BGP based Virtual Private Network (VPN) Services over SRv6+ enabled IPv6 networks
draft-ssangli-idr-bgp-vpn-srv6-plus-01

Abstract

This document defines BGP protocol extensions for encoding and carrying SRv6+ Per-Path Service Instruction information to support Virtual Private Network services. This is applicable when the VPN services are offered in a SRv6+ enabled IPv6 network such that the VPN payload is transported over IPv6. The Per-Path Service Instruction information is encoded in the IPv6 Destination Option Header in the IPv6 data packets.

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 January 6, 2020.

Copyright Notice

Copyright (c) 2019 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
1. Introduction

Virtual Private Network (VPN) technologies allow network providers to emulate private networks with shared infrastructure. For example, assume that a set of red sites, set of blue sites and a set of green sites connect to a provider network. Furthermore, assume that red sites and blue sites wish to interconnect, exchange packets. However, the green sites wish to communicate with green sites only. The provider should allow its infrastructure network to scale to both the requirements without having to create multiple parallel network infrastructures. The IETF has standardized many VPN technologies viz. Layer 3 VPN (L3VPN) [RFC4364], Layer 2 VPN (L2VPN) [RFC6624], Virtual Private LAN Service (VPLS) [RFC4761], [RFC4762], Ethernet VPN (EVPN) [RFC7432], Pseudowires [RFC8077] to enable Layer 3 and Layer 2 VPN services.
The aforementioned technologies leverage MPLS network architecture:

- to establish a MPLS tunnel from ingress PE to egress PE, thus making all P routers agnostic of VPN state.
- to provide demultiplexing abstraction in the tunnelled packet so the payload packet can be forwarded at the egress router based on Routing table and/or interface.

In pure IPv6 deployments where there may be non-MPLS capable routers, it would be desirable to have alternate mechanism to provide VPN connectivity. This document describes BGP extensions and procedures applicable for SRv6+ enabled IPv6 networks, to provide VPN services over BGP.

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.

3. Per-Path Service Instruction Information

A SRv6+ [I-D.bonica-spring-srv6-plus] segment provides unidirectional connectivity from an ingress node to an egress node. A SRv6+ path contains one or more such segments. SRv6+ introduces the concept of Per-Segment Service Instruction and Per-Path Service Instruction. These instructions describe the additional packet processing performed on a node. The Per-Segment Service Instruction is executed on the segment egress node while the Per-Path Service Instruction is executed on the path egress node. The SR Path egress node advertises the reachability information to SR Path ingress node via Multi Protocol extensions in BGP [RFC4760].

For providing VPN services, aforementioned BGP extensions rely on MPLS architecture [RFC3031]. The BGP extensions specify the new encoding for Network Layer Reachability Information (NLRI) to include the MPLS VPN labels [RFC8277]. Such a MPLS VPN label is associated with a forwarding decision in the VPN Routing Instance on the egress BGP Router. The ingress BGP router will push the VPN label on the data packet destined to the egress BGP router. The transport tunnel from ingress router to egress router can be MPLS or GRE or L2TPv3, but inner payload is a MPLS packet as described in [RFC4023], [RFC4817], [RFC7510]. The intermediate routers do not process the VPN label [a.k.a.] embedded label as described in [I-D.ietf-idr-tunnel-encaps].
To provide BGP based VPN services on a non-MPLS IPv6 networks, it would be beneficial to retain the benefits of BGP protocol extensions while leveraging the benefits of IPv6 [RFC8200].

[I-D.bonica-6man-vpn-dest-opt] describes SRv6+ paths as programmable with Per-Path Service Instructions (PPSI) that determine how egress nodes process SRv6+ payloads. The PPSIs are carried in the PPSI Option encoded in the IPv6 Destination Option Header [RFC8200].

The Per-Path Service Instruction (PPSI) Identifier is defined as follows:

- 32 bit quantity.

The PPSI Identifier have node-local significance and is assigned by the egress BGP router. The value of zero is reserved. The PPSI Identifier will serve 2 purposes.

- It MUST uniquely identify the VPN Routing Instance for L3VPN or identify an Ethernet Segment for EVPN or identify a leaf property for EVPN TREE upon which forwarding decision can be taken.

- It MAY provide information for special processing before the packet is forwarded.

The structure of 3 octet PPSI Identifier will be updated in the next version of this document.

The encoding of the Per-Path Service Instruction Identifier for VPNs is described in Section 7 and Section 8.

4. Usage of Tunnel Encapsulation Attribute

This document defines a new Tunnel type : SRv6+. The format is as per below.

- Tunnel Type (2 Octets) : To be assigned
- Tunnel Length (2 Octets) : 1
- Value : List of Sub-TLVs

[I-D.ietf-idr-tunnel-encaps] defines many sub-TLVs for the tunnels. The encoding for them are as follows:

- Remote Endpoint sub-TLV : As per [I-D.ietf-idr-tunnel-encaps]
- Encapsulation sub-TLV : Not needed.
Internet-Draft

IPv4 DS Field sub-TLV : Not needed.

UDP Destination Port sub-TLV : Not needed.

Protocol type sub-TLV : As per [I-D.ietf-idr-tunnel-encaps].

Color Sub-TLV : As per [I-D.ietf-idr-tunnel-encaps].

Embedded Label Handling sub-TLV : 2.

MPLS Label Stack Sub-TLV : Not needed.

Prefix SID Sub-TLV : Not needed.

The Tunnel Encapsulation Attribute is an Optional Transitive attribute as described in [I-D.ietf-idr-tunnel-encaps]. This attribute with SRv6+ tunnel type MUST be present in the BGP update carrying the Network Layer Reachability Information encoded with the PPSI Information. This document refers to the NLRI that is associated with SRv6+ tunnel Encapsulation attribute as SRv6+_NLRI. The document [I.D.ietf-idr-tunnel-encaps-12] defines the encoding for sub-TLV as follows.

Sub-TLV Type : 1 octet

Sub-TLV Length : 1 or 2 octets

Sub-TLV Value : defined per Sub-TLV as per below.

The Remote Endpoint sub-TLV can specify the IPv6 address of the egress router as the final destination address of SRv6+ packet which is also referred to as SR Path destination address. The sub-fields on this sub-TLV is encoded as below.

Autonomous System Number : AS number of the IPv6 SR domain.

Address Family : 2 (refers to IPv6).

Address : IPv6 address of the egress interface present in SRv6+ domain.

The Value field may be set to 0 which indicates that next hop value in the NLRI should be chosen for the SRv6+ Path destination address.

The Embedded Label Handling sub-TLV describes how the label field in the NLRI should be interpreted.

Value : MUST be set to 2.
The value of 2 indicates that the label field in the NLRI MUST be ignored at the ingress router.

5. Procedures for Egress BGP Speaker

The PPSI Information instructs the egress router to de-encapsulate the packet and forward the newly exposed payload inner packet through the specified interface or forward using the specified Routing Instance. The PPSI Identifier described in Section 3 will be assigned by the egress BGP Router except in the case of EVPN per ES AD route when P2MP tunnel is used for delivering BUM traffic in EVPN. If P2MP tunnel is used to deliver BUM traffic for EVPN, the PPSI Identifier used to identify an Ethernet Segment is assigned by the upstream ingress BGP Router. Otherwise, it is downstream assigned by the egress BGP router.

When the egress BGP Speaker advertises the NLRI, it will include the PPSI Information in the encoding described in Section 7 and Section 8. The egress BGP Speaker MUST include the Tunnel Encapsulation Attribute with Route type SRv6+ as described in Section 4 in such BGP updates.

By tagging the BGP update with Tunnel Encapsulation attribute of SRv6+ type, the BGP Speaker informs how the SRv6+_NLRI should be decoded and processed by the receiving BGP Speaker.

Via the Remote Tunnel Endpoint Sub-TLV encoding, the egress BGP router may specify the SRv6+ Path Destination Address. The Protocol type Sub-TLV and the Color Sub-TLV may be used by the egress BGP router to influence the payload packets to be put on SRv6+ path. The Embedded Label Handling Sub-TLV MUST be set to 2 to inform that the MPLS label field should be ignored.

A single PPSI Identifier may be associated with all the prefixes in a Routing Instance or a unique PPSI Identifier may be associated for each prefix in the Routing Instance. Similarly, a PPSI Identifier may be assigned to identify an Ethernet segment or leaf AC property by EVPN. The choice is left to the Network Operator and is outside the scope of this document.

6. Procedures for Ingress BGP Speaker

Upon receiving a BGP update, the receiving BGP Speaker will look for Tunnel Encapsulation attribute. If the tunnel type carried in the Tunnel Encapsulation attribute is SRv6+, the BGP updates is said to be carrying the SRv6+_NLRI and the Label field in the Network Layer Reachability Information is treated as Per-Path Service Instruction (PPSI) Identifier.
The tuple (PPSI Identifier, Prefix) is programmed in the forwarding infrastructure of the router. The manner in which this tuple is stored in the router is outside the scope of this document. If SRv6+ has been enabled on the router, such a tuple SHOULD be used for encoding the Destination Options Header as described in [I-D.bonica-6man-vpn-dest-opt].

[I.D.ietf-idr-tunnel-encaps-12] describes how Remote Tunnel Endpoint Sub-TLV has to be processed. It also describes the usage of the Protocol type Sub-TLV and the Color Sub-TLV. This may be used by the ingress BGP router to select the payload packets that should be put on SRv6+ path.

The Embedded Label Handling Sub-TLV value that is set to 2 indicates that ingress BGP router to ignore the MPLS label field.

7. BGP based L3 VPN services over IPv6

The Egress and Ingress BGP speakers form a BGP peering session to exchange a set of prefixes described in [RFC4271] and Multi protocol extensions [RFC4760]. The BGP Router capable of SRv6+ that is enabled to carry L3 VPN services over IPv6 networks should follow the procedures mentioned in Section 5 and Section 6. The manner in which a BGP Router is configured for SRv6+ underlay and L3 VPN overlay is outside the scope of this document.

7.1. IPv4 VPN on SRv6+ enabled IPv6 Core

The IPv4 L3 VPN over IPv6 is defined in [RFC5549]. The MP_REACH NLRI and Tunnel Encapsulation attribute encoding is as per below:

- AFI : 1; SAFI : 128
- Length of the Next Hop : 16 (or 32 if Link Local)
- Network address of the Next Hop : IPv6 address of the egress BGP Router
- NLRI : IPv4-VPN routes
- Label : Per-Path Service Instruction Identifier
- Tunnel Encapsulation Path Attribute : SRv6+ Type as described in Section 4

The PPSI Identifier is associated with VPN Routing Instance on the Egress PE. The Tunnel Encapsulation attribute with SRv6+ type MUST be appended to the Path attributes associated with the NLRI.
7.2. IPv6 VPN on SRv6+ enabled IPv6 Core

The IPv6 L3 VPN over IPv6 is defined in [RFC4659]. The MP_REACH NLRI and Tunnel Encapsulation attribute encoding is as per below:

- AFI : 2; SAFI : 128
- Length of the Next Hop : 16 (or 32 if Link Local)
- Network address of the Next Hop : IPv6 address of the egress BGP Router
- NLRI : IPv6-VPN routes
- Label : Per-Path Service Instruction Identifier
- Tunnel Encapsulation Path Attribute : SRv6+ Type as described in Section 4

The PPSI Identifier is associated with VPN Routing Instance on the Egress PE. The Tunnel Encapsulation attribute with SRv6+ type MUST be appended to the Path attribute associated with the NLRI.

7.3. IPv4 Global Routes on SRv6+ enabled IPv6 Core

The IPv4 L3 VPN over IPv6 is defined in [RFC5549]. The MP_REACH NLRI and Tunnel Encapsulation attribute encoding is per below:

- AFI : 1; SAFI : 1
- Length of the Next Hop : 16 (or 32 if Link Local)
- Network address of the Next Hop : IPv6 address of the egress BGP Router
- NLRI : IPv4 routes
- Label : Per-Path Service Instruction Identifier
- Tunnel Encapsulation Path Attribute : SRv6+ Type as described in Section 4

The PPSI Identifier is associated with VPN Routing Instance on the Egress PE. The Tunnel Encapsulation attribute with SRv6+ type MUST be appended to the Path attribute associated with the NLRI.
8. BGP based Ethernet VPN services over IPv6

The [RFC7432] describes the BGP extensions for carrying the Ethernet Virtual Private Network Overlay on MPLS network. It defines 4 types of EVPN NLRI. This document specifies changes to certain fields for those NLRIs.

- Ethernet Auto-Discovery (A-D) route
- MAC/IP Advertisement route
- Inclusive Multicast Ethernet Tag route
- IP Prefix route

8.1. Ethernet Per ES Auto-Discovery (A-D) route

The MP_REACH and MP_UNREACH attributes will carry this route in the NLRI encoding described in [RFC7432]. In addition to Tunnel Encapsulation attribute encoding, this document recommends to follow the [RFC4732] encoding except the following. For MPLS label carried in the Ethernet A-D per ESI route:

- MPLS label: Per [RFC7432], it is set to zero.
- Tunnel Encapsulation Path Attribute: SRv6+ Type as described in Section 4

The MPLS label field is not part of the route but treated as route attribute. For procedures and usage of this route, refer to [RFC7432]. The Tunnel Encapsulation attribute with SRv6+ type MUST be appended to the Path attribute associated with the NLRI.

An EVPN Ethernet per ES A-D route is usually signaled together with an ESI label extended community. For ESI Label carried in the ESI label extended community:

- ESI Label: Per-Path Service Instruction Identifier

The Per-Path Service Instruction Identifier is used to identify an Ethernet segment attached to the BGP PE for EVPN.

If P2MP tunnel is used to deliver BUM traffic, then this PPSI Identifier is upstream assigned by the ingress router, otherwise it is downstream assigned by the egress router.
8.2. Ethernet per EVI Auto-Discovery (A-D) route

The MP_REACH and MP_UNREACH attributes will carry this route in the NLRI encoding described in [RFC7432]. In addition to Tunnel Encapsulation attribute encoding, this document recommends to follow the [RFC4732] encoding except the following:

- MPLS label: Per-Path Service Instruction Identifier
- Tunnel Encapsulation Path Attribute: SRv6+ Type as described in Section 4

The MPLS label field is not part of the route but treated as route attribute. For procedures and usage of this route, refer to [RFC7432]. The Tunnel Encapsulation attribute with SRv6+ type MUST be appended to the Path attribute associated with the NLRI.

In addition, for EVPN E-tree service, this route may be signaled together with an E-Tree Extended Community as it is specified in [RFC8317]. For the leaf label carried in the E-Tree Extended Community:

- Leaf Label: Per-Path Service Instruction Identifier

In case of EVPN E-tree service, the per-path service identifier carried in the E-Tree extended community is used to signal a leaf AC property.

In the data plane, this PPSI identifier specified in the Destination Option header is used by an egress router to identify that a data packet is ingressed from a leaf AC such that appropriate forwarding decision can be made.

If P2MP tunnel is used to deliver BUM traffic, then this PPSI Identifier is upstream assigned by the ingress router. Otherwise it is downstream assigned by the egress router.

8.3. MAC/IP Advertisement route

The MP_REACH and MP_UNREACH attributes will carry this route in the NLRI encoding described in [RFC7432]. In addition to Tunnel Encapsulation attribute encoding, this document recommends to follow the [RFC4732] encoding except the following:

- MPLS label1: Per-Path Service Instruction Identifier1
- MPLS label2: Per-Path Service Instruction Identifier2
8.4. Inclusive Multicast Ethernet Route

The MP_REACH and MP_UNREACH attributes will carry this route in the NLRI encoding described in [RFC7432]. In addition to Tunnel Encapsulation attribute encoding, this document recommends to follow the [RFC4732] encoding except the following.

- If MPLS label field in the PMSI Tunnel Attributed is non-zero, it is set to Per-Path Service Instruction Identifier.

- Tunnel Encapsulation Path Attribute : SRv6+ Type as described in Section 4

The Tunnel Encapsulation attribute with SRv6+ type MUST be appended to the Path attribute associated with the NLRI.

8.5. IP Prefix Route

The MP_REACH and MP_UNREACH attributes will carry this route in the NLRI encoding described in [I-D draft-ietf-bess-evpn-prefix-advertisement]. In addition to Tunnel Encapsulation attribute encoding, this document recommends the following change:

- MPLS label: if it is non-zero, it is set to Per-Path Service Instruction Identifier.

- Tunnel Encapsulation Path Attribute : SRv6+ Type as described in Section 4

The MPLS label field is not part of the route but treated as route attribute. For procedures and usage of this route, refer to [I-D draft-ietf-bess-evpn-prefix-advertisement]. The Tunnel Encapsulation attribute with SRv6+ type MUST be appended to the Path attribute associated with the NLRI.

9. Deployment Considerations

This document proposes to reuse the NLRI encoding for BGP L3VPN and EVPN Network Layer Routing Information. However, care should be taken when BGP VPN overlay services are enabled on SRv6+ underlay.
such that Tunnel Encapsulation Path attribute with SRv6+ type MUST be appended. When a BGP router advertises SRv6+_NLRI, it MUST NOT remove the Tunnel Encapsulation Path attribute.

The SRv6+ underlay is similar to other "tunnel" technologies viz MPLS, GRE, IP-in-IP, L2TPv3. The egress and ingress BGP routers can be connected via one or more such underlay technologies. A BGP speaker can advertise the VPN NLRI with the nexthop reachable via one or more such underlay paths. Each such mechanism can co-exist together as ships-in-night. However, when SRv6+_NLRI is advertised by a egress BGP speaker and received by an ingress BGP speaker, they MUST follow the procedures mentioned in this document.

For migrating a BGP router to SRv6+ the following procedures can be followed.

- Operator will enable SRv6+ underlay on the ingress and egress routers identifying the SRv6+ path from ingress router’s interface to egress router’s interface. The way to configure the ingress and egress routers are outside the scope of this document.

- SRv6+ enabled ingress BGP router will setup the additional information in the forwarding table such that it can append an IPv6 tunnel header and encode the PPSI Option in the Destination Options Header.

- SRv6+ enabled egress BGP router will setup the additional information in the forwarding table such that PPSI Identifier can be used to lookup to find the Routing Instance and make the forwarding decision.

- Operator will enable BGP VPN overlay over SRv6+ underlay on ingress router. This means that ingress router will start looking for SRv6+_NLRI in the BGP updates. The way to enable the BGP VPN overlay over SRv6+ underlay is outside the scope of this document.

- The operator will enable BGP VPN overlay over SRv6+ underlay on egress router. With this, the egress router will create PPSI Identifier and associate it with Routing Instances. It then advertises the SRv6+_NLRIs to the ingress BGP router.

- The ingress router will interpret the SRv6+_NLRIs and use PPSI identifier and follow the procedures in [I.D. bonica-spring-srv6-plus-00.txt] to encode the Destination Options Header to forward the data packet.

- Now that SRv6+ path is setup between ingress and egress BGP routers, on the egress BGP router the Operator can migrate the
Routing Instances from MPLS VPN set of Instances to SRv6+ enabled set of Instances. The way to configure Routing Instances to achieve the above is outside the scope of this document.

10. Backward Compatibility

The extension proposed in this document is backward compatible with procedures described for BGP enabled services.

11. Security Considerations

This document does not introduce any new security considerations beyond those already specified in [RFC4271], [RFC8277] and [I.D.ietf-idr-tunnel-encaps-12].

12. IANA Considerations

IANA is requested to assign a code point for SRv6+ Route Type for BGP Tunnel Encapsulation Path Attribute from BGP Tunnel Encapsulation Attribute Tunnel Types Registry.

13. Acknowledgements

The authors would like to thank Jeff Haas and Wen Lin for careful review and suggestions.

14. References

14.1. Normative References

[I-D.bonica-6man-vpn-dest-opt]

[I-D.bonica-spring-srv6-plus]

[I-D.ietf-idr-tunnel-encaps]


14.2. Informative References


Authors' Addresses

Srihari Sangli
Juniper Networks Inc.
Exora Business Park
Bangalore, KA 560103
India

Email: ssangli@juniper.net

Ron Bonica
Juniper Networks Inc.
2251 Corporate Park Drive
Herndon, Virginia 20171
USA

Email: rbonica@juniper.net