Workgroup: Inter-Domain Routing
Internet-Draft:
draft-ls-idr-bgp-ls-service-metadata-01
Published: 24 February 2023
Intended Status: Standards Track
Expires: 28 August 2023
Authors: C. Li, Ed. H. Shi, Ed.
Huawei Technologies Huawei Technologies
T. He R. Pang G. Qian
China Unicom China Unicom Huawei Technologies

Distribution of Service Metadata in BGP-LS

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 IP layer, and the route of it with potential service metatdata 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 the 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

Copyright Notice

Table of Contents

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 steer the user traffic to a another light-loaded sites may improve the QoE.

[I-D.ietf-idr-5g-edge-service-metadata] descirbes the BGP extension of distributing service route with network and computing-related metrics. The router connected to the site will received the service routes and service metadata sent from devices inside the egdge site, and then generates the corresponding routes and distributes 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 the information of the service prefix and metadata of the service, such as network metrics and compute metrics. A service is identified by an prefix, and this information is carried by existing prefix NLRI TLV. Other information including service metadata are 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.

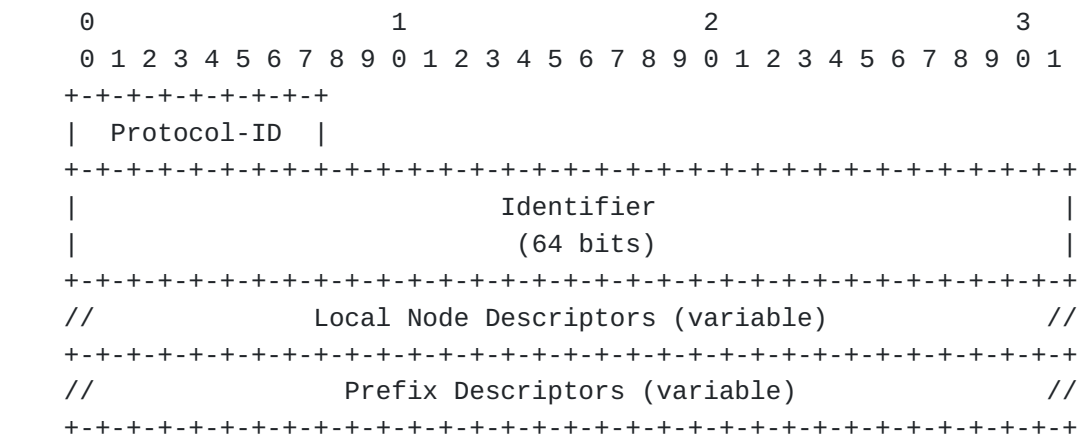

Figure 1: The IPv4/IPv6 Topology Prefix NLRI Format

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.

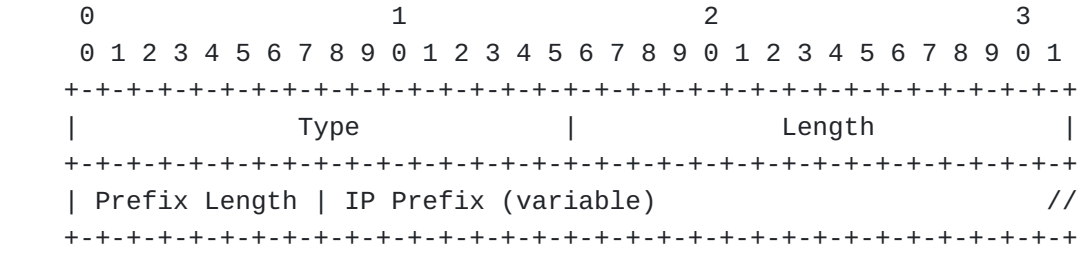

Figure 2: IP Reachability Information TLV Format

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 compute metric of the service instance such as site preference, capacity index and load measurement defined in [I-D.ietf-idr-5g-edge-service-metadata].

2. Prefix SID TLV carries a Prefix SID associated to 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 define a new TLV in BGP-LS, which reuse the name and the format of Metadata Path Attribute TLV.

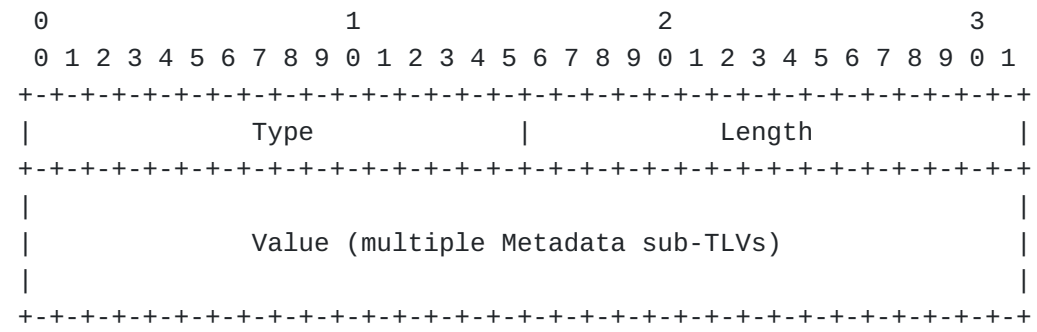

Figure 3: Metadata Path Attribute TLV format

*Type: identify the Metadata Path Attribute, to be assigned by IANA.

*Length: the total number of the octets of the value field.

*Value: contains multiple sub-TLVs.

There are three types of Edge Service Metadata sub-TLVs defined in [I-D.ietf-idr-5g-edge-service-metadata]:

1. Site Preference Index indicates the preference to choose the site.

2. Capacity Index indicates the capability of a site. One Edge Site can be in full capacity, reduced capacity, or completely out of service.

3. Load Measurement indicates the load level of the site.

To collect these information, this document defines TLVs reusing the name and format of the TLVs defined in [I-D.ietf-idr-5g-edge-service-metadata].

2.3. Prefix SID Attribute TLV

In some cases, there may be multiple sites connect to one Edge(egress) router through different interfaces. Generally, a overlay path, such a 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 to the edge site. The TLV format is illustrated in Figure 4.

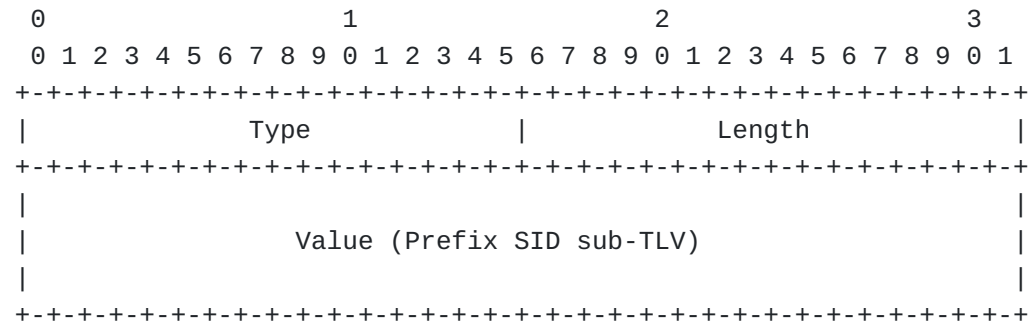

Figure 4: Prefix-SID TLV format

*Type: 1158, identify the Prefix SID Attribute.

*Length: the total number of the 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 site may have different level of service capability which is taken into account of by the controller when calculate 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].

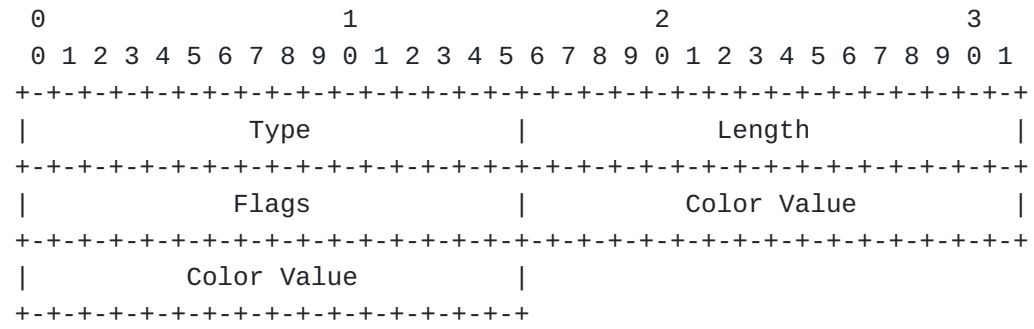

Figure 5: Color Attribute TLV format

*Type: identify the Color Attribute, to be assigned by IANA.

*Length: 6, length of Flags + Color Value.

*Flags and Color is 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	Site Preference Sub-Type	Section 2.2.1
TBD3	Capacity Sub-Type	Section 2.2.1
TBD4	Load Measurement Sub-Type1: Aggregated-Cost	Section 2.2.1
TBD5	Load Measurement Sub-Type2: Raw-Measurements	Section 2.2.1
TBD6	Color Attribute Type	Section 2.3.1

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., Wang, H., and G. S. Mishra, "BGP Extension for 5G Edge Service Metadata", Work in Progress, Internet-Draft, draft-ietf-idr-5g-edge-service-metadata-00, 2 December 2022, <https://datatracker.ietf.org/doc/html/draft-ietf-idr-5g-edge-service-metadata-00>.

[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>.

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