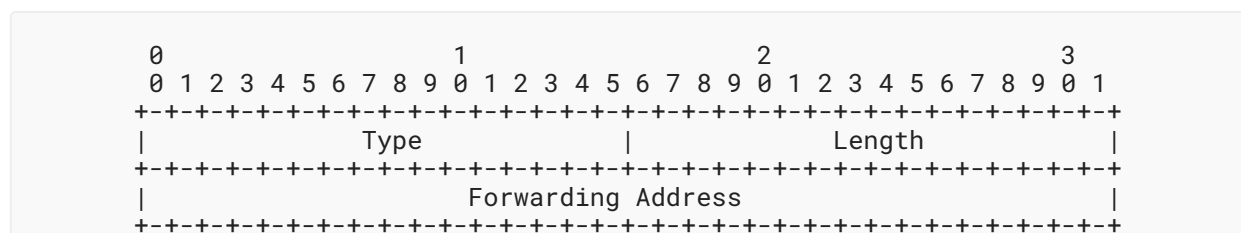The OSPFv2 IP Forwarding Address Sub-TLV has the following format:

```
    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            |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                      Forwarding Address                       |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 6: OSPFv2 IP Forwarding Address Sub-TLV

Type (2 octets):   The value is 7

Length (2 octets):   4

Forwarding Address (4 octets):   The same as defined in Appendix A.4.5 of [RFC2328]

The OSPFv2 IP Forwarding Address Sub-TLV **MUST NOT** be used for computing algorithm 0 prefix reachability and **MUST** be ignored for algorithm 0 prefixes.

The OSPFv2 IP Forwarding Address Sub-TLV is optional. If it is not present, the forwarding address for computing the IP Algorithm Prefix reachability is assumed to be equal to 0.0.0.0.

The OSPFv2 IP Forwarding Address Sub-TLV is only applicable to AS External and Not-So-Stubby Area (NSSA) External route types. If the OSPFv2 IP Forwarding Address Sub-TLV is advertised in the OSPFv2 Extended Prefix TLV that has the Route Type field set to any other type, the OSPFv2 IP Forwarding Address Sub-TLV **MUST** be ignored.

6.4. The OSPFv3 IP Algorithm Prefix Reachability Sub-TLV

The OSPFv3 [RFC5340] IP Algorithm Prefix Reachability Sub-TLV is defined for advertisement of the IP Algorithm Prefix Reachability in OSPFv3.

The OSPFv3 IP Algorithm Prefix Reachability Sub-TLV is a sub-TLV of the following OSPFv3 TLVs defined in [RFC8362]:

- Intra-Area-Prefix TLV
- Inter-Area-Prefix TLV
- External-Prefix TLV

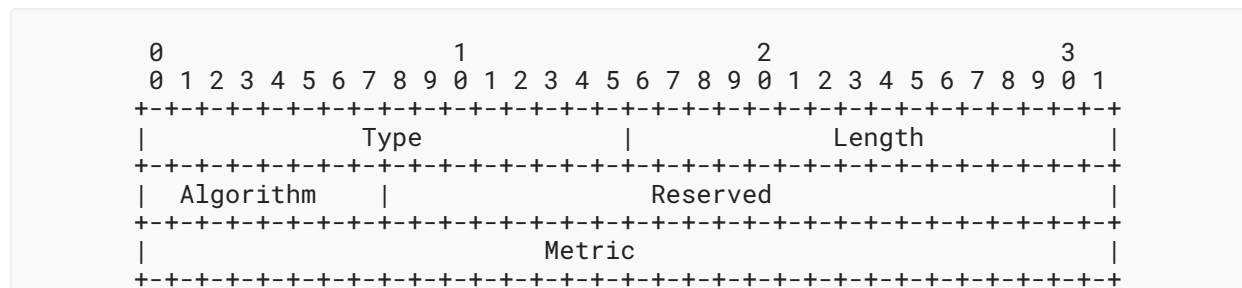The format of OSPFv3 IP Algorithm Prefix Reachability Sub-TLV is shown below:

```
    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           |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |   Algorithm   |                   Reserved                    |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                             Metric                            |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 7: OSPFv3 IP Algorithm Prefix Reachability Sub-TLV

Where:

Type (2 octets):   The value is 35

Length (2 octets):   8

Algorithm (1 octet):   Associated Algorithm from 128 to 255

Reserved (3 octets):   **SHOULD** be set to 0 on transmission and **MUST** be ignored on reception.

Metric (4 octets):   The algorithm-specific metric value. The metric value of 0XFFFFFFFF **MUST** be considered unreachable.

If the Algorithms in the OSPFv3 IP Algorithm Prefix Reachability Sub-TLV are outside the Flex-Algorithm range (128-255), the OSPFv3 IP Algorithm Prefix Reachability Sub-TLV **MUST** be ignored by the receiver. This situation **SHOULD** be logged as an error.

When the OSPFv3 IP Algorithm Prefix Reachability Sub-TLV is present, the NU-bit in the PrefixOptions field of the parent TLV **MUST** be set. This is needed to prevent the OSPFv3 IP Algorithm Prefix Reachability advertisement from contributing to the base algorithm reachability. If the NU-bit in the PrefixOptions field of the parent TLV is not set, the OSPFv3 IP Algorithm Prefix Sub-TLV **MUST** be ignored by the receiver.

The metric value in the parent TLV is **RECOMMENDED** to be set to LSInfinity [RFC2328]. This recommendation is provided as a network troubleshooting convenience; if it is not followed, the protocol will still function correctly.

An OSPFv3 router receiving multiple OSPFv3 IP Algorithm Prefix Reachability Sub-TLVs in the same parent TLV **MUST** select the first advertisement of this sub-TLV and **MUST** ignore all remaining occurrences of this sub-TLV in the parent TLV.

An OSPFv3 router receiving multiple OSPFv3 IP Algorithm Prefix Reachability TLVs for the same prefix from different originators where all of them do not advertise the same algorithm **MUST** ignore all of them and **MUST NOT** install any forwarding entries based on these advertisements. This situation **SHOULD** be logged as an error.

In cases where a prefix advertisement is received in any of the LSAs advertising the prefix reachability for algorithm 0 and in an OSPFv3 OSPFv3 IP Algorithm Prefix Reachability Sub-TLV, only the prefix reachability advertisement for algorithm 0 **MUST** be used, and all occurrences of the OSPFv3 IP Algorithm Prefix Reachability Sub-TLV **MUST** be ignored.

In cases where a prefix advertisement is received in both an OSPFv3 SRv6 Locator TLV and in an OSPFv3 IP Algorithm Prefix Reachability Sub-TLV, the receiver **MUST** ignore both of them and **MUST NOT** install any forwarding entries based on these advertisements. This situation **SHOULD** be logged as an error.

6.5. The OSPF IP Flexible Algorithm ASBR Metric Sub-TLV

[RFC9350] defines the OSPF Flexible Algorithm ASBR Metric (FAAM) Sub-TLV that is used by an OSPFv2 or an OSPFv3 Area Border Router (ABR) to advertise a Flex-Algorithm-specific metric associated with the corresponding ASBR LSA.

As described in [RFC9350], each data plane signals its participation independently. IP Flexible Algorithm participation is signaled independent of SR Flexible Algorithm participation. As a result, the calculated topologies for SR and IP Flexible Algorithm could be different. Such a difference prevents the usage of FAAM for the purpose of the IP Flexible Algorithm.

The OSPF IP Flexible Algorithm ASBR Metric (IPFAAM) Sub-TLV is defined for the advertisement of the IP Flex-Algorithm-specific metric associated with an ASBR by the ABR.

The IPFAAM Sub-TLV is a sub-TLV of the:

- OSPFv2 Extended Inter-Area ASBR TLV, as defined in [RFC9350]
- OSPFv3 Inter-Area-Router TLV, as defined in [RFC8362]

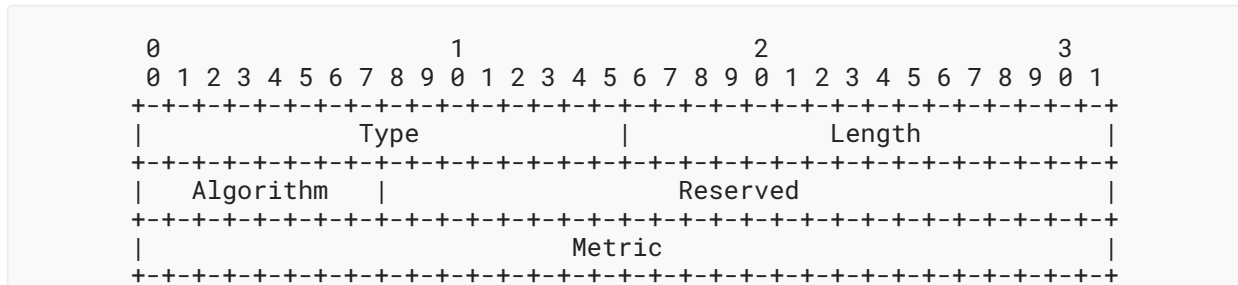The OSPF IPFAAM Sub-TLV has the following format:

```
  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             |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |   Algorithm   |                  Reserved                     |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |                            Metric                             |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 8: OSPF IP Flexible Algorithm ASBR Metric Sub-TLV

Where:

Type (2 octets):   2 (allocated by IANA) for OSPFv2, 36 for OSPFv3

Length (2 octets):   8

Algorithm (1 octet):   Associated Algorithm from 128 to 255

Reserved (3 octets):   **SHOULD** be set to 0 on transmission and **MUST** be ignored on reception

Metric (4 octets):   The algorithm-specific metric value

If the Algorithms in the OSPF IP Flexible Algorithm ASBR Metric Sub-TLV are outside the Flex-Algorithm range (128-255), the OSPF IP Flexible Algorithm ASBR Metric Sub-TLV **MUST** be ignored by the receiver. This situation **SHOULD** be logged as an error.

The usage of the IPFAAM Sub-TLV is similar to the usage of the FAAM Sub-TLV defined in [RFC9350], but it is used to advertise IP Flexible Algorithm metric.

An OSPF ABR **MUST** include the OSPF IPFAAM Sub-TLVs as part of any IP Flexible Algorithm ASBR reachability advertisement between areas.

The FAAM Sub-TLV as defined in [RFC9350] **MUST NOT** be used during IP Flexible Algorithm path calculation; the IPFAAM Sub-TLV **MUST** be used instead.

7. Calculating of IP Flexible Algorithm Paths

The IP Flexible Algorithm is considered as yet another data plane of the Flexible Algorithm as described in [RFC9350].

Participation in the IP Flexible Algorithm is signaled as described in Section 5 and is specific to the IP Flexible Algorithm data plane.

Calculation of IP Flexible Algorithm paths follows what is described in [RFC9350]. This computation uses the IP Flexible Algorithm data plane participation and is independent of the Flexible Algorithm calculation done for any other Flexible Algorithm data plane (e.g., SR, SRv6).

The IP Flexible Algorithm data plane only considers participating nodes during the Flexible Algorithm calculation. When computing paths for a given Flex-Algorithm, all nodes that do not advertise participation for such IP Flex-Algorithm, as described in Section 5, **MUST** be pruned from the topology.

8. IP Flexible Algorithm Forwarding

The IP Algorithm Prefix Reachability advertisement as described in Section 5 includes the MTID value that associates the prefix with a specific topology. Algorithm Prefix Reachability advertisement also includes an Algorithm value that explicitly associates the prefix with a specific Flex-Algorithm. The paths to the prefix **MUST** be calculated using the specified Flex-Algorithm in the associated topology.

Forwarding entries for the IP Flex-Algorithm prefixes advertised in IGPs **MUST** be installed in the forwarding plane of the receiving IP Flex-Algorithm prefix capable routers when they participate in the associated topology and algorithm. Forwarding entries for IP Flex-Algorithm prefixes associated with Flex-Algorithms in which the node is not participating **MUST NOT** be installed in the forwarding plane.

9. Deployment Considerations

IGP Flexible Algorithm can be used by many data planes. The original specification was done for SR and SRv6; this specification adds IP as another data plane that can use IGP Flexible Algorithm. Other data planes may be defined in the future. This section provides some details about the coexistence of the various data planes of an IGP Flexible Algorithm.

Flexible Algorithm Definition (FAD), as described in [RFC9350], is data plane independent and is used by all Flexible Algorithm data planes.

Participation in the Flexible Algorithm, as described in [RFC9350], is data plane specific.

Calculation of the Flexible Algorithm paths is data plane specific and uses data-plane-specific participation advertisements.

Data-plane-specific participation and calculation guarantee that the forwarding of the traffic over the Flex-Algorithm data-plane-specific paths is consistent between all nodes that apply the IGP Flex-Algorithm to the data plane.

Multiple data planes can use the same Flex-Algorithm value at the same time and, and as such, share the FAD for it. For example, SR-MPLS and IP can both use a common Flex-Algorithm. Traffic for SR-MPLS will be forwarded based on Flex-Algorithm-specific SR SIDs. Traffic for IP Flex-Algorithm will be forwarded based on Flex-Algorithm-specific prefix reachability advertisements. Note that for a particular Flex-Algorithm, for a particular IP prefix, there will only be path(s) calculated and installed for a single data plane.

10. Protection

In many networks where IGP Flexible Algorithms are deployed, IGP restoration will be fast and additional protection mechanisms will not be required. IGP restoration may be enhanced by Equal Cost Multipath (ECMP).

In other networks, operators can deploy additional protection mechanisms. The following are examples:

- Loop-Free Alternates (LFAs) [RFC5286]
- Remote Loop-Free Alternates (R-LFAs) [RFC7490]

LFA and R-LFA computations **MUST** be restricted to the Flex-Algorithm topology and the computed backup next hops should be programmed for the IP Flex-Algorithm prefixes.

11. IANA Considerations

This specification updates the "OSPF Router Information (RI) TLVs" registry as follows:

Value	TLV Name	Reference
21	IP Algorithm	RFC 9502, Section 5.2

Table 1

This document also updates the "IS-IS Sub-TLVs for IS-IS Router CAPABILITY TLV" registry as follows:

Value	TLV Name	Reference
29	IP Algorithm	RFC 9502, Section 5.1

Table 2

This document also updates the "IS-IS Top-Level TLV Codepoints" registry as follows:

Value	TLV Name	IIH	LSP	SNP	Purge	Reference
126	IPv4 Algorithm Prefix Reachability	n	y	n	n	RFC 9502, Section 6.1

Value	TLV Name	IIH	LSP	SNP	Purge	Reference
127	IPv6 Algorithm Prefix Reachability	n	y	n	n	RFC 9502, Section 6.2

Table 3

Since the above TLVs share the sub-TLV space managed in the "IS-IS Sub-TLVs for TLVs Advertising Prefix Reachability" registry, IANA has added "IPv4 Algorithm Prefix Reachability TLV (126)" and "IPv6 Algorithm Prefix Reachability TLV (127)" to the list of TLVs in the description of that registry.

In addition, columns headed "126" and "127" have been added to that registry, as follows:

Type	Description	126	127
1	32-bit Administrative Tag Sub-TLV	y	y
2	64-bit Administrative Tag Sub-TLV	y	y
3	Prefix Segment Identifier	n	n
4	Prefix Attribute Flags	y	y
5	SRv6 End SID	n	n
6	Flexible Algorithm Prefix Metric (FAPM)	n	n
11	IPv4 Source Router ID	y	y
12	IPv6 Source Router ID	y	y
32	BIER Info	n	n

Table 4

This document registers the following in the "OSPFv2 Extended Prefix TLV Sub-TLVs" registry:

Value	TLV Name	Reference
6	OSPFv2 IP Algorithm Prefix Reachability	RFC 9502, Section 6.3
7	OSPFv2 IP Forwarding Address	RFC 9502, Section 6.3.1

Table 5

IANA has created the "IP Algorithm Prefix Reachability Sub-TLV Flags" registry within the "Open Shortest Path First v2 (OSPFv2) Parameters" group of registries. The new registry defines the bits in the 8-bit Flags field in the OSPFv2 IP Algorithm Prefix Reachability Sub-TLV (Section 6.3). New bits can be allocated via IETF Review or IESG Approval [RFC8126]

Bit	Name	Reference
0	E bit	RFC 9502, Section 6.3
1-7	Unassigned	

Table 6

This document registers the following in the "OSPFv3 Extended-LSA Sub-TLVs" registry:

Value	Description	L2BM	Reference
35	OSPFv3 IP Algorithm Prefix Reachability	X	RFC 9502, Section 6.4
36	OSPFv3 IP Flexible Algorithm ASBR Metric	X	RFC 9502, Section 6.5

Table 7

This document registers the following in the "OSPFv2 Extended Inter-Area ASBR Sub-TLVs" registry:

Value	Description	Reference
2	OSPF IP Flexible Algorithm ASBR Metric	RFC 9502, Section 6.5

Table 8

12. Security Considerations

This document inherits security considerations from [RFC9350].

This document adds one new way to disrupt IGP networks that are using Flexible Algorithm: an attacker can suppress reachability for a given prefix whose reachability is advertised by a legitimate node for a particular IP Flex-Algorithm X by advertising the same prefix in Flex-Algorithm Y from another malicious node. (To see why this is, consider, for example, the rule given in the second-to-last paragraph of Section 6.1).

This attack can be addressed by the existing security extensions, as described in [RFC5304] and [RFC5310] for IS-IS, in [RFC2328] and [RFC7474] for OSPFv2, and in [RFC4552] and [RFC5340] for OSPFv3.

If a node that is authenticated is taken over by an attacker, such a rogue node can perform the attack described above. Such an attack is not preventable through authentication, and it is not different from advertising any other incorrect information through IS-IS or OSPF.

13. References

13.1. Normative References

[ISO10589] ISO, "Information technology - Telecommunications and information exchange between systems - Intermediate System to Intermediate System intra-domain routeing information exchange protocol for use in conjunction with the protocol for providing the connectionless-mode network service (ISO 8473)", Second Edition, ISO/IEC 10589:2002, November 2002.

[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>.

[RFC4552]   Gupta, M. and N. Melam, "Authentication/Confidentiality for OSPFv3", RFC 4552, DOI 10.17487/RFC4552, June 2006, <https://www.rfc-editor.org/info/rfc4552>.

[RFC4915]   Psenak, P., Mirtorabi, S., Roy, A., Nguyen, L., and P. Pillay-Esnault, "Multi-Topology (MT) Routing in OSPF", RFC 4915, DOI 10.17487/RFC4915, June 2007, <https://www.rfc-editor.org/info/rfc4915>.

[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>.

[RFC5304]   Li, T. and R. Atkinson, "IS-IS Cryptographic Authentication", RFC 5304, DOI 10.17487/RFC5304, October 2008, <https://www.rfc-editor.org/info/rfc5304>.

[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>.

[RFC5308]   Hopps, C., "Routing IPv6 with IS-IS", RFC 5308, DOI 10.17487/RFC5308, October 2008, <https://www.rfc-editor.org/info/rfc5308>.

[RFC5310]   Bhatia, M., Manral, V., Li, T., Atkinson, R., White, R., and M. Fanto, "IS-IS Generic Cryptographic Authentication", RFC 5310, DOI 10.17487/RFC5310, February 2009, <https://www.rfc-editor.org/info/rfc5310>.

[RFC5340]   Coltun, R., Ferguson, D., Moy, J., and A. Lindem, Ed., "OSPF for IPv6", RFC 5340, DOI 10.17487/RFC5340, July 2008, <https://www.rfc-editor.org/info/rfc5340>.

[RFC7474]   Bhatia, M., Hartman, S., Zhang, D., and A. Lindem, Ed., "Security Extension for OSPFv2 When Using Manual Key Management", RFC 7474, DOI 10.17487/RFC7474, April 2015, <https://www.rfc-editor.org/info/rfc7474>.

[RFC7770]   Lindem, A., Ed., Shen, N., Vasseur, JP., Aggarwal, R., and S. Shaffer, "Extensions to OSPF for Advertising Optional Router Capabilities", RFC 7770, DOI 10.17487/RFC7770, February 2016, <https://www.rfc-editor.org/info/rfc7770>.

[RFC7981]   Ginsberg, L., Previdi, S., and M. Chen, "IS-IS Extensions for Advertising Router Information", RFC 7981, DOI 10.17487/RFC7981, October 2016, <https://www.rfc-editor.org/info/rfc7981>.

[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>.

[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>.

[RFC9352]   Psenak, P., Ed., Filsfils, C., Bashandy, A., Decraene, B., and Z. Hu, "IS-IS Extensions to Support Segment Routing over the IPv6 Data Plane", RFC 9352, DOI 10.17487/RFC9352, February 2023, <https://www.rfc-editor.org/info/rfc9352>.

13.2. Informative References

[IANA-ALG] IANA, "IGP Algorithm Types", <https://www.iana.org/assignments/igp-parameters>.

[RFC5286]   Atlas, A., Ed. and A. Zinin, Ed., "Basic Specification for IP Fast Reroute: Loop-Free Alternates", RFC 5286, DOI 10.17487/RFC5286, September 2008, <https://www.rfc-editor.org/info/rfc5286>.

[RFC7490]   Bryant, S., Filsfils, C., Previdi, S., Shand, M., and N. So, "Remote Loop-Free Alternate (LFA) Fast Reroute (FRR)", RFC 7490, DOI 10.17487/RFC7490, April 2015, <https://www.rfc-editor.org/info/rfc7490>.

[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>.

[RFC8402]   Filsfils, C., Ed., Previdi, S., Ed., Ginsberg, L., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing Architecture", RFC 8402, DOI 10.17487/RFC8402, July 2018, <https://www.rfc-editor.org/info/rfc8402>.

[RFC8986]   Filsfils, C., Ed., Camarillo, P., Ed., Leddy, J., Voyer, D., Matsushima, S., and Z. Li, "Segment Routing over IPv6 (SRv6) Network Programming", RFC 8986, DOI 10.17487/RFC8986, February 2021, <https://www.rfc-editor.org/info/rfc8986>.

[TS.23.501-3GPP] 3GPP, "System architecture for 5G System (5GS)", Release 18.3.0, 3GPP TS 23.501, September 2023.

Acknowledgements

Authors' Addresses

William Britto
Juniper Networks
Elnath-Exora Business Park Survey
Bangalore 560103
Karnataka
India
Email: bwilliam@juniper.net

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

Parag Kaneriya
Juniper Networks
Elnath-Exora Business Park Survey
Bangalore 560103
Karnataka
India
Email: pkaneria@juniper.net

Rejesh Shetty
Juniper Networks
Elnath-Exora Business Park Survey
Bangalore 560103
Karnataka
India
Email: mrajesh@juniper.net

Ron Bonica
Juniper Networks
2251 Corporate Park Drive
Herndon, Virginia 20171
United States of America
Email: rbonica@juniper.net

Peter Psenak
Cisco Systems
Apollo Business Center
Mlynske nivy 43
82109 Bratislava
Slovakia
Email: ppsenak@cisco.com