BGP-LS with Flex-Algo for Segment Routing based Virtual Transport Networks
draft-zhu-idr-bgpls-sr-vtn-flexalgo-01

Abstract

   Enhanced VPN (VPN+) aims to provide enhanced VPN service to support
   some application's needs of enhanced isolation and stringent
   performance requirements. VPN+ requires integration between the
   overlay VPN and a particular set of resources in the underlay
   network. A Virtual Transport Network (VTN) is a virtual underlay
   network which has a customized network topology and a set of network
   resources allocated from the physical network. A VTN could be used
   to support one or a group of VPN+ services.

   When Segment Routing is used as the data plane to provide VTNs, each
   VTN can be allocated with a group of SIDs to identify the topology
   and resource attributes of network segments in the VTN. The
   association between the network topology, the network resource
   attributes and the SR SIDs may need to be distributed to a
   centralized network controller. For network scenarios where each VTN
   can be identified by a unique Flex-Algo ID, this document describes a
   mechanism to distribute the information of SR based VTNs using BGP-LS
   with Flex-Algo.

Status of This Memo

This Internet-Draft will expire on August 26, 2021.

Table of Contents

1. Introduction

   Enhanced VPN (VPN+) is an enhancement to VPN services to support the
   needs of new applications, particularly including 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. Thus these
   properties require integration between the overlay VPN and the
   underlay 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 other
   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-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 VTNs. To allow
the network controller and network nodes to perform VTN-specific
explicit path computation and/or shortest path computation, the group
of resource-aware SIDs allocated by network nodes to each VTN and the
associated topology and resource attributes need to be distributed in
the control plane. When a centralized network controller is used for
VTN-specific path computation, especially when a VTN spans multiple
IGP areas or multiple Autonomous Systems (ASes), BGP-LS
[I-D.ietf-idr-rfc7752bis] is needed to advertise the VTN 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 VTNs.

In some network scenarios, each VTN can be identified by a unique
Flex-Algo ID [I-D.ietf-lsr-flex-algo].
[I-D.zhu-lsr-isis-sr-vtn-flexalgo] describes an IGP mechanism to
advertise the association between the Flex-Algo and the resource
attributes and the SR SIDs associated with each VTN. This document
describes a mechanism to distribute the information of SR based VTNs
to the network controller using BGP-LS with Flex-Algo.

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
BCP14 RFC 2119 [RFC2119] [RFC8174] when, and only when, they appear
in all capitals, as shown here.

2. Advertisement of VTN Topology Attribute

[I-D.zhu-lsr-isis-sr-vtn-flexalgo] describes the mechanism of using
ISIS Flex-Algo to distribute the topological constraints of SR based
VTNs. This section describes the corresponding BGP-LS mechanism to

   distribute both the intra-domain and inter-domain topology attributes
   of SR based VTNs.

2.1. Intra-domain Topology Advertisement

   [I-D.ietf-lsr-flex-algo] specifies the mechanism to provide
   distributed constraint-path computation, and the usage of SR-MPLS
   prefix-SIDs and SRv6 locators for steering traffic along the
   constrained paths. The Flex-Algo definition is the combination of
   calculation-type, metric-type and the topological constraints used
   for path computation. As specified in
   [I-D.zhu-lsr-isis-sr-vtn-flexalgo], the topology attribute of a VTN
   can be described by applying Flex-Algo constraints on a particular
   topology. If each VTN is associated with a unique Flex-Algo, the
   Flex-Algo ID could be reused as the identifier of the VTN in the
   control plane.

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

   In[I-D.ietf-idr-bgp-ls-segment-routing-ext], 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 the Prefix NLRI with the SRv6 Locator TLV
   carried in the associated BGP-LS Attribute, 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 .

2.2. Inter-Domain Topology Advertisement

   In some network scenarios, a VTNs which span multiple areas or ASes
   needs to be created. The multi-domain VTN 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 VTNs using segment
   routing, it is necessary to advertise the topology and resource
   attribute of VTN on the inter-domain links and the associated BGP
   Peering SIDs.

   [I-D.ietf-idr-bgpls-segment-routing-epe] and
   [I-D.ietf-idr-bgpls-srv6-ext] defines the BGP-LS extensions for

advertisement of BGP topology information between ASes and the 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 VTNs, different mechanisms can be used to specify the inter-domain connections of VTNs.

o 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 VTNs 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 several 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 VTN identifier associated with each BGP Peer-Adj-SID.

o 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 multi-domain VTN, while some other BGP sessions are used for another multi-domain VTN. In this case, different BGP peer-node-SIDs are allocated to each BGP session, and ASBRs SHOULD advertise the VTN identifier associated with each BGP Peer-node-SIDs.

o At the AS-level topology, different multi-domain VTNs 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 multi-domain VTN.

When Flex-Algo is used consistently in multiple ASes covered by a VTN, 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 Flex-Algo among multiple ASes can not be expected, a global-significant VTN-ID can be used to define the AS level VTN topologies. Within each domain, the Flex-Algo based mechanism could be used for intra-domain topology advertisement. The detailed mechanism is specified in [I-D.dong-idr-bgpls-sr-enhanced-vpn].

3. Advertisement of VTN Resource Attribute

   [I-D.zhu-lsr-isis-sr-vtn-flexalgo] specifies the mechanism to
   advertise the resource information associated with each VTN. It is
   based on the extensions to the advertisement of L2 bundle member
   links information[RFC8668]. This section defines the corresponding
   BGP-LS extensions.

   A new TLVs is defined to specify the attribute flags of either a
   Layer-3 link or a L2 bundle member 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 V is defined in this
      document. When the V flag is set, it indicates this link is a
      virtual link.

   In order to correlate the virtual or physical member links with the
   Flex-Algo used to identify the VTN, each VTN SHOULD be assigned with
   a unique Admin Group (AG) or Extended Admin Group (EAG), and the
   virtual or physical member link associated with this VTN SHOULD be
   configured with the AG or EAG assigned to the VTN. The AG or EAG of
   the Layer-3 link SHOULD be set to the union of all the AGs or EAGs of
   its virtual or physical member links. In the definition of the Flex-
   Algo corresponding to the VTN, It MUST use the Include-Any Admin
   Group rule with only the AG or EAG assigned to the VTN as the link
   constraints, the Include-All Admin Goup rule or the Exclude Admin
   Group rule MUST NOT be used. This ensures that the Layer-3 link is
   included in the Flex-Algo specific constraint path computation for
   each VTN it participates in.

The TE attributes of each Layer-3 link or Layer 2 bundle member link, such as the bandwidth, the Adj-SIDs or the SRv6 End.X SIDs, can be advertised using the mechanism as defined in [I-D.ietf-idr-bgp-ls-segment-routing-ext][I-D.ietf-idr-bgpls-segment-routing-epe] and [I-D.ietf-idr-bgpls-srv6-ext].

4. Scalability Considerations

The mechanism described in this document assumes that each VTN is associated with an unique Flex-Algo, so that the Flex-Algo IDs can be reused to identify the VTNs in the control plane. While this brings the benefit of simplicity, it also has some limitations. For example, it means that even if multiple VTNs share the same topological constraints, they would still need to be identified using different Flex-Algo IDs in the control plane, then independent path computation needs to be executed for each VTN. The number of VTNs supported in a network may be dependent on the number of Flex-Algos supported, which is related to the control plane computation overhead. Another aspect which may impact the number of VTNs supported with this mechanism is that at most 128 Flex-Algos can be used in a network.

Based on the above considerations, this mechanism is suitable for networks where a relatively small number of VTNs are needed.

5. 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.

6. IANA Considerations

TBD

7. Acknowledgments

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

8. References

8.1. Normative References

   [I-D.ietf-idr-bgp-ls-flex-algo]
              Talaulikar, K., Psenak, P., Zandi, S., and G. Dawra,
              "Flexible Algorithm Definition Advertisement with BGP
              Link-State", draft-ietf-idr-bgp-ls-flex-algo-05 (work in
              progress), November 2020.

   [I-D.ietf-idr-bgp-ls-segment-routing-ext]
              Previdi, S., Talaulikar, K., Filsfils, C., Gredler, H.,
              and M. Chen, "BGP Link-State extensions for Segment
              Routing", draft-ietf-idr-bgp-ls-segment-routing-ext-16
              (work in progress), June 2019.

   [I-D.ietf-idr-bgpls-segment-routing-epe]
              Previdi, S., Talaulikar, K., Filsfils, C., Patel, K., Ray,
              S., and J. Dong, "BGP-LS extensions for Segment Routing
              BGP Egress Peer Engineering", draft-ietf-idr-bgpls-
              segment-routing-epe-19 (work in progress), May 2019.

   [I-D.ietf-idr-bgpls-srv6-ext]
              Dawra, G., Filsfils, C., Talaulikar, K., Chen, M.,
              daniel.bernier@bell.ca, d., and B. Decraene, "BGP Link
              State Extensions for SRv6", draft-ietf-idr-bgpls-
              srv6-ext-05 (work in progress), November 2020.

   [I-D.ietf-idr-rfc7752bis]
              Talaulikar, K., "Distribution of Link-State and Traffic
              Engineering Information Using BGP", draft-ietf-idr-
              rfc7752bis-05 (work in progress), November 2020.

   [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", draft-ietf-spring-resource-aware-segments-01
              (work in progress), January 2021.

   [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", February 2021,
              <https://tools.ietf.org/html/draft-ietf-spring-sr-for-
              enhanced-vpn>.

   [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>.

   [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>.

   [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>.

8.2. Informative References

   [I-D.dong-idr-bgpls-sr-enhanced-vpn]
              Dong, J., Hu, Z., Li, Z., Tang, X., and R. Pang, "BGP-LS
              Extensions for Segment Routing based Enhanced VPN", draft-
              dong-idr-bgpls-sr-enhanced-vpn-02 (work in progress), June
              2020.

   [I-D.ietf-lsr-flex-algo]
              Psenak, P., Hegde, S., Filsfils, C., Talaulikar, K., and
              A. Gulko, "IGP Flexible Algorithm", draft-ietf-lsr-flex-
              algo-13 (work in progress), October 2020.

   [I-D.ietf-teas-enhanced-vpn]
              Dong, J., Bryant, S., Li, Z., Miyasaka, T., and Y. Lee, "A
              Framework for Enhanced Virtual Private Networks (VPN+)
              Service", draft-ietf-teas-enhanced-vpn-06 (work in
              progress), July 2020.

   [I-D.zhu-lsr-isis-sr-vtn-flexalgo]
              Zhu, Y., Dong, J., and Z. Hu, "Using Flex-Algo for Segment
              Routing based VTN", draft-zhu-lsr-isis-sr-vtn-flexalgo-01
              (work in progress), September 2020.

   [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>.

Authors' Addresses

   Yongqing Zhu
   China Telecom

   Email: zhuyq8@chinatelecom.cn

Jie Dong
Huawei Technologies

Email: jie.dong@huawei.com

Zhibo Hu
Huawei Technologies

Email: huzhibo@huawei.com