Workgroup: IDR Working Group
Internet-Draft:
draft-dong-idr-bgpls-sr-enhanced-vpn-05
Published: 23 October 2023
Intended Status: Standards Track
Expires: 25 April 2024
Authors: J. Dong, Huawei Technologies
Z. Hu, Huawei Technologies
Z. Li, Huawei Technologies
X. Tang, China Unicom
R. Pang, China Unicom

BGP-LS Extensions for Scalable Segment Routing based Enhanced VPN

Abstract

Enhanced VPN (VPN+) aims to provide enhanced VPN services to support some applications' needs of enhanced isolation and stringent performance requirements. VPN+ requires integration between the overlay VPN connectivity and the resources and characteristics provided by the underlay network. A Virtual Transport Network (VTN) is a virtual underlay network which can be used to support one or a group of VPN+ services. In the context of network slicing, a VTN could be instantiated as a network resource partition (NRP).

This document specifies the BGP-LS mechanisms with necessary extensions to advertise the information of scalable Segment Routing (SR) based NRPs to a centralized network controller. Each NRP can have a customized topology and a set of network resources allocated from the physical network. Multiple NRPs may shared the same topology, and multiple NRPs may share the same set of network resources on specific network segments. This allows flexible combination of network topology and network resource attributes to build a large number of NRPs with a relatively small number of logical topologies. The proposed mechanism is applicable to both segment routing with MPLS data plane (SR-MPLS) and segment routing with IPv6 data plane (SRv6).

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 RFC 2119 [RFC2119].

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

Enhanced VPN (VPN+) is an enhancement to VPN services to support the needs of new applications, particularly the applications that are associated with 5G services. These applications require enhanced isolation and have more stringent performance requirements than that can be provided with traditional overlay VPNs. These properties require integration between the underlay and the overlay networks. [I-D.ietf-teas-enhanced-vpn] specifies the framework of enhanced VPN and describes the candidate component technologies in different network planes and layers. An enhanced VPN can be used for 5G network slicing, and will also be of use in more generic scenarios.

To meet the requirement of enhanced VPN services, a number of virtual underlay networks need to be created, each with a subset of the underlay network topology and a set of network resources allocated to meet the requirement of a specific VPN+ service or a group of VPN+ services. Such a virtual underlay network is called Virtual Transport Network (VTN) in [I-D.ietf-teas-enhanced-vpn]. [I-D.ietf-teas-ietf-network-slices] introduces the concept Network Resource Partition (NRP) as a set of network resources that are available to carry traffic and meet the SLOs and SLEs. In order to allocate network resources to an NRP, the NRP is associated with a network topology to define the set of links and nodes. Thus VTN and NRP are similar concepts, and NRP can be seen as an instantiation of VTN in the context of network slicing. For clarity, the rest of this document uses NRP in the description of the proposed mechanisms and protocol extensions.

[I-D.ietf-spring-resource-aware-segments] introduces resource-awareness to Segment Routing (SR) [RFC8402] by associating existing type of SIDs with network resource attributes (e.g. bandwidth, processing or storage resources). These resource-aware SIDs retain their original functionality, with the additional semantics of identifying the set of network resources available for the packet processing action. [I-D.ietf-spring-sr-for-enhanced-vpn] describes the use of resource-aware segments to build SR based NRPs. To allow the network controller and network nodes to perform NRP-specific explicit path computation and/or shortest path computation, the group of resource-aware SIDs allocated by network nodes to each NRP and the associated topology and resource attributes need to be distributed in the control plane.

When an NRP spans multiple IGP areas or multiple Autonomous Systems (ASes), BGP-LS is needed to advertise the NRP information in each IGP area or AS to the network controller, so that the controller could use the collected information to build the view of inter-area or inter-AS SR NRPs.

This document describes BGP-LS [RFC7752] based mechanism with necessary extensions to advertise the topology and resource attribute of inter-area and inter-domain SR based NRPs. Each NRP can have a customized topology and a set of network resources allocated. Multiple NRPs may shared the same topology, and some of the NRPs may share the same set of network resources on specific network segments. This allows flexible combination of network topology and network resource attributes to build a large number of NRPs with a relatively small number of logical topologies. The definition of NRP is advertised as a node attribute using BGP-LS. The associated network topology and resources attributes of a NRP are advertised as link attributes using BGP-LS.

2. Advertisement of NRP Definition

According to [I-D.ietf-teas-ietf-network-slices], an NRP consists of a set of dedicated or shared network resources, and is associated with a customized network topology. Thus a NRP can be defined as the combination of a set of network attributes, which include the topology attribute and other attributes, such as the associated network resources.

The Network Resource Partition Definition (NRPD) TLV is a new TLV of the optional BGP-LS Attribute which is associated with the node NLRI.

The format of NRPD TLV is as follows:

```
 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            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                            NRP ID                             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|             MT-ID             |   Algorithm   |     Flags     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                           Sub-TLVs                            |
~                              ...                              ~
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Where:

* Type: To be assigned by IANA.

* Length: the length of the value field of the TLV. It is variable dependent on the included Sub-TLVs.

* NRP ID: A global significant 32-bit identifier which is used to identify an NRP.

* MT-ID: 16-bit identifier which contains the multi-topology identifier of the IGP topology.

* Algorithm: 8-bit identifier which indicates the algorithm which applies to this virtual transport network. It can be either a normal algorithm in [RFC8402] or a Flex-Algorithm [RFC9350].

* Flags: 8-bit flags. Currently all the flags are reserved for future use. They SHOULD be set to zero on transmission and MUST be ignored on receipt.

* Sub-TLVs: optional sub-TLVs to specify the additional attributes of an NRP. Currently no sub-TLV is defined in this document.

3. Advertisement of NRP Topology Attribute

[I-D.dong-lsr-sr-enhanced-vpn] describes the IGP mechanisms to distribute the topology attributes of SR based NRPs. This section describes the BGP-LS mechanism to distribute both the intra-domain and inter-domain topology attributes of SR based NRPs.

3.1. Intra-domain Topology Advertisement

The intra-domain topology attribute of an NRP can be determined by the MT-ID and/or the algorithm ID included in the NRP definition. In practice, it could be described using two optional approaches.

The first approach is to use Multi-Topology Routing (MTR) [RFC4915] [RFC5120] with the segment routing extensions to advertise the topology associated with the SR based NRPs. Different algorithms MAY be used to further specify the computation algorithm or the metric type used for path computation within the topology. Multiple NRPs can be associated with the same <topology, algorithm> tupple, and the IGP computation with the <topology, algorithm> tuple can be shared by these NRPs.

The second approach is to use Flex-Algo [RFC9350] to describe the topological constraints of SR based NRPs on a network topology (e.g. the default topology). Multiple NRPs can be associated with the same

Flex-Algo, and the IGP computation result with this Flex-Algo can be shared.

This section describes the two optional approaches to advertise the intra-domain topology of an NRP using BGP-LS.

3.1.1. MTR based Topology Advertisement

In section 4.2.2.1 of [I-D.ietf-idr-rfc7752bis], Multi-Topology Identifier (MT-ID) TLV is defined, which can contain one or more IS-IS or OSPF Multi-Topology IDs. The MT-ID TLV MAY be present in a Link Descriptor, a Prefix Descriptor, or the BGP-LS Attribute of a Node NLRI.

[RFC9085] defines the BGP-LS extensions to carry the segment routing information using TLVs of BGP-LS Attribute. When MTR is used with SR-MPLS data plane, topology-specific prefix-SIDs and topology-specific Adj-SIDs can be carried in the BGP-LS Attribute associated with the prefix NLRI and link NLRI respectively, the MT-ID TLV is carried in the prefix descriptor or link descriptor to identify the corresponding topology of the SIDs.

[I-D.ietf-idr-bgpls-srv6-ext] defines the BGP-LS extensions to advertise SRv6 segments along with their functions and attributes. When MTR is used with SRv6 data plane, the SRv6 Locator TLV is carried in the BGP-LS Attribute associated with the prefix-NLRI, the MT-ID TLV can be carried in the prefix descriptor to identify the corresponding topology of the SRv6 Locator. The SRv6 End.X SIDs are carried in the BGP-LS Attribute associated with the link NLRI, the MT-ID TLV can be carried in the link descriptor to identify the corresponding topology of the End.X SIDs. The SRv6 SID NLRI is defined to advertise other types of SRv6 SIDs, in which the SRv6 SID Descriptors can include the MT-ID TLV so as to advertise topology-specific SRv6 SIDs.

[I-D.ietf-idr-rfc7752bis] also defines the rules of the usage of MT-ID TLV:

"In a Link or Prefix Descriptor, only a single MT-ID TLV containing the MT-ID of the topology where the link or the prefix is reachable is allowed. In case one wants to advertise multiple topologies for a given Link Descriptor or Prefix Descriptor, multiple NLRIs MUST be generated where each NLRI contains a single unique MT-ID."

Editor's note: the above rules indicates that only one MT-ID is allowed to be carried the Link or Prefix descriptors. When a link or prefix needs to be advertised in multiple topologies, multiple NLRIs needs to be generated to report all the topologies the link or prefix participates in, together with the topology-specific segment routing information and link attributes. This may increase the number of BGP Updates needed for advertising MT-specific topology attributes, and may introduce additional processing burden to both the sending BGP speaker and the receiving network controller. When the number of topologies in a network is not a small number, some optimization may be needed for the reporting of multi-topology information and the associated segment routing information in BGP-LS. Based on the WG's opinion, this will be elaborated in a future version.

3.1.2. Flex-Algo based Topology Advertisement

The Flex-Algo definition [RFC9350] can be used to describe the calculation-type, the metric-type and the topological constraints for path computation on a network topology. As specified in [I-D.dong-lsr-sr-enhanced-vpn], the topology of a NRP can be determined by applying Flex-Algo constraints on a network topology.

BGP-LS extensions for Flex-Algo [RFC9351] provide the mechanisms to advertise the Flex-Algo definition information. BGP-LS extensions for SR-MPLS [RFC9085] and SRv6 [I-D.ietf-idr-bgpls-srv6-ext] provide the mechanism to advertise the algorithm-specific segment routing information.

In [RFC9085], algorithm-specific prefix-SIDs can be advertised in BGP-LS attribute associated with Prefix NLRI. In [I-D.ietf-idr-bgpls-srv6-ext], algorithm-specific SRv6 Locators can be advertised in BGP-LS Attribute associated with the corresponding Prefix NLRI, and algorithm-specific End.X SID can be advertised in BGP-LS Attribute associated with the corresponding Link NLRI. Other types of SRv6 SIDs can also be algorithm-specific and are advertised using the SRv6 SID NLRI.

3.2. Inter-Domain Topology Advertisement

In some network scenarios, an NRP which spans multiple areas or ASes needs to be created. The multi-domain NRP could have different inter-domain connectivity, and may be associated with different set of network resources in each domain and also on the inter-domain links. In order to build the multi-domain NRPs using segment routing, it is necessary to advertise the topology and resource attribute of NRP on the inter-domain links and the associated BGP Peering SIDs.

[RFC9086] and [I-D.ietf-idr-bgpls-srv6-ext] defines the BGP-LS extensions for advertisement of BGP topology information between ASes and the associated BGP Peering Segment Identifiers. Such information could be used by a network controller for the computation and instantiation of inter-AS traffic engineering SR paths.

Depending on the network scenarios and the requirement of inter-domain NRPs, different mechanisms can be used to specify the inter-domain connections of NRPs.

* One EBGP session between two ASes can be established over multiple underlying links. In this case, different underlying links can be used for different inter-domain NRPs which requires link isolation between each other. In another similar case, the EBGP session is established over a single link, while the network resource (e.g. bandwidth) on this link can be partitioned into different pieces, each of which can be considered as a virtual member link. In both cases, different BGP Peer-Adj-SIDs SHOULD be allocated to each underlying physical or virtual member link, and ASBRs SHOULD advertise the NRP identifier associated with each BGP Peer-Adj-SID.

* For inter-domain connection between two ASes, multiple EBGP sessions can be established between different set of peering ASBRs. It is possible that some of these BGP sessions are used for one inter-domain NRP, while some other BGP sessions are used for another inter-domain NRP. In this case, different BGP peer-node-SIDs are allocated to each BGP session, and ASBRs SHOULD advertise the NRP identifier associated with each BGP Peer-node-SIDs.

* At the AS-level topology, different inter-domain NRPs may have different inter-domain connectivity. Different BGP Peer-Set-SIDs can be allocated to represent the groups of BGP peers which can be used for load-balancing in each inter-domain NRP.

In network scenarios where the MT-ID or Flex-Algo is used consistently in multiple areas or ASes covered by a NRP. the approaches to advertise topology-specific BGP peering SIDs are described as below:

* Using MT-based mechanism, the topology-specific BGP peering SIDs can be advertised with the MT-ID associated with the NRP carried in the corresponding link NLRI. This can be supported with the existing mechanisms defined in [RFC7752][RFC9086] and [I-D.ietf-idr-bgpls-srv6-ext].

* Using Flex-Algo based mechanism, the topology-specific BGP peering SIDs can be advertised together with the Admin Group (color) of the corresponding Flex-Algo in the BGP-LS attribute.

In network scenarios where consistent usage of MT-ID or Flex-Algo among multiple ASes can not be expected, then the global-significant NRP-ID can be used to define the AS level topologies. Within each domain, the MT or Flex-Algo based mechanism could still be used for topology advertisement.

3.2.1. NRP IDs TLV

A new NRP IDs TLV is defined to describe the identifiers of one or more NRPs an intra-domain or inter-domain link belongs to. It can be carried in BGP-LS attribute which is associated with a Link NLRI, or it could be carried as a sub-TLV in the L2 Bundle Member Attribute TLV.

The format of NRP IDs TLV is as below:

```
 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            |            Reserved           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                            NRP ID-1                           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                              ...                              ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                            NRP ID-n                           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Where:

* Type: To be assigned by IANA.

* Length: The length of the value field of the sub-TLV. It is variable dependent on the number of NRP IDs included.

* Flags: 16 bit flags. All the bits are reserved, which MUST be set to 0 on transmission and SHOULD be ignored on receipt.

* Reserved: this field is reserved for future use. MUST be set to 0 on transmission and SHOULD be ignored on receipt.

* NRP IDs: One or more 32-bit identifiers to specify the NRPs this link belongs to.

4. Advertisement of NRP Resource Attribute

[I-D.dong-lsr-sr-enhanced-vpn] specifies the optional mechanism to advertise the resource information associated with each NRP. One approach is to use the L2 bundle mechanism [RFC8668] to advertise the set of link resources allocated to an NRP as a L2 physical or virtual member link. Another approach is to advertise the set of network resources as per NRP link TE attributes. This section defines the corresponding BGP-LS extensions for both approaches.

Two new TLVs are defined to carry the NRP ID and the link attribute flags of either a Layer-3 link or the L2 bundle member links. The NRP ID TLV is defined in section 3.2.1 of this document, and a new Link Attribute Flags TLV is defined in this section. The TE attributes of each Layer 3 link or the L2 bundle member link, such as the bandwidth and the SR SIDs, can be advertised using the mechanism as defined in [RFC9085][RFC9086] and [I-D.ietf-idr-bgpls-srv6-ext].

4.1. Option 1: L2 Bundle based Approach

On an Layer-3 interface, each NRP can be allocated with a subset of link resources (e.g. bandwidth). A subset of link resources may be dedicated to an NRP, or may be shared by a group of NRPs. Each subset of link resource can be instantiated as a virtual layer-2 member link under the Layer-3 interface, and the Layer-3 interface is considered as a virtual Layer-2 bundle. The Layer-3 interface may also be a physical Layer 2 link bundle, in this case a subset of link resources allocated to an NRP may be provided by one of the physical Layer-2 member links.

The NRP ID TLV defined in section 3.2.1 of this document is used to carry the NRP IDs associated with the L2 bundle member links. The TE attributes of the L2 bundle member links, such as the maximum link bandwidth, and the SR SIDs, can be advertised using the mechanism as defined in [RFC9085][RFC9086] and [I-D.ietf-idr-bgpls-srv6-ext].

A new Link attribute Flags TLV is defined to specify the characteristics of a link. It can be carried in BGP-LS attribute which is associated with a Link NLRI, or it could be carried as a sub-TLV in the L2 Bundle Member Attribute TLV. The format of the sub-TLV is as below:

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

Where:

Type: TBD

Length: 4 octets.

Flags: 16-bit flags. This field is consistent with the Flag field in IS-IS Link Attribute sub-TLV in [RFC5029]. In addition to the flags defined in [RFC5029], A new Flag "E" is defined in this document.

-Link excluded from load balancing. When the flag is set, it indicates this link is only used for the associated NRPs.

.

4.2. Option 2: Per-NRP Link TE Attributes

An Layer-3 interface can participate in multiple NRPs, each of which is allocated with a subset of the resources of the interface. For each NRP, the associated resources can be described using per-NRP TE attributes. A new NRP-specific TE attribute TLV is defined to advertise the link attributes associated with an NRP. This sub-TLV MAY be carried in the BGP-LS Attribute associated with a Link NLRI. The format of the NRP-specific TE attribute TLV is shown as below:

```
 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             |            Reserved           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                         NRP IDs Sub-TLV                       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                         Other Sub-TLVs                        ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Where:

*Type: To be assigned by IANA.

*Length: The length of the value field of the TLV. It is variable dependent on the length of the Sub-TLVs field.

*Flags: 16-bit flags. All the 16 bits are reserved for future use, which SHOULD be set to 0 on transmission and MUST be ignored on receipt.

*Reserved: 16-bit field reserved for future use, SHOULD be set to 0 on transmission and MUST be ignored on receipt.

The NRP IDs TLV as defined in section 3.2.1 is used as the NRP IDs Sub-TLV in the per-NRP Link TE Attribute TLV.

Other Sub-TLVs are optional and can be used to carry the TE attributes associated with the NRPs. The existing Link TE Attribute TLVs as defined in [I-D.ietf-idr-rfc7752bis] can be reused as sub-TLVs here. New sub-TLVs may be defined in the future.

5. Advertisement of NRP specific Data Plane Identifiers

In network scenarios where each NRP is associated with an independent topology or Flex-Algo, the topology or Flex-Algo specific SR SIDs or Locators could be used to identify the NRP in data plane, so that the set of network resources associated with the NRP can be determined. In network scenarios where multiple NRPs share the same topology or Flex-Algo, additional data plane identifiers are needed to identify different NRPs.

This section describes the mechanisms to advertise the NRP identifiers with different data plane encapsulations.

5.1. NRP-specific SR-MPLS SIDs

With SR-MPLS data plane, the NRP identifier can be implicitly determined by the SR SIDs associated with the NRP. Each node SHOULD allocate NRP-specific Prefix-SIDs for each NRP it participates in. Similarly, NRP-specific Adj-SIDs MAY be allocated for each link which participates in the NRP.

5.1.1. NRP-specific Prefix-SID TLV

A new NRP-specific Prefix-SID TLV is defined to advertise the relationship between the prefix-SID and its associated NRP. It is derived from NRP-specific Prefix-SID sub-TLV of IS-IS [I-D.dong-lsr-sr-enhanced-vpn]. The format of the sub-TLV is as below:

```
 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             |            Reserved           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                             NRP ID                            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                    SID/Index/Label(Variable)                  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Where:

*Type: TBD

*Length: The length of the value field of the sub-TLV. It is variable dependent on the length of the SID/Index/Label field.

*Flags: 16-bit flags. The high-order 8 bits are the same as in the Prefix-SID sub-TLV defined in [RFC8667]. The lower-order 8 bits are reserved for future use, which SHOULD be set to 0 on transmission and MUST be ignored on receipt.

*Reserved: 16-bit field reserved for future use, SHOULD be set to 0 on transmission and MUST be ignored on receipt.

*NRP ID: A 32-bit local identifier to identify the NRP this prefix-SID is associated with.

*SID/Index/Label: The same as defined in [RFC8667].

One or more of NRP-specific Prefix-SID TLVs MAY be carried in BGP-LS attribute of the associated Prefix NLRI. The MT-ID in the Prefix descriptors SHOULD be the same as the MT-ID in the definition of the NRP.

5.1.2. NRP-specific Adj-SID TLV

A new NRP-specific Adj-SID TLV is defined to advertise between the Adj-SID and its associated NRP. It is derived from NRP specific Adj-SID sub-TLV of IS-IS [I-D.dong-lsr-sr-enhanced-vpn]. The format of the sub-TLV is as below:

```
 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             |            Reserved           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                             NRP ID                            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      SID/Index/Label(Variable)                |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Where:

*Type: TBD

*Length: The length of the value field of the sub-TLV. It is variable dependent on the length of the SID/Index/Label field.

*Flags: 16-bit flags. The high-order 8 bits are the same as in the Adj-SID sub-TLV defined in [RFC8667]. The lower-order 8 bits are reserved for future use, which SHOULD be set to 0 on transmission and MUST be ignored on receipt.

* Reserved: 16-bit field reserved for future use, SHOULD be set to 0 on transmission and MUST be ignored on receipt.

* NRP ID: A 32-bit global unique identifier to identify the NRP this Adj-SID is associated with.

* SID/Index/Label: The same as defined in [RFC8667].

Multiple NRP-specific Adj-SID TLVs MAY be carried in BGP-LS attribute of the associated Link NLRI. The MT-ID in the Link descriptors SHOULD be the same as the MT-ID in the definition of these NRPs.

5.2. NRP-specific SRv6 SIDs

5.2.1. NRP-specific SRv6 Locators and End SIDs

With SRv6 data plane, the NRP identifier can be implicitly or explicitly determined using the SRv6 Locators associated with the NRP, this is to ensure that all network nodes (including both the SRv6 End nodes and Transit nodes) can identify the NRP to which a packet belongs. Network nodes SHOULD allocate NRP-specific Locators for each NRP it participates in. The NRP-specific Locators are used as the covering prefix of NRP-specific SRv6 End SIDs, End.X SIDs and other types of SIDs.

Each NRP-specific SRv6 Locator MAY be advertised in a separate Prefix NLRI. If multiple NRPs share the same topology/algorithm, the topology/algorithm specific Locator is the covering prefix of a group of NRP-specific Locators. Then the advertisement of NRP-specific locators can be optimized to reduce the amount of information advertised in the control plane.

A new NRP locator-block sub-TLV under the SRv6 Locator TLV is defined to advertise a set of sub-blocks which follows the topology/algorithm specific Locator. Each NRP locator-block value is assigned to one of the NRPs which share the same topology/algorithm.

```
     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             |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    | Number of NRPs|  Block Length |           Reserved            |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                           NRP ID #1                           |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    ~                      Locator Block Value                      ~
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    ~                              ...                              ~
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                           NRP ID #n                           |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    ~                      Locator Block Value                      ~
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Where:

* Type: TBD

* Length: The length of the value field of the sub-TLV. It is variable dependent on the number of NRPs and the Block Length.

* Number of NRPs: The number of NRPs which share the same topology/algorithm specific Locator as the covering prefix.

* Block Length: The length of the NRP locator-block which follows the length of the topology/algorithm specific Locator.

* NRP ID: A 32-bit identifier to identify the NRP the locator-block is associated with.

* Block Value: The value of the NRP locator-block for each NRP.

With the NRP locator-block sub-TLV, the NRP-specific Locator can be obtained by concatenating the topology/algorithm specific locator and the locator-block value advertised for the NRP.

5.2.2. NRP-specific SRv6 End.X SID

The SRv6 End.X SIDs are advertised in the BGP-LS attribute with Link NLRI.In order to distinguish the End.X SIDs which belong to different NRPs, a new "NRP ID Sub-TLV" is introduced under the SRv6 End.X SID TLV and SRv6 LAN End.X SID TLV defined in [I-D.ietf-idr-bgpls-srv6-ext]. Its format is shown as below:

```
 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             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                            NRP ID                             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Where:

- *Type: TBD.
- *Length: the length of the Value field of the TLV. It is set to 4.
- *NRP ID: A 32-bit global identifier to identify the NRP this End.X SID is associated with.

5.3. Dedicated NRP ID in Data Plane

As the number of NRPs increases, with the mechanism described in [I-D.ietf-spring-sr-for-enhanced-vpn], the number of SR SIDs and SRv6 Locators allocated for different NRPs would also increase. In network scenarios where the number of SIDs or Locators becomes a concern, some data plane optimization may be needed to reduce the amount of SR SIDs and Locators allocated. As described in [I-D.ietf-teas-nrp-scalability], one approach is to decouple the data plane identifiers used for topology based forwarding and the identifiers used for the NRP-specific processing. Thus a new data plane global NRP-ID could be introduced and encapsulated in the packet. One possible encapsulation of NRP-ID in IPv6 data plane is proposed in [I-D.ietf-6man-enhanced-vpn-vtn-id]. One possible encapsulation of NRP-ID in MPLS data plane is proposed in [I-D.li-mpls-enhanced-vpn-vtn-id].

In that case, the NRP ID encapsulated in data packet can be the same value as the NRP ID used in the control protocols, so that the overhead of advertising the mapping relationship between the NRP IDs in the control plane and the corresponding data plane identifiers could be saved.

6. Security Considerations

This document introduces no additional security vulnerabilities to BGP-LS.

The mechanism proposed in this document is subject to the same vulnerabilities as any other protocol that relies on BGP-LS.

7. IANA Considerations

TBD

8. Acknowledgments

The authors would like to thank Shunwan Zhuang and Zhenbin Li for the review and discussion of this document.

9. References

9.1. Normative References

[I-D.ietf-idr-bgpls-srv6-ext]
Dawra, G., Filsfils, C., Talaulikar, K., Chen, M., Bernier, D., and B. Decraene, "BGP Link State Extensions for SRv6", Work in Progress, Internet-Draft, draft-ietf-idr-bgpls-srv6-ext-14, 17 February 2023, <https://datatracker.ietf.org/doc/html/draft-ietf-idr-bgpls-srv6-ext-14>.

[I-D.ietf-idr-rfc7752bis]
Talaulikar, K., "Distribution of Link-State and Traffic Engineering Information Using BGP", Work in Progress, Internet-Draft, draft-ietf-idr-rfc7752bis-17, 25 August 2023, <https://datatracker.ietf.org/doc/html/draft-ietf-idr-rfc7752bis-17>.

[I-D.ietf-spring-resource-aware-segments]
Dong, J., Bryant, S., Miyasaka, T., Zhu, Y., Qin, F., Li, Z., and F. Clad, "Introducing Resource Awareness to SR Segments", Work in Progress, Internet-Draft, draft-ietf-spring-resource-aware-segments-07, 31 May 2023, <https://datatracker.ietf.org/doc/html/draft-ietf-spring-resource-aware-segments-07>.

[I-D.ietf-spring-sr-for-enhanced-vpn]
Dong, J., Bryant, S., Miyasaka, T., Zhu, Y., Qin, F., Li, Z., and F. Clad, "Segment Routing based Virtual Transport Network (VTN) for Enhanced VPN", Work in Progress, Internet-Draft, draft-ietf-spring-sr-for-enhanced-vpn-05, 31 May 2023, <https://datatracker.ietf.org/doc/html/draft-ietf-spring-sr-for-enhanced-vpn-05>.

[I-D.ietf-teas-enhanced-vpn] Dong, J., Bryant, S., Li, Z., Miyasaka, T., and Y. Lee, "A Framework for Enhanced Virtual Private Network (VPN+)", Work in Progress, Internet-Draft, draft-ietf-teas-enhanced-vpn-14, 28 July 2023, <https://datatracker.ietf.org/doc/html/draft-ietf-teas-enhanced-vpn-14>.

[I-D.ietf-teas-ietf-network-slices]
Farrel, A., Drake, J., Rokui, R., Homma, S., Makhijani, K., Contreras, L. M., and J. Tantsura, "A Framework for Network Slices in Networks Built from IETF Technologies", Work in Progress, Internet-Draft, draft-ietf-teas-ietf-network-slices-25, 14 September 2023, <https://datatracker.ietf.org/doc/html/draft-ietf-teas-ietf-network-slices-25>.

[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>.

[RFC5029] Vasseur, JP. and S. Previdi, "Definition of an IS-IS Link Attribute Sub-TLV", RFC 5029, DOI 10.17487/RFC5029, September 2007, <https://www.rfc-editor.org/info/rfc5029>.

[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/info/rfc7752>.

[RFC8402] Filsfils, C., Ed., Previdi, S., Ed., Ginsberg, L., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing Architecture", RFC 8402, DOI 10.17487/RFC8402, July 2018, <https://www.rfc-editor.org/info/rfc8402>.

[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] Previdi, S., Talaulikar, K., Ed., Filsfils, C., Patel, K., Ray, S., and 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>.

[RFC9351] Talaulikar, K., Ed., Psenak, P., Zandi, S., and G. Dawra, "Border Gateway Protocol - Link State (BGP-LS) Extensions for Flexible Algorithm Advertisement", RFC 9351, DOI 10.17487/RFC9351, February 2023, <https://www.rfc-editor.org/info/rfc9351>.

9.2. Informative References

[I-D.dong-lsr-sr-enhanced-vpn] Dong, J., Hu, Z., Li, Z., Tang, X., Pang, R., and S. Bryant, "IGP Extensions for Scalable Segment Routing based Virtual Transport Network (VTN)", Work in Progress, Internet-Draft, draft-dong-lsr-sr-enhanced-vpn-09, 10 July 2023, <https://datatracker.ietf.org/doc/html/draft-dong-lsr-sr-enhanced-vpn-09>.

[I-D.ietf-6man-enhanced-vpn-vtn-id] Dong, J., Li, Z., Xie, C., Ma, C., and G. S. Mishra, "Carrying Virtual Transport Network (VTN) Information in IPv6 Extension Header", Work in Progress, Internet-Draft, draft-ietf-6man-enhanced-vpn-vtn-id-05, 6 July 2023, <https://datatracker.ietf.org/doc/html/draft-ietf-6man-enhanced-vpn-vtn-id-05>.

[I-D.ietf-teas-nrp-scalability] Dong, J., Li, Z., Gong, L., Yang, G., Guichard, J., Mishra, G. S., Qin, F., Saad, T., and V. P. Beeram, "Scalability Considerations for Network Resource Partition", Work in Progress, Internet-Draft, draft-ietf-teas-nrp-scalability-02, 2 June 2023, <https://datatracker.ietf.org/doc/html/draft-ietf-teas-nrp-scalability-02>.

[I-D.li-mpls-enhanced-vpn-vtn-id] Li, Z. and J. Dong, "Carrying Virtual Transport Network (VTN) Information in MPLS Packet", Work in Progress, Internet-Draft, draft-li-mpls-enhanced-vpn-vtn-id-03, 16 October 2022, <https://datatracker.ietf.org/doc/html/draft-li-mpls-enhanced-vpn-vtn-id-03>.

[RFC4915] Psenak, P., Mirtorabi, S., Roy, A., Nguyen, L., and P. Pillay-Esnault, "Multi-Topology (MT) Routing in OSPF", RFC 4915, DOI 10.17487/RFC4915, June 2007, <https://www.rfc-editor.org/info/rfc4915>.

[RFC5120] Przygienda, T., Shen, N., and N. Sheth, "M-ISIS: Multi Topology (MT) Routing in Intermediate System to Intermediate Systems (IS-ISs)", RFC 5120, DOI 10.17487/RFC5120, February 2008, <https://www.rfc-editor.org/info/rfc5120>.

[RFC8667] Previdi, S., Ed., Ginsberg, L., Ed., Filsfils, C., Bashandy, A., Gredler, H., and B. Decraene, "IS-IS Extensions for Segment Routing", RFC 8667, DOI 10.17487/

RFC8667, December 2019, <https://www.rfc-editor.org/info/rfc8667>.

[RFC8668] Ginsberg, L., Ed., Bashandy, A., Filsfils, C., Nanduri, M., and E. Aries, "Advertising Layer 2 Bundle Member Link Attributes in IS-IS", RFC 8668, DOI 10.17487/RFC8668, December 2019, <https://www.rfc-editor.org/info/rfc8668>.

[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>.

Authors' Addresses

Jie Dong
Huawei Technologies

Email: jie.dong@huawei.com

Zhibo Hu
Huawei Technologies

Email: huzhibo@huawei.com

Zhenbin Li
Huawei Technologies

Email: lizhenbin@huawei.com

Xiongyan Tang
China Unicom

Email: tangxy@chinaunicom.cn

Ran Pang
China Unicom

Email: pangran@chinaunicom.cn