

routing  
Internet-Draft  
Intended status: Standards Track  
Expires: April 28, 2022

Z. Zhang  
Juniper Networks  
October 25, 2021

MVPN Inter/Intra-region Tunnel Segmentation  
draft-zzhang-bess-mvpn-regional-segmentation-01

## Abstract

[RFC7524](#) specifies procedures for Inter-Area Point-to-Multipoint Segmented Label Switched Paths (aka MVPN tunnel segmentation). This document updates [RFC7524](#) by extending the inter-area segmentation concept to inter-region and intra-region segmentation where a region is no longer tied to an IGP area.

## 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 April 28, 2022.

## Copyright Notice

Copyright (c) 2021 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 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.

---

Internet-Draft      MVPN Inter/Intra-region Segmentation      October 2021

## Table of Contents

<a href="#">1.</a>	Introduction . . . . .	<a href="#">2</a>
<a href="#">1.1.</a>	Tunnel Segmentation . . . . .	<a href="#">2</a>
<a href="#">1.2.</a>	Intra-region Segmentation . . . . .	<a href="#">3</a>
<a href="#">1.3.</a>	Bud Node Support . . . . .	<a href="#">3</a>
<a href="#">2.</a>	Specifications . . . . .	<a href="#">3</a>
<a href="#">2.1.</a>	Inter-region Segmentation . . . . .	<a href="#">3</a>
<a href="#">2.2.</a>	Intra-region Segmentation . . . . .	<a href="#">4</a>
<a href="#">2.3.</a>	Bud Node Support . . . . .	<a href="#">5</a>
<a href="#">3.</a>	Security Considerations . . . . .	<a href="#">6</a>
<a href="#">4.</a>	IANA Considerations . . . . .	<a href="#">6</a>
<a href="#">5.</a>	Acknowledgements . . . . .	<a href="#">6</a>
<a href="#">6.</a>	References . . . . .	<a href="#">6</a>
<a href="#">6.1.</a>	Normative References . . . . .	<a href="#">6</a>
<a href="#">6.2.</a>	Informative References . . . . .	<a href="#">7</a>
	Author's Address . . . . .	<a href="#">7</a>

## [1.](#) Introduction

### [1.1.](#) Tunnel Segmentation

[RFC6514] specifies (among other things) inter-AS MVPN tunnel segmentation procedures and [\[RFC7524\]](#) specifies inter-area MVPN tunnel segmentation procedures. The procedures for inter-AS and inter-area are similar in that the segmentation points - ASBRs or ABRs - change the PMSI Tunnel Attribute (PTA) attached to I/S-PMSI routes to specify the type and identification of tunnel to be used in the next AS/area, when they re-advertise the PMSI routes into the next AS/area.

This change of tunnel at the segmentation points and stitching of upstream and downstream tunnel segments not only allows different tunnel technology/instance to be used in different AS/area, but also limits the replication of traffic to only PEs and segmentation points in the local AS/area, instead of to all PEs.

The inter-area segmentation points are route reflectors and when they re-advertise the x-PMSI routes to different downstream areas they may use different BGP neighbor groups so that different tunnel type/identification can be encoded in PTA for different downstream areas. If the ABR is also responsible for reflecting the routes to PEs in the same area, the ABR does not modify the PTA (because of that those

local PEs are also put into a different neighbor group).

As a result, a segmentation point will likely have different neighbor groups (one group for each area) so that the PTA and Inter-Area P2MP

Segmented Next-Hop Extended Community (referred to as Segmentation EC) can be set accordingly when it re-advertise the x-PMSI routes.

The provisioning of a RR with these different neighbor groups for segmentation purpose can actually be done on any router (as a segmentation point) – not necessarily on an ABR. As a result, the procedures in [RFC7524](#), while specified for inter-area, can be extended inter-region as well – the segmentation points can be any border routers between arbitrarily defined "regions".

This concept is already described in Section 6 of [\[I-D.ietf-bess-evpn-bum-procedure-updates\]](#), but specified formally in this document for MVPN.

## [1.2.](#) Intra-region Segmentation

Even with the inter-area segmentation extended to inter-region, when a regional border router (RBR) reflects routes to PEs in the same region, it does not modify the PTA or Segmentation EC. But if the RBR also modifies the two attributes when reflecting routes to the local PEs, tunnel segmentation is achieved even intra-region – both the upstream and downstream tunnel segments are in the same region.

This Intra-region Segmentation is one way to achieve Assisted Replication in MVPN: a PE sends traffic to assisting replicators who will then relay traffic to other PEs (even in the same region).

## [1.3.](#) Bud Node Support

A segmentation point may have both local receivers off a VRF and downstream receivers off a remote PE for traffic arriving on an upstream segment. This segmentation point is referred to a bud node, just like that a node can be both a transit and leaf node for a P2MP tree.

Depending on implementation, a bud node may need to receive two

copies of a packet, one for local delivery and one for remote delivery. If so, the bud node may request the upstream PE or segmentation point to send two copies.

## [2.](#) Specifications

### [2.1.](#) Inter-region Segmentation

The procedures in [RFC7524](#) are extended to beyond IGP area-based. A provider network can be arranged into "regions" connected by "Regional Border Routers" (RBRs). On a segmentation point a region MAY be defined as a BGP neighbor group - all peers in the group are

Zhang

Expires April 28, 2022

[Page 3]

---

Internet-Draft

MVPN Inter/Intra-region Segmentation

October 2021

subject to the same export policy, which can be used to control the modification of attributes for the purpose of segmentation.

[RFC7524](#) procedures apply as is, though "area" is replaced with "region" and "Area Border Router" (ABR) is replaced with "Regional Border Router" (RBR).

The concept of Per-region Aggregation, as explained in Section 6.1 of [\[I-D.ietf-bess-evpn-bum-procedure-updates\]](#), is also applicable to MVPN. A future revision of this document will specify details of Per-region Aggregation for MVPN.

### [2.2.](#) Intra-region Segmentation

The following procedures are applicable for intra-region segmentation. One use of intra-region segmentation is for Assisted Replication where PE-PE traffic goes through a relay point (assisting replicator).

If it is known that the local PEs are only peered with the RBRs (as RRs and segmentation points), the PEs and RBRs follow the procedures in [RFC7524](#). In addition, the local RBRs modify the PTA and Segmentation EC even when they re-advertise x-PMSI routes to PEs in the ingress region, thus achieving Intra-region Segmentation.

Otherwise (i.e., if a local PE may import BGP-MVPN routes directly unless with the modified procedures specified below), the following modified procedures apply:

- o When an ingress PE advertises an x-PMSI route, it attaches an Extended Community (EC) derived from the Route Target for the VPN (RT-VPN) [[I-D.zzhang-idr-rt-derived-community](#)] but not the RT-VPN itself. Call this EC as EC-VPN. The route also carry a Segmentation EC as specified in [RFC7524](#).
- o When the local RBRs (as RRs and segmentation points) receive this route, it replaces the EC-VPN with the corresponding RT-VPN (the EC-VPN and RT-VPN can be derived from each other), and then re-advertise the route to its peers, with the Segmentation EC and PTA modified as specified in [RFC7524](#). The modification applies even when re-advertising to peers in the same ingress region.

This is to ensure that local egress PEs will only import the routes re-advertised by the RBRs after the modification of PTA and Segmentation EC.

Additionally, if there are intermediate RRs between the ingress PE and local RBRs, and Route Target Constrain [[RFC4684](#)] is in use, the

ingress PE MUST also attach a Route Target (referred to as RT-RBR) and the local RBRs MUST be provisioned to import routes with RT-RBR (otherwise the intermediate RRs will not re-advertise the routes towards the RBRs because the routes carry only EC-VPN but not RT-VPN). The local RBRs MUST remove the RT-RBR when they re-advertise the routes.

### [2.3.](#) Bud Node Support

This section applies only if the segmentation point can not both route traffic arriving on the upstream segment to local receivers and label switch the traffic to downstream segments due to implementation limitation.

If a segmentation point is a bud node for a segmented x-PMSI tunnel with the above mentioned limitation, it SHOULD request an additional copy to be sent by the upstream RSVP neighbor if the upstream segment is a RSVP-TE P2MP tunnel, or by the upstream PE/RBR when the upstream segment is an IR or mLDP tunnel.

The RSVP-TE P2MP case is outside the scope of this document (though there are known implementations). For the IR/mLDP case, it is done

by including a Tunnel Encapsulation Attribute (TEA) [[RFC9012](#)] in the Leaf A-D route in response to the x-PMSI route for the upstream segment. Note that the leaf A-D route is sent for this purpose even if the Leaf Information Required (LIR) flag is not set in the x-PMSI route (e.g. for mLDP tunnel).

The TEA includes one tunnel of a desired type (e.g. MPLS or Any Encapsulation [[I-D.ietf-bess-bgp-multicast-controller](#)]) that is used for the upstream PE/RBR to send the additional copy to this bud node. The tunnel MUST include a Tunnel Egress Endpoint sub-TLV set to a local address on the bud node, and MUST include a Tree Label Stack sub-TLV that includes a single label. The node MUST program a label forwarding entry to pop the label and forward packet based on IP lookup in a VRF identified by the label (while the tunnel label for the upstream segment or the label in the PTA of the x-PMSI/Leaf route for the upstream segment is used to stitch the upstream and downstream segments together).

When the upstream PE/RBR decodes the TEA in the Leaf A-D route in response to an x-PMSI A-D route that it (re-)advertises (even if it did set the LIR flag in the x-PMSI A-D route), it SHOULD send an extra copy via unicast tunneling with the label encoded in the Tree Label Stack sub-TLV. If the extra copy is not sent the downstream bud node segmentation point will not be able to send traffic to its local receivers.

### [3.](#) Security Considerations

No additional security considerations are needed beyond what are discussed in [RFC7524](#).

### [4.](#) IANA Considerations

This document requests the IANA to create a "PMSI Tunnel Attribute Extension sub-TLV Type Registry". Allocation from the registry is First Come First Serve, with an initial allocation for "Additional Label".

### [5.](#) Acknowledgements

The author thanks Sanoj Vivekanandan for his review and comments.

## [6.](#) References

### [6.1.](#) Normative References

- [I-D.ietf-bess-bgp-multicast-controller]  
Zhang, Z., Raszuk, R., Pacella, D., and A. Gulko,  
"Controller Based BGP Multicast Signaling", [draft-ietf-bess-bgp-multicast-controller-07](#) (work in progress), July 2021.
- [I-D.zzhang-idr-rt-derived-community]  
Zhang, Z., "Extended Communities Derived from Route Targets", [draft-zzhang-idr-rt-derived-community-01](#) (work in progress), March 2021.
- [RFC6514] Aggarwal, R., Rosen, E., Morin, T., and Y. Rekhter, "BGP Encodings and Procedures for Multicast in MPLS/BGP IP VPNs", [RFC 6514](#), DOI 10.17487/RFC6514, February 2012, <<https://www.rfc-editor.org/info/rfc6514>>.
- [RFC7524] Rekhter, Y., Rosen, E., Aggarwal, R., Morin, T., Grosclaude, I., Leymann, N., and S. Saad, "Inter-Area Point-to-Multipoint (P2MP) Segmented Label Switched Paths (LSPs)", [RFC 7524](#), DOI 10.17487/RFC7524, May 2015, <<https://www.rfc-editor.org/info/rfc7524>>.
- [RFC9012] Patel, K., Van de Velde, G., Sangli, S., and J. Scudder, "The BGP Tunnel Encapsulation Attribute", [RFC 9012](#), DOI 10.17487/RFC9012, April 2021, <<https://www.rfc-editor.org/info/rfc9012>>.

### [6.2.](#) Informative References

- [I-D.ietf-bess-evpn-bum-procedure-updates]  
Zhang, Z., Lin, W., Rabadan, J., Patel, K., and A. Sajassi, "Updates on EVPN BUM Procedures", [draft-ietf-bess-evpn-bum-procedure-updates-11](#) (work in progress), October 2021.

[RFC4684] Marques, P., Bonica, R., Fang, L., Martini, L., Raszuk, R., Patel, K., and J. Guichard, "Constrained Route Distribution for Border Gateway Protocol/MultiProtocol Label Switching (BGP/MPLS) Internet Protocol (IP) Virtual Private Networks (VPNs)", [RFC 4684](https://www.rfc-editor.org/info/rfc4684), DOI 10.17487/RFC4684, November 2006, <<https://www.rfc-editor.org/info/rfc4684>>.

#### Author's Address

Zhaohui Zhang  
Juniper Networks

Email: [zzhang@juniper.net](mailto:zzhang@juniper.net)