Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: January 18, 2018

P. Psenak, Ed.
Cisco Systems
S. Hegde, Ed.
Juniper Networks, Inc.
C. Filsfils
Cisco Systems, Inc.
A. Gulko
Thomson Reuters

July 17, 2017

# ISIS Segment Routing Flexible Algorithm
**draft-hegdeppsenak-isis-sr-flex-algo-00.txt**

## Abstract

IGP protocols traditionally compute best paths over the network based on the IGP metric assigned to the links. Many network deployments use RSVP based or Segment Routing based Traffic Engineering to enforce traffic over a path that is computed using different metrics or constrains then IGP path. Various mechanisms are used to steer the traffic towards such traffic engineered paths. This document proposes a solution that allows IGPs itself to compute constrained based path over the network without the usage of the traffic engineering.

## Status of This Memo



## Copyright Notice

Table of Contents



## 1. Introduction

IGP computed path based on the shortest IGP metric must often be replaced by traffic engineered path due to the traffic requirements which are not reflected in the IGP metric. Some networks engineer the IGP metric assignments in a way that the IGP Metric reflects the link bandwidth or delay. If, for example, the IGP metric is reflecting the bandwidth on the link and the application traffic is delay sensitive, the best IGP path may not reflect the best path from such application's perspective.

To overcome such IGP limitation, various sorts of traffic engineering has been deployed, including RSVP-TE or SR-TE, in which case the TE component is responsible for computing the path based on some other or additional metrics and/or constrains. Such paths need to be installed in the forwarding and replace the original paths computed by IGPs. Tunnels are often used to represent the engineered paths

and mechanisms like one described in [RFC3906] are used to replace the native IGP paths with such tunnel paths.

Segment Routing (SR) allows a flexible definition of end-to-end paths within IGP topologies by encoding paths as sequences of topological sub-paths, called segments. It also defines an algorithm that defines how the path is computed. It also provides a way to associate Prefix-SID with an algorithm. This allows IGPs to compute the path based on various algorithms and forward the traffic on a such path using the algorithm specific segments.

This document describes the IS-IS extension to support Segment Routing Flexible Algorithm on an MPLS data-plane.

### 1.1. Requirements notation

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119].

## 2. Flexible Algorithm

Many possible constrains may be used to compute a path over a network. Some networks are deployed as multiple planes. A simple form of constrain may be to use a particular plane. A more sophisticated form of constrain can include some extended metric as described in [RFC7810]. Even more advanced case could be to restrict the path and avoid links with certain affinities. Combinations of these are also possible.

To provide a maximum flexibility we do not want to provide a strict mapping between the set of constrains and the algorithm that is associated with it. We want the mapping between the algorithm value and it's meaning to be flexible and defined by the user. As far as all routers in the domain has the common understanding what the particular algorithm value represents, the computation for such algorithm is consistent and traffic is not subject to the looping.

Because the meaning of the algorithm is not defined by any standard, but is defined by the user, we call it Flex-Algorithm.

## 3. Flexible Algorithm Advertisement

[I-D.ietf-isis-segment-routing-extensions] defines an SR-Algorithm. This algorithm defines how the best path is computed by IGP. Routers advertise the support for the algorithm as a node capability. Prefix SIDs are also advertised with an algorithm value and as such are tightly coupled with the algorithm.

Existing advertisement of the SR-Algorithm is used for the Flex-Algorithm advertisements as defined in [I-D.ietf-isis-segment-routing-extensions].

SR-Algorithm is a one octet value. We propose to split the range of values as follows:

0-127 - standardised values provided by IANA

128-255 - user defined values

## 4. Flexible Algorithm Definition Advertisement

To guarantee the loop free forwarding for paths computed for a particular Flex-Algorithm, all routers in the network MUST share the same definition of the Flex-Algorithm. This can be achieved by each router advertising its definition of each Flex-Algorithm that is locally defined and detect any conflicts in the Flex-Algorithm definition between routers.

Alternatively, the central entity in the network can advertise the definition of the Flex-Algorithm and let all routers to use it.

Two definitions of the Flex-Algorithm are considered to match if all of the following conditions are met:

Metric Type for both definitions is the same.

The set of Admin Groups that are excluded is exactly the same in both definitions.

### 4.1. Flexible Algorithm Definition TLV

Flexible Algorithm Definition TLV (FAD TLV) is used to advertise the definition of the Flex-Algorithm.

FAD TLV can be advertised as:

Sub-TLV of the IS-IS Router Capability TLV-242 that is defined in [RFC7981]. When advertised as Sub-TLV of the IS-IS Router Capability TLV-242, it is used to advertise the local definition of the Flex-Algorithm on the originating router.

ISIS top-level TLV. When advertised as top-level TLV, it is used to inform routers in entire domain about the definition of the Flex-Algorithm.

FAD 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     |     Flags     |   Algorithm   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  Metric Type  |             MTID              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                Flex-Algorithm Exclude sub-TLVs                |
+-                                                             -+
|                            ...                                |

|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

where:

Type:

When advertised as Sub-TLV of the IS-IS Router Capability TLV-242: TBD1

When advertised as ISIS top-level TLV: TBD2

Length: variable, dependent on the number of Sub-TLVs

Flags: Single octet field containing the following flags:

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

where:

S-Flag: If set, the FAD top-level TLV SHOULD be flooded across the entire routing domain.  If the S flag is not set, the FAD TLV MUST NOT be leaked between levels.  This bit MUST NOT be altered during the TLV leaking.  This bit MUST be ignored in the FAD Sub-TLV of the IS-IS Router Capability TLV-242.

D-Flag: when the FAD top-level TLV is leaked from level-2 to level-1, the D bit MUST be set.  Otherwise, this bit MUST be clear.  FAD top-level TLVs with the D bit set MUST NOT be leaked from level-1 to level-2.  This is to prevent TLV looping across levels.  This bit MUST be ignored in the FAD Sub-TLV of the IS-IS Router Capability TLV-242.

Algorithm: Flex-Algorithm number. Value between 128 and 255 inclusive.

Metric Type: Type of metric to be used during the calculation. Following values are defined:

0: IGP Metric

1: Min Unidirectional Link Delay as defined in [RFC7810].

2: TE metric as defined in [RFC5305].

MTID: Multitopology identifier defined as::

```
 0                   1
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| RESVD |         MTID          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

where:

RESVD: reserved bits. MUST be reset on transmission and ignored on receive.

MTID: a 12-bit field containing the non-zero ID of the topology being announced. The TLV MUST be ignored if the ID is zero. This is to ensure the consistent view of the standard unicast topology.

Flex-Algorithm Exclude sub-TLVs - optional sub-TLVs as described in section Section 4.2.

When the router is configured with the local definition of the Flex-Algorithm, the router SHOULD advertise its local definition in the FAD Sub-TLV of the IS-IS Router Capability TLV-242. If the local definition of the Flex-Algorithm is not advertised, the inconsistency in the configuration of the Flex-Algorithm on various nodes can not be detected and traffic routed based on a Flex-Algorithm path may loop permanently.

When the router receives the FAD TLV as top-level TLV, it uses it as a definition of the Flex-Algorithm. If the local definition of the same Flex-Algorithm exists on the router and is in conflict with the definition received over top-level FAD TLV, the router MUST NOT compute any path for such Flex-Algorithm and MUST stop advertising support for such Flex-Algorithm in its SR-Algorithm Sub-TLV ([I-D.ietf-isis-segment-routing-extensions]).

When router receives the FAD Sub-TLV of the IS-IS Router Capability TLV-242 from multiple sources and the Flex-Algorithm definition in these advertisements are conflicting, it MUST NOT compute any path for such Flex-Algorithm and MUST stop advertising support for such Flex-Algorithm in its SR-Algorithm Sub-TLV ([I-D.ietf-isis-segment-routing-extensions]).

When router receives the FAD Sub-TLV of the IS-IS Router Capability TLV-242 from another router and the definition is in conflict with either the local definition of the Flex-Algorithm OR the definition received in the FAD top-level TLV, it MUST NOT compute any path for such Flex-Algorithm.

The FAD Sub-TLV of the IS-IS Router Capability TLV-242 MUST be propagated throughout the level and MUST be advertised across level boundaries.  Therefore Router Capability TLV distribution flags SHOULD be set accordingly, i.e., the S flag in the Router Capability TLV MUST be set.

### 4.2. Flexible Algorithm Exclude Admin Group Sub-TLV

To provide even more granularity, the Flexible-Algorithm can include link 'colors' that the operator wants to exclude from the computation.  This provides a per link granularity for the Flex-Algorithm definition.

Flexible Algorithm Exclude Admin Group Sub-TLV (FAEAG Sub-TLV) is a Sub-TLV of the 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     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      Extended Admin Group                      |
+-                                                              -+
|                             ...                                |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

where:

Type: TBD3

Length: variable, dependent on the size of the Extended Admin Group.  MUST be a multiple of 4 octets.

Extended Administrative Group: Extended Administrative Group as defined in [RFC7308].

FAEAG Sub-TLV SHOULD only appear once in FAD TLV. If it appears more then once, FAD TLV MUST be ignored by the receiver.

## 5. Calculation of Flexible Algorithm Paths

A router may compute path for multiple Flex-Algorithms.

A router MUST be configured to support Flex-Algorithm K before it can compute any path for Flex-Algorithm K.

A router MUST either be configured with a local definition of Flex-Algorithm K or receive the definition via the FAD top-level TLV as described in Section 4.1 from the central entity that acts as the Flex-Algorithm definition holder before it can compute any path for Flex-Algorithm K.

If any conflicts in the Flex-Algorithm K definition exists, as described in Section 4.1, the router MUST NOT compute any path for Flex-Algorithm K.

When computing the Shortest Path Tree for Flex-Algorithm K, all nodes that do not advertise support for Flex-Algorithm K in SR-Algorithm Sub-TLV ([I-D.ietf-isis-segment-routing-extensions]), MUST be pruned from the topology.

When computing the Shortest Path Tree for Flex-Algorithm K, any link advertised with any of the corresponding bits in both (Extended) Administrative Groups sub-TLV and FAEAG Sub-TLV set to 1, MUST be pruned from the topology.

When computing the Shortest Path Tree for Flex-Algorithm K, router MUST use the metric that is part of the Flex-Algorithm definition. If the metric is not advertised for the particular link, such link MUST be pruned from the topology. A metric of value 0 MUST NOT be assumed in such case.

Flex-Algorithm K path to any prefix MUST be installed in the forwarding using the Prefix-SID that was advertised for algorithm K. If the Prefix SID for algorithm K is not known, Flex-Algorithm K path to such prefix MUST NOT be installed in the forwarding.

Loop Free Alternate (LFA) paths for Flex-Algorithm K path MUST be computed using the same constrains as the calculation of the primary paths for Flex-Algorithm K. LFA path MUST only use Prefix-SIDs advertised specifically for algorithm K to enforce the traffic over such path.

Any Shortest Path Tree calculation is limited to a single area. Same applies to Flex-Algorithm calculations. Given that the computing router may not have the visibility to the topology of remote areas, the Flex-Algorithm K path to inter-area prefix will only be computed for the local area. The 'exit' L1/L2 router will be selected based on the best path for the Flex-Algorithm K in the local area and such 'exit' L1/L2 router will be responsible to compute the best Flex-Algorithm K path over the next area. This may produce end-to-end path, which is not the best from the Flex-Algorithm K perspective. If the best end-to-end path for Flex-Algorithm K needs to be used for inter-area destinations, paths for such destinations need to be computed by the entity that has the topological information about all areas.

## 6. Backward Compatibility

This extension brings no new backward compatibility issues.

## 7. Security Considerations

This extension adds no new security considerations.

## 8. IANA Considerations

This documents request allocation for the following TLVs and subTLVs.

### 8.1. Sub TLVs for Type 242

This document makes the following registrations in the "sub-TLVs for TLV 242" registry.

Type: TBD1 (suggested value 24).

Description: Flexible Algorithm Definition Sub-TLV.

Reference: This document (Section 4.1).

### 8.2. New TLV Codepoint and Sub-TLV registry

This document registers the following TLV:

Type: TBD2 (suggested value 151)

name: Flexible Algorithm Definition TLV.

IIH: no

LSP: yes

SNP: no

Purge: no

Reference: This document (Section 4.1).

This document creates the following sub-TLV Registry:

Registry: sub-TLVs for TLV 151

Registration Procedure: Expert review

Reference: This document (Section 4.1)

This document resisters following TLV in the "sub-TLVs for TLV 151" registry

Type: TBD3, suggested value 1.

Description: Flexible Algorithm Exclude Admin Group Sub-TLV.

Reference: This document (Section 4.2).

## 9. Acknowledgments

This draft, among other things, is also addressing the problem that the [I-D.gulkohegde-routing-planes-using-sr] was trying to solve. All authors of that draft agreed to join this draft.

Thanks to Les Ginsberg for review and useful comments on the initial version of the draft.

## 10. References

### 10.1. Normative References

[I-D.ietf-isis-segment-routing-extensions]
Previdi, S., Filsfils, C., Bashandy, A., Gredler, H., Litkowski, S., Decraene, B., and j. jefftant@gmail.com, "IS-IS Extensions for Segment Routing", draft-ietf-isis-segment-routing-extensions-13 (work in progress), June 2017.

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

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

[RFC7308]  Osborne, E., "Extended Administrative Groups in MPLS Traffic Engineering (MPLS-TE)", RFC 7308, DOI 10.17487/RFC7308, July 2014, <http://www.rfc-editor.org/info/rfc7308>.

[RFC7810]  Previdi, S., Ed., Giacalone, S., Ward, D., Drake, J., and Q. Wu, "IS-IS Traffic Engineering (TE) Metric Extensions", RFC 7810, DOI 10.17487/RFC7810, May 2016, <http://www.rfc-editor.org/info/rfc7810>.

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

## 10.2. Informative References

[I-D.gulkohegde-routing-planes-using-sr]
Hegde, S. and a. arkadiy.gulko@thomsonreuters.com, "Separating Routing Planes using Segment Routing", draft-gulkohegde-routing-planes-using-sr-00 (work in progress), March 2017.

[RFC3906]  Shen, N. and H. Smit, "Calculating Interior Gateway Protocol (IGP) Routes Over Traffic Engineering Tunnels", RFC 3906, DOI 10.17487/RFC3906, October 2004, <http://www.rfc-editor.org/info/rfc3906>.

Authors' Addresses

Peter Psenak (editor)
Cisco Systems
Apollo Business Center
Mlynske nivy 43
Bratislava, 82109
Slovakia

Email: ppsenak@cisco.com

Shraddha Hegde (editor)
Juniper Networks, Inc.
Embassy Business Park
Bangalore, KA, 560093
India

Email: shraddha@juniper.net

Clarence Filsfils
Cisco Systems, Inc.
Brussels
Belgium

Email: cfilsfil@cisco.com

Arkadiy Gulko
Thomson Reuters

Email: arkadiy.gulko@thomsonreuters.com