Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: 4 September 2025

Y. Wang
G. Xu
X. Geng
J. Dong
Huawei
P. Psenak
Cisco Systems
3 March 2025

IGP Flexible Algorithm with Link Loss
draft-wang-lsr-flex-algo-link-loss-04

Abstract

This document proposes extensions to the IGP Flexible Algorithms framework defined in [RFC9350]. It introduces a mechanism to exclude links exceeding a specified packet loss rate threshold during path computation. The solution leverages existing link loss measurements advertised via IS-IS [RFC8570] and OSPF [RFC7471], and defines new constraints for Flex-Algorithm path calculation.

Requirements Language

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

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

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

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

This Internet-Draft will expire on 4 September 2025.

Table of Contents

1. Introduction

Link packet loss rate (hereafter "link loss") refers to the percentage of data packets that are lost during transmission over a network. It is a critical metric for network performance evaluation. High loss rates directly impact service quality, congestion management, and operational efficiency. To maintain optimal forwarding paths, it is essential to avoid links with excessive packet loss during IGP path computation.

The IGP Flexible Algorithms enable IGPs to compute constraint-based paths [RFC9350]. Current path computation methods focus on determining the minimum cost of the path from the source to the destination. Flex-Algorithm already supports path computation based on IGP cost, minimum link delay, and traffic-engineering metrics. [I-D.ietf-lsr-flex-algo-bw-con] defines a family of generic metrics

(e.g., bandwidth-based metric type) and bandwidth-related constraints to enable path computation based on bandwidth. However, current flexible algorithm definitions lack native support for path computation based on link loss, as the path cost should be defined as the maximum or minimum value among all links in the path.

To address this issue, two solutions are considered. First, new operators, such as a maximum value operator, can be defined to function as a step function. Specifically, when the link loss exceeds a threshold, the link cost is set to the maximum value. Second, new Flexible Algorithm Definition (FAD) constraints can be defined to exclude links that do not meet the link loss requirements during path calculation. The second method is specifically demonstrated in this document, and the general ideas are as follows:

1. The link loss is used as a link constraint for path computation. That is, links with a loss rate exceeding the specified value are excluded.

2. Metric-type remains unchanged: igp, te, and delay.

This document proposes the method to exclude links exceeding a specified packet loss rate by defining:

a) A new Flexible Algorithm Definition (FAD) constraint to exclude links exceeding a configured maximum loss threshold (Section 2).

b) Operational procedures for integrating loss constraints into Flex-Algorithm path computations (Section 3).

c) Mechanisms to stabilize routing during loss metric fluctuations (Section 4).

The solution reuses existing link loss advertisements defined in [RFC8570] for IS-IS and [RFC7471] for OSPF, ensuring backward compatibility with deployed networks. The link packet loss rate can be measured using methods such as TWAMP [RFC5357] and STAMP [RFC8762]. However, these measurement techniques are beyond the scope of this document. It is important to ensure that link-loss measurements are consistent throughout the IGP routing domain.

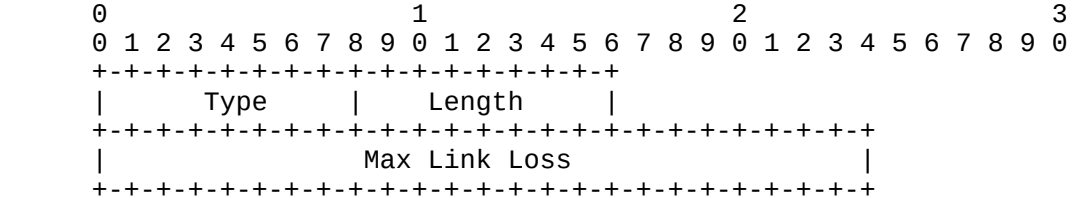
2. Exclude Maximum Link Loss Sub-TLV

A new sub-TLV, the Exclude Maximum Link Loss Sub-TLV, is defined as part of the FAD TLV. To ensure loop-free forwarding, all routers participating in a Flex-Algorithm MUST agree on the FAD definition. Selected nodes within the IGP domain MUST advertise FADs by including them in their routing updates, as specified in Sections 5, 6, and 7 of [RFC9350].

The Exclude Maximum Link Loss Sub-TLV is introduced to define the maximum allowable link loss value. When this Sub-TLV is carried within the FAD TLV, all network links with packet loss rates exceeding the specified maximum value are excluded from the Flex-Algorithm path computation.

2.1. IS-IS Exclude Maximum Link Loss Sub-TLV

The IS-IS Flex-Algorithm Exclude Maximum Link Loss Sub-TLV (FAEML) is defined as a sub-TLV of the IS-IS FAD Sub-TLV. The format follows standard TLV structure:



Type: 252(TBA by IANA)

Length: 3 octets

Max Link Loss: 24-bit unsigned integer representing the maximum allowable loss percentage. Encoded with a resolution of 0.000003% per unit, providing a maximum expressible value of 50.331642% (0xFFFFFF * 0.000003). Values exceeding this cap MUST be advertised as 0xFFFFFF.

Figure 1: IS-IS FAEML Sub-TLV

The FAEML sub-TLV MUST appear at most once in the FAD Sub-TLV. If it appears more than once, the IS-IS FAD Sub-TLV MUST be ignored by the receiving node.

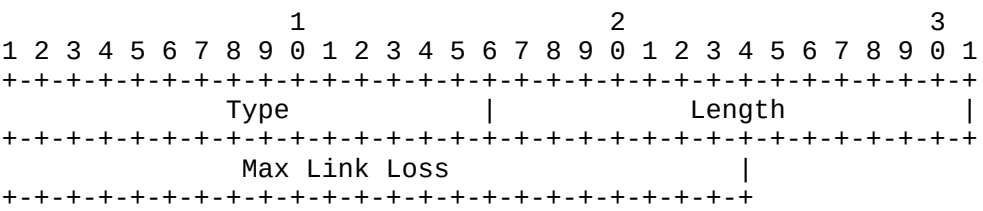
The maximum link loss advertised in the FAEML Sub-TLV MUST be compared with the link loss advertised in Sub-Sub-TLV 36 [RFC8570] of ASLA Sub-TLV [RFC9479]. If the L-Flag is set in the ASLA sub-TLV,

the maximum link loss advertised in the FAEML sub-TLV MUST be compared with the link loss advertised by the sub-TLV 36 of the TLV 22/222/23/223/141 [RFC5305] as defined in [RFC9479] Section 4.2.

If the link loss exceeds the maximum link loss advertised in the FAEML sub-TLV, the link MUST be excluded from the Flex-Algorithm topology. However, if a link does not advertise the link loss but the FAD contains the FAEML sub-TLV, the link MUST NOT be excluded from the Flex-Algorithm topology.

2.2. OSPF Exclude Maximum Link Loss Sub-TLV

The OSPF Flex-Algorithm Exclude Maximum Link Loss Sub-TLV (FAEML) is defined as a sub-TLV of the OSPF FAD Sub-TLV. The format follows standard TLV structure:



Type: 252(TBA)

Length: 3 octets

Max Link Loss: 24-bit unsigned integer representing the maximum allowable loss percentage. Encoded with a resolution of 0.000003% per unit, providing a maximum expressible value of 50.331642% (0xFFFFFF * 0.000003). Values exceeding this cap MUST be advertised as 0xFFFFFF.

Figure 2: OSPF FAEML Sub-TLV

The FAEML sub-TLV MUST appear at most once in the FAD Sub-TLV. If it appears more than once, the OSPF FAD Sub-TLV MUST be ignored by the receiving node.

The maximum link loss advertised in the FAEML Sub-TLV MUST be compared with the link loss advertised in Sub-Sub-TLV 30 [RFC7471] of the ASLA Sub-TLV [RFC9492]. The ASLA Sub-TLV is advertised in Extended Link Opaque LSAs [RFC7684] for OSPFv2 and E-Router-LSAs [RFC8362] for OSPFv3.

If the link loss exceeds the maximum link loss advertised in the FAEML sub-TLV, the link MUST be excluded from the Flex-Algorithm topology. However, if a link does not advertise the link loss but the FAD contains the FAEML sub-TLV, the link MUST NOT be excluded from the Flex-Algorithm topology.

3. Calculation of Flexible Algorithm Paths

The following rule is added to the topology pruning rules in Section 13 of [RFC9350]:

1. Check if any exclude FAEML rule is part of the Flex-Algorithm definition. If such exclude rule exists and the link has link loss advertised, check if the link satisfies the FAEML rule. If not, the link MUST be pruned from the computation.

4. Operational Considerations

In certain scenarios, the link status may fluctuate between available and unavailable due to the link packet loss rate oscillating around the threshold value. Consequently, Flex-Algorithm computation may be triggered repeatedly. Several mechanisms are considered to address this issue:

1. Delayed collection: The IGP-advertised loss can be calculated over a sufficiently long interval, such as 10 minutes, to reduce the frequency of updates.

2. Averaging and normalization: The IGP-advertised loss should be derived from a form of averaging, such as an exponential weighted average, of the collected loss values. The advertised loss can be normalized to prevent the dissemination of non-significant changes in loss metrics.

3. Flapping suppression: If frequent changes in the IGP-advertised loss are detected, a timer can be implemented to delay the update process, thereby stabilizing the routing computations.

5. IANA Considerations

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

Type: 252(TBA)

Description: IS-IS Exclude Maximum Link Loss Sub-TLV

Reference: This document Section 2.1

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

Type: 252(TBA)

Description: OSPF Exclude Maximum Link Loss Sub-TLV

Reference: This document Section 2.2

6. References

6.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://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, <https://www.rfc-editor.org/info/rfc5305>.

[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>.

[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>.

[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>.

6.2. Informative References

[I-D.ietf-lsr-flex-algo-bw-con] Hegde, S., Britto, W., Shetty, R., Decraene, B., Psenak, P., and T. Li, "IGP Flexible Algorithms: Bandwidth, Delay, Metrics and Constraints", Work in Progress, Internet-Draft, draft-ietf-lsr-flex-algo-bw-con-22, 13 February 2025, <https://datatracker.ietf.org/doc/html/draft-ietf-lsr-flex-algo-bw-con-22>.

[RFC5357] Hedayat, K., Krzanowski, R., Morton, A., Yum, K., and J. Babiarz, "A Two-Way Active Measurement Protocol (TWAMP)", RFC 5357, DOI 10.17487/RFC5357, October 2008, <https://www.rfc-editor.org/info/rfc5357>.

[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>.

[RFC8762] Mirsky, G., Jun, G., Nydell, H., and R. Foote, "Simple Two-Way Active Measurement Protocol", RFC 8762, DOI 10.17487/RFC8762, March 2020, <https://www.rfc-editor.org/info/rfc8762>.

Authors' Addresses

Yifan Wang
Huawei
Huawei Bld., No. 156 Beiqing Rd.
Beijing
100095
China
Email: wangyifan82@huawei.com

Guoqi Xu
Huawei

Email: xuguoqi@huawei.com

Xuesong Geng
Huawei
Email: gengxuesong@huawei.com

Jie Dong
Huawei
Email: jie.dong@huawei.com

Peter Psenak
Cisco Systems
Email: ppsenak@cisco.com