

lsr
Internet-Draft
Intended status: Standards Track
Expires: 8 September 2022

Z. Zhang
Juniper Networks
7 March 2022

Anycast Affiliation Advertisement
draft-zzhang-lsr-anycast-affiliation-00

Abstract

This document specifies the advertisement of addresses affiliated with an anycast address in ISIS/OSPF/BGP, and describes one example use of such advertisement in VxLAN interconnect with EVPN.

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 8 September 2022.

Copyright Notice

Copyright (c) 2022 IETF Trust and the persons identified as the document authors. All rights reserved.

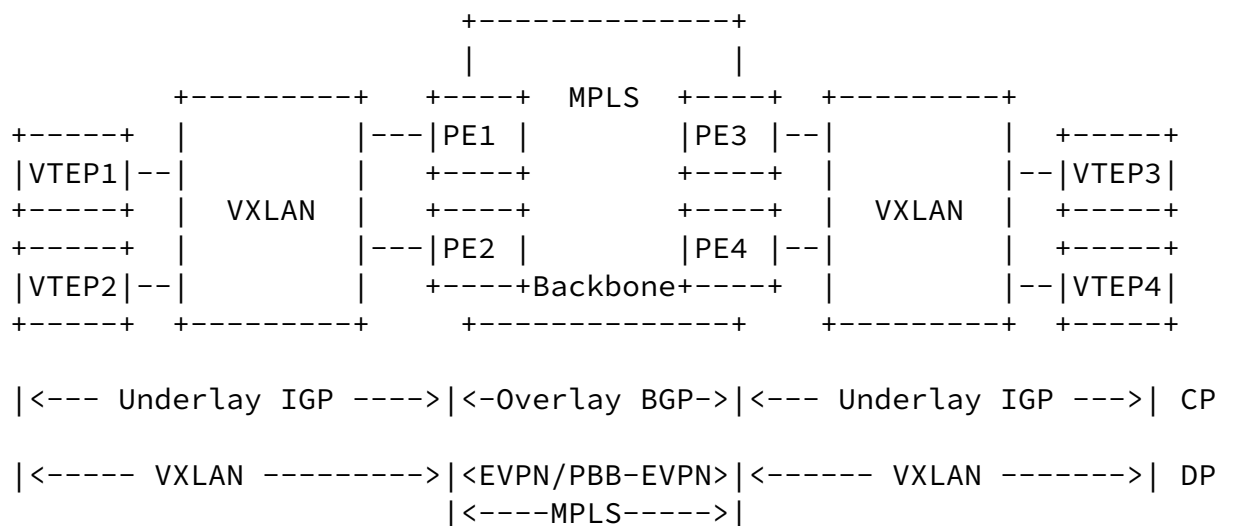
This document is subject to [BCP 78](#) and the IETF Trust's Legal Provisions Relating to IETF Documents (<https://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Revised BSD License text as described in Section 4.e of the [Trust Legal Provisions](#) and are provided without warranty as described in the Revised BSD License.

Table of Contents

1.	Background	2
2.	Specifications	4
2.1.	ISIS Encoding	4
2.2.	OSPF Encoding	4
2.3.	BGP Encoding	4
3.	Security Considerations	4
4.	IANA Considerations	5
5.	Acknowledgements	5
6.	References	5
6.1.	Normative References	5
6.2.	Informative References	5
	Author's Address	6

1. Background

[I-D.boutros-bess-vxlan-evpn] describes how Ethernet VPN (EVPN) [RFC7432] can be used to interconnect VXLAN/NVGRE networks over an MPLS/IP network. In the following diagram copied from that draft, a pair of gateways PE1/PE2, which are both VXLAN/NVGRE VTEPs and EVPN PEs, connect a VXLAN/NVGRE site to the EVPN Data Center Interconnect (DCI).



Legend: CP = Control Plane View

DP = Data Plane View

PE1/PE2 use a common anycast address as the source address of VXLAN/

NVGRE packets towards other VXLAN/NVGRE VTEPs. This serves two purposes as stated in [[I-D.boutros-bess-vxlan-evpn](#)]:

- * It prevents any flip/flopping in the forwarding tables for the MAC-to-VTEP associations

Zhang

Expires 8 September 2022

[Page 2]

Internet-Draft

Anycast Affiliation

March 2022

- * It enables load-balancing via ECMP for DCI traffic among the multi-homed PEs

Notice that load-balancing is only possible with ECMP - a VTEP uses the anycast address as the destination address in its VXLAN/NVGRE packets towards the DCI, and ECMP-based load-balancing can happen on any router from the source VTEP towards PE1/PE2.

However, if the source VTEP knows that PE1/PE2's specific non-anycast address that are associated with the anycast address, it can load-balance traffic using those specific addresses, even when there is no ECMP to PE1/PE2.

There are two ways for a VTEP to learn the affiliation of an address to an anycast address:

- * Provisioning on the VTEPs - statically or from a controller
- * Dynamic advertisement via IGP/BGP from the anycast address "owners" (PE1/PE2)

This document specifies the dynamic advertisement of anycast address affiliation via IGP/BGP, which can actually serve other general purposes as well.

Even though if a node advertises anycast and non-anycast local addresses using existing methods (e.g. with ISIS's N-bit) then others can derive that those addresses are owned by the same advertiser, it is desired to explicitly announce the non-anycast addresses affiliated with an anycast address for the following reasons:

- * Just because a node advertises two addresses A and B it does not mean A and B are exchangeable when another node needs to send traffic to it. Explicit association should be known for the source node to choose an "alias" address to send traffic.

- * The affiliation may have to be one-way - if a node needs to send traffic to an anycast address it can use the affiliated non-anycast addresses (or even a transit node may choose to change the destination address from anycast to an affiliated address) but the other way may not be desired (or allowed at all in the case of transit node changing destination address - otherwise loops could happen).
- * A node may have different anycast addresses and it may be necessary to associate different non-anycast addresses to different non-anycast addresses.

- * A node may also withdraw the affiliation advertisement even when all the relevant addresses are still reachable. In that case, only the affiliation is withdrawn, not the reachability.

For example, in the above VXLAN-EVPN interconnect scenario, if PE2 loses all its EVPN peers, it may withdraw anycast address affiliation advertisement and then the VTEPs can stop sending traffic to it.

[2.](#) Specifications

[2.1.](#) ISIS Encoding

A new "Anycast Affiliation sub-TLV" is defined for ISIS extended Reachability TLVs [[RFC7794](#)]. It MAY be advertised with an anycast host prefix. Its value part includes a list of affiliated local addresses. The number of affiliated addresses is determined from the length of the sub-TLV.

[2.2.](#) OSPF Encoding

A new "Anycast Affiliation sub-TLV" is defined for OSPFv2 Extended Prefix TLVs [[RFC7684](#)]. It MAY be advertised with an anycast host prefix. Its value part includes a list of affiliated local addresses. The number of affiliated addresses is determined from the length of the sub-TLV.

A new "Anycast Affiliation sub-TLV" is defined for OSPFv3 Intra-Area-Prefix TLV and Inter-Area-Prefix TLV [[RFC8362](#)]. It MAY be advertised

with an anycast host prefix. Its value part includes a list of affiliated local addresses. The number of affiliated addresses is determined from the length of the sub-TLV.

[2.3.](#) BGP Encoding

IPv4 (or IPv6) Address Specific Extended Communities with an "Anycast Affiliation" sub-type (or type in case of IPv6) MAY be attached to the NLRI of an anycast host prefix, one such extended community for each affiliated address. The Global Administrator field of the extended community is set to the affiliated address, and the Local Administrator field is set to 0.

[3.](#) Security Considerations

To be provided.

[4.](#) IANA Considerations

IANA is requested to allocate the following:

- * A new sub-TLV type for "Anycast Affiliation" from the ISIS "Sub-TLVs for TLVs 135, 235, 236, and 237" registry [[RFC7794](#)]
- * A new sub-TLV type for "Anycast Affiliation" from the OSPFv2 Extended Prefix TLV Sub-TLVs" registry [[RFC7684](#)]
- * A new sub-TLV type for "Anycast Affiliation" from the OSPFv3 Extended-LSA sub-TLV registry [[RFC8362](#)]
- * A new sub-type for "Anycast Affiliation" from the Transitive IPv4-Address-Specific Extended Community Sub-Types registry and Non-Transitive IPv4-Address-Specific Extended Community Sub-Types registry
- * A new type for "Anycast Affiliation" from the Transitive IPv6-Address-Specific Extended Community Types registry and Non-Transitive IPv6-Address-Specific Extended Community Types registry

5. Acknowledgements

The author thanks Shraddha Hegde for her inputs and comments.

6. References

6.1. Normative References

- [RFC7684] Psenak, P., Gredler, H., Shakir, R., Henderickx, W., Tantsura, J., and A. Lindem, "OSPFv2 Prefix/Link Attribute Advertisement", [RFC 7684](#), DOI 10.17487/RFC7684, November 2015, <<https://www.rfc-editor.org/info/rfc7684>>.
- [RFC7794] Ginsberg, L., Ed., Decraene, B., Previdi, S., Xu, X., and U. Chunduri, "IS-IS Prefix Attributes for Extended IPv4 and IPv6 Reachability", [RFC 7794](#), DOI 10.17487/RFC7794, March 2016, <<https://www.rfc-editor.org/info/rfc7794>>.
- [RFC8362] Lindem, A., Roy, A., Goethals, D., Reddy Vallem, V., and F. Baker, "OSPFv3 Link State Advertisement (LSA) Extensibility", [RFC 8362](#), DOI 10.17487/RFC8362, April 2018, <<https://www.rfc-editor.org/info/rfc8362>>.

6.2. Informative References

Zhang

Expires 8 September 2022

[Page 5]

Internet-Draft

Anycast Affiliation

March 2022

[I-D.boutros-bess-vxlan-evpn]

Boutros, S., Sajassi, A., Salam, S., Cai, D., Thoria, S., Singh, T., Drake, J., and J. Tantsura, "VXLAN DCI Using EVPN", Work in Progress, Internet-Draft, [draft-boutros-bess-vxlan-evpn-02](#), 21 October 2016, <<https://www.ietf.org/archive/id/draft-boutros-bess-vxlan-evpn-02.txt>>.

- [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, <<https://www.rfc-editor.org/info/rfc7432>>.

Author's Address

Zhaohui Zhang
Juniper Networks
Email: zzhang@juniper.net