Inter-Domain Routing
Internet-Draft
Intended status: Standards Track
Expires: September 13, 2017

G. Dawra, Ed.
C. Filsfils
D. Dukes
P. Brissette
P. Camarilo
Cisco Systems
J. Leddy
Comcast
D. Voyer
D. Bernier
Bell Canada
D. Steinberg
Steinberg Consulting
R. Raszuk
Bloomberg LP
B. Decraene
Orange
S. Matsushima
SoftBank Telecom Japan

March 12, 2017

BGP Signaling of IPv6-Segment-Routing-based VPN Networks
draft-dawra-bgp-srv6-vpn-00.txt

Abstract

This draft defines procedures and messages for BGP SRv6-based EVPNs and L3 VPNs. It builds on RFC7432 "BGP MPLS-Based Ethernet VPN" and RFC4364 "BGP/MPLS IP Virtual Private Networks (VPNs)" to provide a migration path from MPLS-based VPNs to SRv6 based VPNs.

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 http://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 September 13, 2017.

Copyright Notice

Copyright (c) 2017 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 (http://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 Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction 2
2. Requirements Language 3
3. BGP for SRv6-L3VPN 3
3.1. SRv6-VPN SID TLV 4
3.2. IPv4 VPN Over SRv6 Core 5
3.3. IPv6 VPN Over SRv6 Core 6
4. Migration from L3 MPLS based Segment Routing to SRv6 Segment Routing 6
5. EVPN and SRv6 7
6. Error Handling of BGP SRv6 SID Updates 7
7. IANA Considerations 7
8. Security Considerations 7
9. Conclusions 8
10. References 8
10.1. Normative References 8
10.2. Informative References 9
Appendix A. Contributors 10
Authors' Addresses 10

1. Introduction

SRv6 refers to Segment Routing instantiated on the IPv6 dataplane [I-D.filsfils-spring-srv6-network-programming][I-D.ietf-6man-segment-routing-header].

SRv6-based VPN (SRv6-VPN) refers to the creation of VPN between PE's leveraging the SRv6 dataplane and more specifically the END.DT* (crossconnect to a VRF) and END.DX* (crossconnect to a nexthop) functions defined in the SRv6 network programming document

[I-D.filsfils-spring-srv6-network-programming]. SRv6-L3VPN refers to the creation of Layer3 VPN service between PE's supporting an SRv6 data plane.

SRv6 SID refers to a SRv6 Segment Identifier as defined in [I-D.filsfils-spring-srv6-network-programming].

SRv6-VPN SID refers to an SRv6 SID that MAY be associated with one of the END.DT or END.DX functions as defined in [I-D.filsfils-spring-srv6-network-programming].

To provide SRv6-VPN service with best-effort connectivity, the egress PE signals an SRv6-VPN SID with the VPN route. The ingress PE encapsulates the VPN packet in an outer IPv6 header where the destination address is the SRv6-VPN SID provided by the egress PE. The underlay between the PE's only need to support plain IPv6 forwarding [RFC2460].

To provide SRv6-VPN service in conjunction with an underlay SLA from the ingress PE to the egress PE, the egress PE colors the VPN route with a color extended community. The ingress PE encapsulates the VPN packet in an outer IPv6 header with an SRH that contains the SR policy associated with the related SLA followed by the SRv6-VPN SID associated with the route. The underlay nodes whose SRv6 SID's are part of the SRH must support SRv6 data plane.

BGP is used to advertise the reachability of prefixes in a particular VPN from an egress Provider Edge (egress-PE) to ingress Provider Edge (ingress-PE) nodes.

This document describes how existing BGP messages between PEs may carry SRv6 Segment IDs (SIDs) as the means to interconnect PEs and form VPNs.

2. 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 [RFC2119].

3. BGP for SRv6-L3VPN

BGP egress nodes (egress-PEs) advertise a set of reachable prefixes. Standard BGP update propagation schemes [RFC4271], which MAY make use of route reflectors [RFC4456], are used to propagate these prefixes. BGP ingress nodes (ingress-PE) receive these advertisements and may add the prefix to the RIB in an appropriate VRF.

For PEs supporting SRv6 the egress-PE advertises an SRv6-VPN SID with VPN routes. This SRv6-VPN SID only has local significance at the egress-PE where it is allocated or configured on a per-CE or per-VRF basis. In practice the SID encodes a cross-connect to a specific Address Family table (END.DT) or next-hop/interface (END.DX) as defined in the SRv6 Network Programming Document [I-D.filsfils-spring-srv6-network-programming]

The SRv6 VPN SID MAY be routable within the AS of the egress-PE and serves the dual purpose of providing reachability between ingress-PE and egress-PE while also encoding the VPN identifier.

For each NLRI, the egress-PE includes a new optional, transitive BGP SRv6-VPN SID Path TLV as part of the BGP Prefix-SID Attribute[I-D.ietf-idr-bgp-prefix-sid]. It contains a list of SIDs, for L3VPN only a single SRv6-VPN SID is necessary. See Section 3.1 below for details on the SRv6-VPN SID TLV.

At an ingress-PE, BGP installs the advertised prefix in the correct RIB table, recursive via an SR Policy leveraging the received SRv6-VPN SID.

Assuming best-effort connectivity to the egress PE, the SR policy has a single path with a single SID list made of a single SID: the SRv6-VPN SID received with the related route.

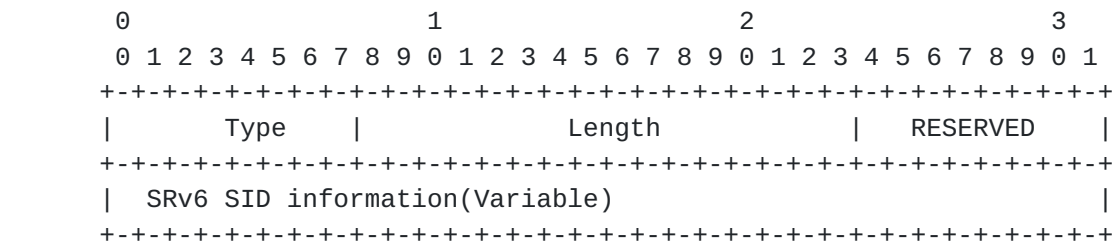
When the VPN route is colored with an extended color community C and the SID is next-hop N and the ingress PE has a valid SRv6 Policy (N, C) associated with SID list <S1,S2, S3> [I-D.filsfils-spring-segment-routing-policy] then the SR Policy is <S1, S2, S3, SRv6-VPN SID>.

Multiple VPN routes MAY recurse on the same SR Policy.

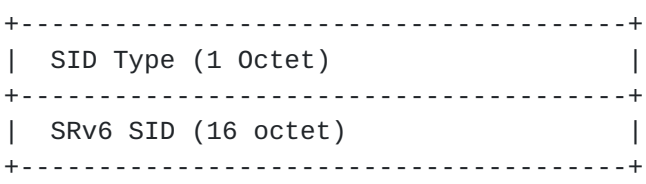
3.1. SRv6-VPN SID TLV

The SRv6-VPN SID TLV is defined as another TLV for BGP-Prefix-SID Attribute [I-D.ietf-idr-bgp-prefix-sid]. The value field of the BGP Prefix SID attribute is defined here to be a set of elements encoded as "Type/Length/Value" (i.e., a set of TLVs). Type for SRv6-VPN SID TLV is defined to be TBD.

The IPv6-SID TLV MUST be present in the Prefix-SID attribute attached to MP-BGP VPN NLRI defined in [RFC4659][RFC5549][RFC7432] when egress-PE is capable of SRv6 data-plane.



SRv6 SID information is encoded as follows:



Where:

- Type is TBD
- Length: 16bit field. The total length of the value portion of the TLV.
- RESERVED: 8 bit field. SHOULD be 0 on transmission and MUST be ignored on reception.

Current Type of SID defined as:

- Type-1 - corresponds to the equivalent functionality provided by an VPN MPLS Label attribute when received with a route containing a MPLS label[RFC4364].

3.2. IPv4 VPN Over SRv6 Core

IPv4 VPN Over IPv6 Core is defined in [RFC5549], the MP_REACH_NLRI is encoded as follows for an SRv6 Core:

- AFI = 1
- SAFI = 128
- Length of Next Hop Network Address = 16 (or 32)
- Network Address of Next Hop = IPv6 address of the egress PE
- NLRI = IPv4-VPN routes
- Label = Implicit-Null

SRv6-VPN SID are encoded as part of the SRv6-VPN SID TLV defined in Section 3.1. The function of the SRv6 SID is entirely up to the originator of the advertisement. In practice the function would likely be End.DX4 or End.DT4.

3.3. IPv6 VPN Over SRv6 Core

IPv6 VPN over IPv6 Core is defined in [RFC4659], the MP_REACH_NLRI is enclosed as follows for an SRv6 Core:

- AFI = 2
- SAFI = 128
- Length of Next Hop Network Address = 16 (or 32)
- Network Address of Next Hop = IPv6 address of the egress PE
- NLRI = IPv6-VPN routes
- Label = Implicit-Null

SRv6-VPN SID are encoded as part of the SRv6-VPN SID TLV defined in Section 3.1. The function of the IPv6 SRv6 SID is entirely up to the originator of the advertisement. In practice the function would likely be End.DX6 or End.DT6.

4. Migration from L3 MPLS based Segment Routing to SRv6 Segment Routing

Migration from IPv4 MPLS based underlay to an SRv6 underlay with BGP speakers is achieved with BGP sessions per BGP instance, one for IPv4 and a one for IPv6. Migration from IPv4 to IPv6 is independent of SRv6 BGP endpoints, and the selection of which route to use (received via the IPv4 or IPv6 session) is a local configurable decision of the ingress-PE, and is outside the scope of this document.

Migration from IPv6 MPLS based underlay to an SRv6 underlay with BGP speakers is achieved with a few simple rules at each BGP speaker.

```
At Egress-PE
  If BGP offers an SRv6-VPN service
      Then BGP allocates an SRv6-VPN SID for the VPN service
      and adds the BGP SRv6-VPN SID TLV while advertising VPN prefixes.
  If BGP offers an MPLS VPN service
      Then BGP allocates an MPLS Label for the VPN service and
      use it in NLRI as normal for MPLS L3 VPNs.

At Ingress-PE
  *Selection of which encapsulation below (SRv6-VPN or MPLS-VPN) is
   defined by local BGP policy
  If BGP supports SRv6-VPN service, and
  receives a BGP SRv6-VPN SID Attribute with an SRv6 SID
      Then BGP programs the destination prefix in RIB recursive via
      the related SR Policy.
  If BGP supports MPLS VPN service, and
  the MPLS Label is not Implicit-Null
      Then the MPLS label is used as a VPN label and inserted with the
      prefix into RIB via the BGP Nexthop.
```

5. EVPN and SRv6

The EVPN SRv6 solution is actively under definition and will be added in a later revision.

6. Error Handling of BGP SRv6 SID Updates

When a BGP Speaker receives a BGP Update message containing a malformed SRv6-VPN SID TLV, it MUST ignore the received BGP attributes and not pass it to other BGP peers. This is equivalent to the -attribute discard- action specified in [RFC7606]. When discarding an attribute, a BGP speaker MAY log an error for further analysis.

7. IANA Considerations

This memo includes no request to IANA.

8. Security Considerations

This document introduces no new security considerations beyond those already specified in [RFC4271] and [RFC3107].

9. Conclusions

This document proposes extensions to the BGP to allow advertising certain attributes and functionalities related to SRv6.

10. References

10.1. Normative References

[I-D.filsfils-spring-segment-routing-policy]
Filsfils, C., Sivabalan, S., Yoyer, D., Nanduri, M., Lin, S., bogdanov@google.com, b., Horneffer, M., Clad, F., Steinberg, D., Decraene, B., and S. Litkowski, "Segment Routing Policy for Traffic Engineering", draft-filsfils-spring-segment-routing-policy-00 (work in progress), February 2017.

[I-D.filsfils-spring-srv6-network-programming]
Filsfils, C., Leddy, J., daniel.voyer@bell.ca, d., daniel.bernier@bell.ca, d., Steinberg, D., Raszuk, R., Matsushima, S., Lebrun, D., Decraene, B., Peirens, B., Salsano, S., Naik, G., Elmalky, H., Jonnalagadda, P., Sharif, M., Ayyangar, A., Mynam, S., Bashandy, A., Raza, K., Dukes, D., Clad, F., and P. Camarillo, "SRv6 Network Programming", draft-filsfils-spring-srv6-network-programming-00 (work in progress), March 2017.

[I-D.ietf-6man-segment-routing-header]
Previdi, S., Filsfils, C., Field, B., Leung, I., Linkova, J., Aries, E., Kosugi, T., Vyncke, E., and D. Lebrun, "IPv6 Segment Routing Header (SRH)", draft-ietf-6man-segment-routing-header-05 (work in progress), February 2017.

[RFC2460] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", RFC 2460, DOI 10.17487/RFC2460, December 1998, <http://www.rfc-editor.org/info/rfc2460>.

[RFC3107] Rekhter, Y. and E. Rosen, "Carrying Label Information in BGP-4", RFC 3107, DOI 10.17487/RFC3107, May 2001, <http://www.rfc-editor.org/info/rfc3107>.

[RFC4364] Rosen, E. and Y. Rekhter, "BGP/MPLS IP Virtual Private Networks (VPNs)", RFC 4364, DOI 10.17487/RFC4364, February 2006, <http://www.rfc-editor.org/info/rfc4364>.

[RFC4456] Bates, T., Chen, E., and R. Chandra, "BGP Route Reflection: An Alternative to Full Mesh Internal BGP (IBGP)", RFC 4456, DOI 10.17487/RFC4456, April 2006, <http://www.rfc-editor.org/info/rfc4456>.

[RFC7432] Sajassi, A., Ed., Aggarwal, R., Bitar, N., Isaac, A., Uttaro, J., Drake, J., and W. Henderickx, "BGP MPLS-Based Ethernet VPN", RFC 7432, DOI 10.17487/RFC7432, February 2015, <http://www.rfc-editor.org/info/rfc7432>.

[RFC7606] Chen, E., Ed., Scudder, J., Ed., Mohapatra, P., and K. Patel, "Revised Error Handling for BGP UPDATE Messages", RFC 7606, DOI 10.17487/RFC7606, August 2015, <http://www.rfc-editor.org/info/rfc7606>.

10.2. Informative References

[I-D.ietf-idr-bgp-prefix-sid]
Previdi, S., Filsfils, C., Lindem, A., Patel, K., Sreekantiah, A., Ray, S., and H. Gredler, "Segment Routing Prefix SID extensions for BGP", draft-ietf-idr-bgp-prefix-sid-04 (work in progress), December 2016.

[I-D.ietf-isis-segment-routing-extensions]
Previdi, S., Filsfils, C., Bashandy, A., Gredler, H., Litkowski, S., Decraene, B., and j. jefftant@gmail.com, "IS-IS Extensions for Segment Routing", draft-ietf-isis-segment-routing-extensions-11 (work in progress), March 2017.

[I-D.ietf-spring-segment-routing]
Filsfils, C., Previdi, S., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing Architecture", draft-ietf-spring-segment-routing-11 (work in progress), February 2017.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <http://www.rfc-editor.org/info/rfc2119>.

[RFC4271] Rekhter, Y., Ed., Li, T., Ed., and S. Hares, Ed., "A Border Gateway Protocol 4 (BGP-4)", RFC 4271, DOI 10.17487/RFC4271, January 2006, <http://www.rfc-editor.org/info/rfc4271>.

[RFC4659] De Clercq, J., Ooms, D., Carugi, M., and F. Le Faucheur, "BGP-MPLS IP Virtual Private Network (VPN) Extension for IPv6 VPN", RFC 4659, DOI 10.17487/RFC4659, September 2006, <http://www.rfc-editor.org/info/rfc4659>.

[RFC5549] Le Faucheur, F. and E. Rosen, "Advertising IPv4 Network Layer Reachability Information with an IPv6 Next Hop", RFC 5549, DOI 10.17487/RFC5549, May 2009, <http://www.rfc-editor.org/info/rfc5549>.

Appendix A. Contributors

Bart Peirens
Proximus
Belgium

Email: bart.peirens@proximus.com

Authors' Addresses

Gaurav Dawra (editor)
Cisco Systems
USA

Email: gdawra@cisco.com

Clarence Filsfils
Cisco Systems
Belgium

Email: cfilsfil@cisco.com

Darren Dukes
Cisco Systems
Canada

Email: ddukes@cisco.com

Patrice Brissette
Cisco Systems
Canada

Email: pbrisset@cisco.com

Pablo Camarilo
Cisco Systems
Spain

Email: pcamaril@cisco.com

Jonn Leddy
Comcast
USA

Email: john_leddy@cable.comcast.com

Daniel Voyer
Bell Canada
Canada

Email: daniel.voyer@bell.ca

Daniel Bernier
Bell Canada
Canada

Email: daniel.bernier@bell.ca

Dirk Steinberg
Steinberg Consulting
Germany

Email: dws@steinberg.net

Robert Raszuk
Bloomberg LP
USA

Email: robert@raszuk.net

Bruno Decraene
Orange
France

Email: bruno.decraene@orange.com

Satoru Matsushima
SoftBank Telecom Japan
Japan

Email: satoru.matsushima@g.softbank.co.jp