

Inter-Domain Routing
Internet-Draft
Intended status: Standards Track
Expires: 16 October 2025

C. Li, Ed.
H. Shi, Ed.
Huawei Technologies
T. He
R. Pang
China Unicom
G. Qian
Huawei Technologies
14 April 2025

Distribution of Service Metadata in BGP-LS
draft-ls-idr-bgp-ls-service-metadata-05

Abstract

In edge computing, a service may be deployed on multiple instances within one or more sites, called edge service. The edge service is associated with an ANYCAST address in the IP layer, and the route of it with potential service metadata will be distributed to the network. The Edge Service Metadata can be used by ingress routers to make path selections not only based on the routing cost but also the running environment of the edge services.

The service route with metadata can be collected by a PCE(Path Compute Element) or an analyzer for calculating the best path to the best site/instance. This draft describes a mechanism to collect information of the service routes and related service metadata in BGP-LS.

About This Document

This note is to be removed before publishing as an RFC.

The latest revision of this draft can be found at <https://VMatrix1900.github.io/draft-service-metadata-in-BGP-LS/draft-ls-idr-bgp-ls-service-metadata.html>. Status information for this document may be found at <https://datatracker.ietf.org/doc/draft-ls-idr-bgp-ls-service-metadata/>.

Discussion of this document takes place on the Inter-Domain Routing Working Group mailing list (<mailto:idr@ietf.org>), which is archived at <https://mailarchive.ietf.org/arch/browse/idr/>. Subscribe at <https://www.ietf.org/mailman/listinfo/idr/>.

Source for this draft and an issue tracker can be found at <https://github.com/VMatrix1900/draft-service-metadata-in-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 16 October 2025.

Copyright Notice

Copyright (c) 2025 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<https://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Revised BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Revised BSD License.

Table of Contents

1. Introduction	3
1.1. Terminology	3
1.2. Requirements Language	3
2. BGP-LS Extension for Service in a Site	3
2.1. Prefix NLRI	4
2.2. Attributes	4
2.2.1. Metadata Path Attribute TLV	5
2.3. Prefix SID Attribute TLV	5
2.3.1. Color Attribute TLV	6
3. Security Considerations	7
4. IANA Considerations	7
5. Contributors	7
6. Normative References	7
Appendix A. Acknowledgements	8
Authors' Addresses	8

1. Introduction

Many services deploy their service instances in multiple sites to get better response time and resource utilization. These sites are often geographically distributed to serve the user demand. For some services such as VR/AR and intelligent transportation, the QoE will depend on both the network metrics and the compute metrics. For example, if the nearest site is overloaded due to the demand fluctuation, then steering the user traffic to another light-loaded site may improve the QoE.

[I-D.ietf-idr-5g-edge-service-metadata] describes the BGP extension of distributing service route with network and computing-related metrics. The router connected to the site will receive the service routes and service metadata sent from devices inside the edge site, and then generate the corresponding routes and distribute them to ingress routers. However, the route with service metadata on the router connected to the site can be also collected by a central Controller for calculating the best path to the best site.

This document defines an extension of BGP-LS to carry the service metadata along with the service route. Using the service metadata and the service route, the controller can calculate the best site for the traffic, giving each user the best QoE.

1.1. Terminology

1.2. 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.

2. BGP-LS Extension for Service in a Site

The goal of the BGP-LS extension is to collect information of the service prefix and metadata of the service, such as network metrics and compute metrics. A service is identified by a prefix, and this information is carried by the existing prefix NLRI TLV. Other information including service metadata is carried by attributes TLVs.

2.1. Prefix NLRI

A service is identified by a prefix, and the Prefix NLRI defined in the [RFC7752] is used to collect the prefix information of the service. The format of the Prefix NLRI is shown in Figure 1 for better understanding.

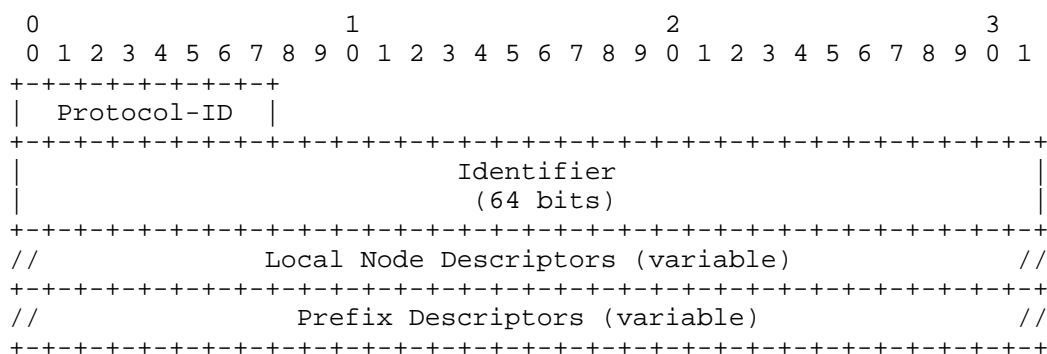


Figure 1: The IPv4/IPv6 Topology Prefix NLRI Format

Specifically, the service prefix is carried by the IP Reachability Information TLV(Figure 2) inside the Prefix Descriptor field. The Prefix Length field contains the length of the prefix in bits. The IP Prefix field contains the most significant octets of the prefix.

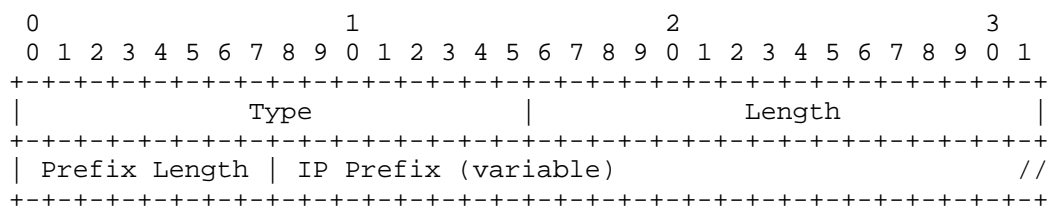


Figure 2: IP Reachability Information TLV Format

2.2. Attributes

The following three prefix attribute TLVs are used to carry the metadata of a service instance:

1. Metadata Path Attribute TLV carries the computing metric of the service instances such as site preference, site physical availability index, service delay prediction, service-oriented capability and service-oriented utilization defined in Section 4 of [I-D.ietf-idr-5g-edge-service-metadata].

2. Prefix SID TLV carries a Prefix SID associated with the edge site.
3. Color Attribute TLV carries the service requirement level information of the service

2.2.1. Metadata Path Attribute TLV

The Metadata Path Attribute TLV is an optional attribute to carry the Edge Service Metadata defined in the [I-D.ietf-idr-5g-edge-service-metadata]. It contains multiple sub-TLVs, with each sub-TLV containing a specific metric of the Edge Service Metadata. This document defines a new TLV in BGP-LS, which reuses the name and the format of Metadata Path Attribute TLV.

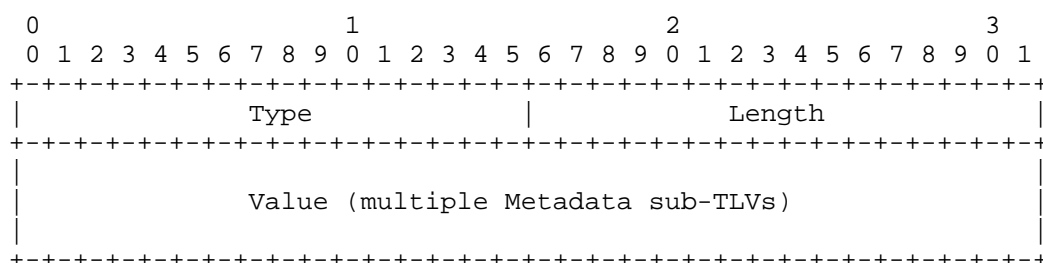


Figure 3: Metadata Path Attribute TLV format

- * Type: identify the Metadata Path Attribute, to be assigned by IANA.
- * Length: the total number of octets of the value field.
- * Value: contains multiple sub-TLVs. The name and format of sub-TLV is the same as [I-D.ietf-idr-5g-edge-service-metadata]

2.3. Prefix SID Attribute TLV

In some cases, there may be multiple sites connected to one Edge(egress) router through different interfaces. Generally, an overlay tunnel will be used between the ingress router and the egress for steering the traffic to the best site correctly. In SR-MPLS networks or SRv6 networks, a prefix SID is needed. For example, some SRv6 Endpoint Behaviors such as End.DX6, End.X can be encoded for each site so that the egress router can steer the traffic to the corresponding site. The Prefix SID TLV defined [RFC9085] can be used to collect this information.

The Prefix SID TLV is an optional TLV to carry the Prefix SID associated with the edge site. The TLV format is illustrated in Figure 4.

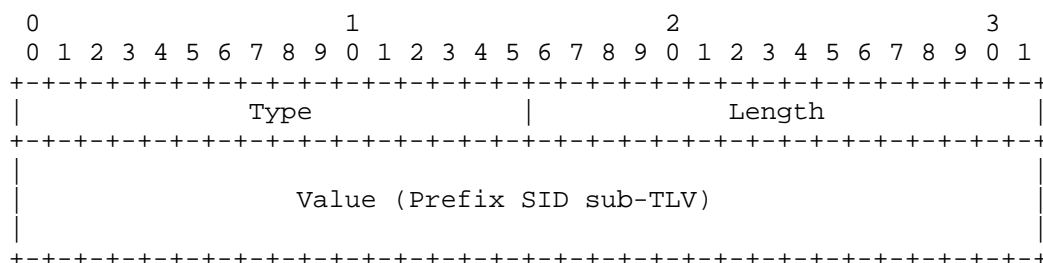


Figure 4: Prefix-SID TLV format

- * Type: 1158, identify the Prefix SID Attribute.
- * Length: the total number of octets of the value field.
- * Value: contains Prefix SID sub-TLV.

2.3.1. Color Attribute TLV

Color is used to indicate the service level. For example, different sites may have different levels of service capability which is taken into account by the controller when calculating the path to the egress router. More details can be added in the future revision.

The TLV format (shown in Figure 5) is similar to the BGP Color Extended Community defined in [RFC9012].

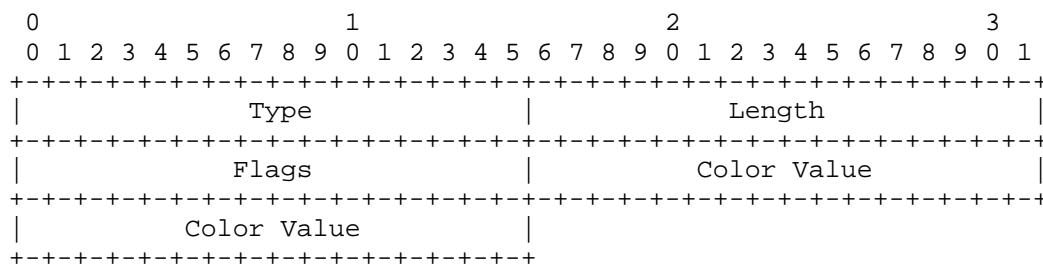


Figure 5: Color Attribute TLV format

- * Type: identify the Color Attribute, to be assigned by IANA.
- * Length: 6, length of Flags + Color Value.

- * Flags and Color are the same as defined in [RFC9012]. Color Value: 32 bit value of color.

3. Security Considerations

TBD

4. IANA Considerations

This document requires IANA to assign the following code points from the registry called "BGP-LS Node Descriptor, Link Descriptor, Prefix Descriptor, and Attribute TLVs":

Value	Description	Reference
TBD1	Metadata Path Attribute Type	Section 2.2.1
TBD2	Color Attribute Type	Section 2.3.1
TBD3	Site Preference Sub-Type	Section 2.2.1
TBD4	Site Physical Availability Index Sub-Type	Section 2.2.1
TBD5	Service Delay Prediction Index Sub-Type	Section 2.2.1
TBD6	Service-Oriented Capability Sub-Type	Section 2.2.1
TBD7	Service-Oriented Utilization Sub-Type	Section 2.2.1

Table 1

5. Contributors

Xiangfeng Ding

email: dingxiangfeng@huawei.com

6. Normative References

- [I-D.ietf-idr-5g-edge-service-metadata]
Dunbar, L., Majumdar, K., Li, C., Mishra, G. S., and Z. Du, "BGP Extension for 5G Edge Service Metadata", Work in

Progress, Internet-Draft, draft-ietf-idr-5g-edge-service-metadata-26, 15 January 2025,
<<https://datatracker.ietf.org/doc/html/draft-ietf-idr-5g-edge-service-metadata-26>>.

- [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/rfc/rfc2119>>.
- [RFC7752] Gredler, H., Ed., Medved, J., Previdi, S., Farrel, A., and S. Ray, "North-Bound Distribution of Link-State and Traffic Engineering (TE) Information Using BGP", RFC 7752, DOI 10.17487/RFC7752, March 2016, <<https://www.rfc-editor.org/rfc/rfc7752>>.
- [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/rfc/rfc8174>>.
- [RFC9012] Patel, K., Van de Velde, G., Sangli, S., and J. Scudder, "The BGP Tunnel Encapsulation Attribute", RFC 9012, DOI 10.17487/RFC9012, April 2021, <<https://www.rfc-editor.org/rfc/rfc9012>>.
- [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/rfc/rfc9085>>.
- [RFC9252] Dawra, G., Ed., Talaulikar, K., Ed., Raszuk, R., Decraene, B., Zhuang, S., and J. Rabadan, "BGP Overlay Services Based on Segment Routing over IPv6 (SRv6)", RFC 9252, DOI 10.17487/RFC9252, July 2022, <<https://www.rfc-editor.org/rfc/rfc9252>>.

Appendix A. Acknowledgements

The authors would like to thank Haibo Wang, Lili Wang, Jianwei Mao for their help.

Authors' Addresses

Cheng Li (editor)
Huawei Technologies
Beijing
China

Email: c.l@huawei.com

Hang Shi (editor)
Huawei Technologies
Beijing
China
Email: shihang9@huawei.com

Tao He
China Unicom
Beijing
China
Email: het21@chinaunicom.cn

Ran Pang
China Unicom
Beijing
China
Email: pangran@chinaunicom.cn

Guofeng Qian
Huawei Technologies
Beijing
China
Email: qianguofeng@huawei.com