IDR Working Group
Internet Draft
Intended status: Standards Track
Expires: October 24, 2025

Y. Liu
China Mobile
C. Lin
New H3C Technologies
J. Li
Y. He
China Mobile

April 24, 2025

BGP - Link State (BGP-LS) Advertisement of BGP Egress Peer Engineering Performance Metric Extensions

draft-liu-idr-bgpls-epe-te-pm-01

Abstract

This document specifies a method for advertising BGP Egress Peer Engineering (EPE) Traffic Engineering (TE) Performance Metric information via BGP-LS.

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 October 24, 2025.

Table of Contents

1. Introduction

In certain networks, such as financial information networks (e.g., stock market data providers), network performance metrics like link propagation delay are becoming as critical for data path selection as traditional metrics.

Consequently, metrics like hop count or cost are becoming less central to routing decisions. Instead, it would be beneficial to make path selection decisions based on network performance information, such as link propagation delay, in a cost-effective and scalable manner.

This document describes extensions to BGP Egress Peer Engineering (EPE) [RFC9086] Traffic Engineering (TE) Performance Metric (hereafter called "BGP EPE TE PM Extensions"), that can be used to distribute network performance information (viz link propagation delay, delay variation, link loss, residual bandwidth, available bandwidth, utilized bandwidth, reverse residual bandwidth, reverse available bandwidth, reverse utilized bandwidth).

Note that the mechanisms described in this document solely focus on distributing network performance information. The methods for measuring this information or acting upon it once distributed are beyond the scope of this document.

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. Problem Statement

2.1. Lack of EPE TE Metric Extensions

[RFC9087] describes a centralized controller-based BGP-EPE solution involving SR path computation using BGP Peering Segments. A centralized controller learns the BGP Peering SIDs via Border Gateway Protocol - Link State (BGP-LS) and then uses this information to implement a BGP-EPE policy. [RFC9086] defines the extension to BGP-LS for advertising BGP Peering Segments along with their BGP peering node information.

[RFC8571] introduces new BGP-LS TLVs to carry the IGP Traffic Engineering Metric Extensions defined in the IS-IS and OSPF protocols. These Traffic Engineering Metric Extensions [RFC8571] are also needed when the controller performs BGP-EPE traffic scheduling.

Currently, BGP-LS lacks the capability to distribute BGP EPE TE PM Extensions information.

2.2. Bidirectional EPE TE Metric Scenario

As shown in Figure 1, an internal router establishes a BGP adjacency with an external router. Typically, both the internal and external routers establish BGP-LS connections with the controller and advertise their respective EPE TE PM Extensions. However, in some cases, due to different operational domains, the controller can only establish a BGP-LS connection with the internal router. In this scenario, the internal router can advertise EPE TE PM Extensions to the controller via BGP-LS, but the controller cannot obtain the EPE TE PM Extensions from the external router.

Therefore, there is a need for border devices to report bidirectional BGP EPE TE PM Extensions.

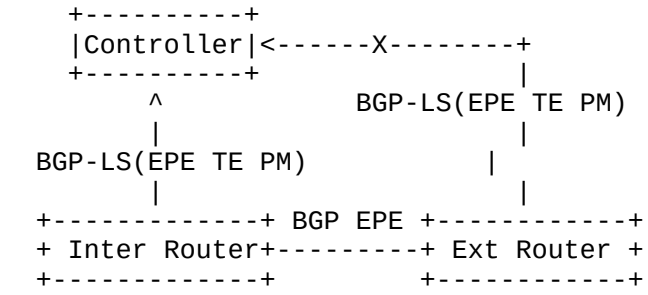


Figure 1

3. Solution

This document updates [RFC8571] to enable the addition of Traffic Engineering Performance Metric Extensions TLVs to the BGP-LS Attribute associated with the Link NLRI of a BGP peering link. It introduces reverse Traffic Engineering Performance Metric Extensions TLVs (as defined in Section 4.2) to advertise BGP EPE reverse TE PM Extensions, addressing the issue outlined in Section 2.2.

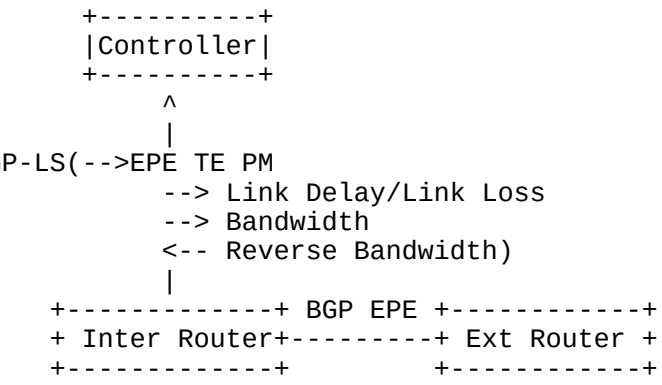


Figure 2

4. Advertising EPE TE PM Extensions in BGP-LS

4.1. Traffic Engineering Performance Metric Extensions

This document reuses the Traffic Engineering Performance Metric Extensions TLVs defined in [RFC8571] to carry BGP EPE TE PM Extensions through BGP-LS in the BGP protocol.

The BGP-EPE Peer Adjacency Traffic Engineering Performance Metric Extensions are advertised with a BGP-LS Link NLRI, where:

* BGP-LS Link NLRI: as described in Section 5.2 of [RFC9086].

* Link Attribute TLVs:

  - include the Unidirectional Link Delay TLV [RFC8571]

  - include the Min/Max Unidirectional Link Delay TLV [RFC8571]

  - include the Unidirectional Delay Variation TLV [RFC8571]

  - include the Unidirectional Link Loss TLV [RFC8571]

  - include the Residual Bandwidth TLV [RFC8571]

  - include the Available Bandwidth TLV [RFC8571]

  - include the Utilized Bandwidth TLV [RFC8571]

4.2. Reverse Traffic Engineering Performance Metric Extensions

This document defines new Reverse Traffic Engineering Performance Metric Extensions TLVs, which can be carried in BGP-LS Peer Link NLRI.

The BGP-EPE Peer Adjacency Reverse Traffic Engineering Performance Metric Extensions are advertised with a BGP-LS Link NLRI, where:

* BGP-LS Link NLRI: as described in Section 5.2 of [RFC9086].

* Link Attribute TLVs:

  - include the Reverse Residual Bandwidth TLV (Section 4.2.1)

  - include the Reverse Available Bandwidth TLV (Section 4.2.2)

  - include the Reverse Utilized Bandwidth TLV (Section 4.2.3)

The following new Link Attribute TLVs are defined:

TLV Code Point	Description
TBD1	Reverse Residual Bandwidth
TBD2	Reverse Available Bandwidth
TBD3	Reverse Utilized Bandwidth

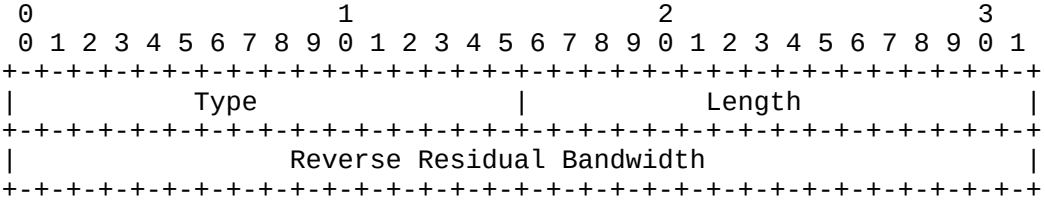
TLV formats are described in detail in the following subsections.

TLV formats follow the rules defined in [RFC9552].

4.2.1. Reverse Residual Bandwidth TLV

This TLV advertises the reverse residual bandwidth between two directly connected BGP neighbors.

The Reverse Residual Bandwidth TLV format is defined as follows:



where:

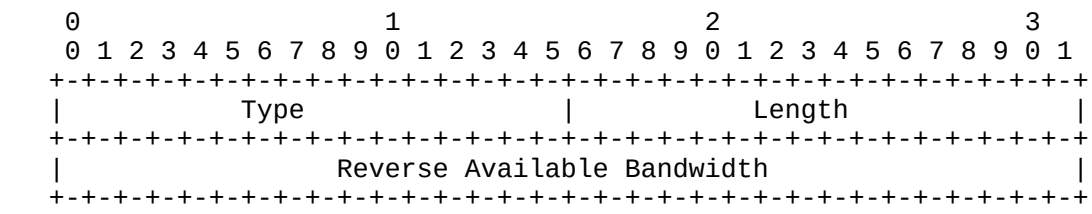
o Type: TBD1.

o Length: 4.

4.2.2. Reverse Available Bandwidth TLV

This TLV advertises the Reverse Available Bandwidth between two directly connected BGP neighbors.

The Reverse Available Bandwidth TLV format is defined as follows:



where:

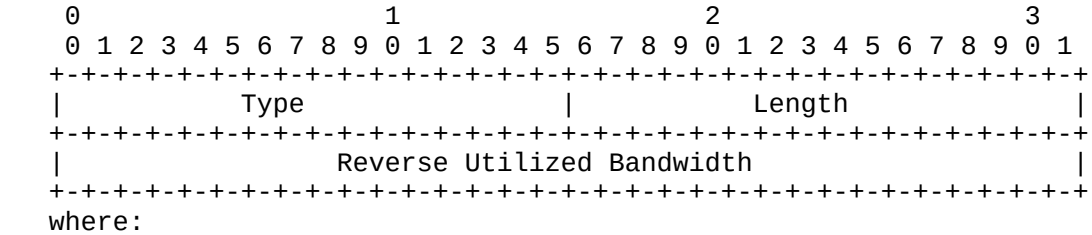
o Type: TBD2.

o Length: 4.

4.2.3. Reverse Utilized Bandwidth TLV

This TLV advertises the Reverse Utilized Bandwidth between two directly connected BGP neighbors.

The Reverse Available Bandwidth TLV format is defined as follows:



where:

o Type: TBD3.

o Length: 4.

5. Security Considerations

The security considerations described in [RFC9552] and [RFC9086] also apply to this document.

This document does not introduce any new security consideration.

6. IANA Considerations

IANA is requested to assign the new Link Attribute TLVs in the "BGP-LS Node Descriptor, Link Descriptor, Prefix Descriptor, and Attribute TLVs" registry as below:

TLV Code Point	Description
TBD1	Reverse Residual Bandwidth
TBD2	Reverse Available Bandwidth
TBD3	Reverse Utilized Bandwidth

7. References

7.1. 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>.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, May 2017.

[RFC8571] L. Ginsberg, Ed., S. Previdi, Q. Wu, J. Tantsura, Apstra, Inc., C. Filsfils, "BGP - Link State (BGP-LS) Advertisement of IGP Traffic Engineering Performance Metric Extensions", RFC 8571, DOI 10.17487/RFC8571, March 2019, <https://www.rfc-editor.org/info/rfc8571>.

[RFC9085] Previdi, S., Talaulikar, K., Ed., Filsfils, C., Gredler, H., and M. Chen, "Border Gateway Protocol - Link State (BGP-LS) Extensions for Segment Routing", RFC 9085, DOI 10.17487/RFC9085, August 2021, <https://www.rfc-editor.org/info/rfc9085>.

[RFC9086] S. Previdi, K. Talaulikar, Ed., C. Filsfils, K. Patel, Arrcus, Inc., S. Ray, Individual, J. Dong, "Border Gateway Protocol - Link State (BGP-LS) Extensions for Segment Routing BGP Egress Peer Engineering", RFC 9086, DOI 10.17487/RFC9086, August 2021, <https://www.rfc-editor.org/info/rfc9086>.

[RFC9552] K. Talaulikar, "Distribution of Link-State and Traffic Engineering Information Using BGP", RFC 9552, DOI 10.17487/RFC9552, December 2023, <https://www.rfc-editor.org/info/rfc9552>.

7.2. Informational References

[RFC9087] Filsfils, C., Ed., Previdi, S., Dawra, G., Ed., Aries, E., and D. Afanasiev, "Segment Routing Centralized BGP Egress Peer Engineering", RFC 9087, DOI 10.17487/RFC9087, August 2021, <https://www.rfc-editor.org/info/rfc9087>.

Authors' Addresses

Yisong Liu
China Mobile
Beijing
China

Email: liuyisong@chinamobile.com

Changwang Lin
New H3C Technologies
Beijing
China

Email: linchangwang.04414@h3c.com

Jinming Li
China Mobile
Beijing
China

Email: lijinming@chinamobile.com

Ying He
China Mobile
Beijing
China

Email: heyingyjy@chinamobile.com