

BESS Workgroup  
Internet-Draft  
Intended status: Standards Track  
Expires: 23 April 2026

J. Rabadan, Ed.  
Nokia  
A. Sajassi  
Cisco  
20 October 2025

Ingress Replication BUM flag in VXLAN  
draft-rabsaj-bess-evpn-vxlan-ir-bum-00

## Abstract

This document proposes the allocation of an “Ingress Replication BUM” flag in the VXLAN Flags IANA registry and defines its use in EVPN networks that employ VXLAN tunnels with ingress replication to forward broadcast, unknown and multicast traffic.

## 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 23 April 2026.

## Copyright Notice

Copyright (c) 2025 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. Introduction . . . . .	2
1.1. Terminology and Conventions . . . . .	3
2. The Ingress Replication BUM Flag in VXLAN tunnels . . . . .	3
3. Use of the Ingress Replication BUM Flag in EVPN Multi-Homing . . . . .	4
4. Security Considerations . . . . .	6
5. IANA Considerations . . . . .	6
6. Contributors . . . . .	6
7. Acknowledgements . . . . .	6
8. References . . . . .	7
8.1. Normative References . . . . .	7
8.2. Informative References . . . . .	7
Authors' Addresses . . . . .	7

## 1. Introduction

Section 8.3.3 of [RFC8365] explains why tunnels used for EVPN ingress replication must include an indication that a given BUM (Broadcast, Unknown Unicast, Multicast) frame was ingress-replicated by the ingress PE. A summary of that section is provided here for the reader's convenience:

In EVPN networks using ingress replication, unknown unicast traffic is flooded with a distinct MPLS label to mark it as BUM traffic. This allows egress PEs to apply Designated Forwarder (DF) filtering for All-Active multihomed sites. If unknown unicast flooding is enabled but no specific unknown-unicast label is used, transient duplicate traffic may occur. This happens when the ingress PE has not yet learned a host MAC via EVPN, while the egress PEs already have. The ingress PE floods the packet to all egress PEs, which then forward multiple copies to the multihomed site. Such duplication occurs only when:

- a. the destination host is multihomed (All-Active)
- b. unknown-unicast flooding is enabled
- c. ingress replication is used
- d. the ingress PE hasn't learned the MAC yet

To prevent this, [RFC8365] recommends the use of VXLAN-GPE encapsulation [I-D.ietf-nvo3-vxlan-gpe], with the ingress PE setting the BUM Traffic Bit (B-bit) to mark the traffic as ingress-replicated BUM.

The goal of this document is twofold. First, it requests the allocation of the B-bit (also referred to as the "Ingress Replication BUM" flag) in the VXLAN Flags registry established by [I-D.ietf-nvo3-rfc7348bis], and updates [RFC8365] to allow the use of this flag in VXLAN tunnels. Second, it analyzes additional use cases where the "Ingress Replication BUM" flag should be applied, beyond the scenario described in Section 8.3.3 of [RFC8365].

## 1.1. Terminology and Conventions

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.

- \* BUM traffic: Broadcast, Unknown unicast and Multicast traffic.
- \* PE: Provider Edge device. In this document a PE can be a Leaf node in a Data Center or a traditional Provider Edge router in an MPLS network.

## 2. The Ingress Replication BUM Flag in VXLAN tunnels

This document requests the allocation of the the Ingress Replication BUM flag (B-bit) in VXLAN tunnels [I-D.ietf-nvo3-rfc7348bis] as follows:

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	2	3	4	5	6	7	8	9
R R R R I R B R										Reserved																													
										VXLAN Network Identifier (VNI)										Reserved																			

This document requests IANA to allocate bit 6 of the VXLAN Flags registry for the "Ingress Replication BUM" flag. The chosen bit position aligns with that used in [I-D.ietf-nvo3-vxlan-gpe], facilitating data path implementations that support both VXLAN-GPE and the procedures defined in this document for VXLAN tunnels.

In an EVPN broadcast domain that uses Ingress Replication to forward BUM traffic:

1. Ingress PEs MUST set the Ingress Replication BUM flag in VXLAN packets that encapsulate broadcast, multicast, or unknown unicast frames.

2. Egress PEs MUST treat any received packet with the Ingress Replication BUM flag set as BUM traffic, regardless of the destination MAC address in the encapsulated frame.
3. Egress PEs MUST treat any received packet with the Ingress Replication BUM flag unset as known unicast traffic, regardless of the destination MAC address in the encapsulated frame.

In an EVPN broadcast domain that uses Assisted Replication [RFC9574], the ingress PE MUST set the Ingress Replication BUM flag in VXLAN packets that encapsulate broadcast, multicast or unknown unicast frames. In this context, the ingress PE may be an Assisted Replication Replicator, Leaf or RNVE (Regular Network Virtualization Edge) node.

The following section illustrates how the use of the Ingress Replication BUM flag addresses issues in EVPN multi-homing networks.

### 3. Use of the Ingress Replication BUM Flag in EVPN Multi-Homing

Consider the scenario in Figure 1, where PE1 and PE2 are multi-homed to CE1 in all-active mode, and ingress replication is used. Without the enhancement described in this document, when CE2 sends traffic to MAC M1, it is possible that PE3 has not yet learned M1, even though it already exists in the FIBs of PE1 and PE2. In this case, PE3 will ingress-replicate the frame in VXLAN packets to both PE1 and PE2, and since both egress PEs forward the traffic to CE1, packet duplication occurs. This is considered a transient scenario, but packet duplication may still be harmful for the applications.

Using the procedures in this document:

- \* In the same example PE3 sets the Ingress Replication BUM flag.
- \* PE1 and PE2 receive the VXLAN packet with the flag set and treat it as BUM traffic. As a result, only the Designated Forwarder (DF) forwards the frame to CE1, thereby preventing packet duplication.

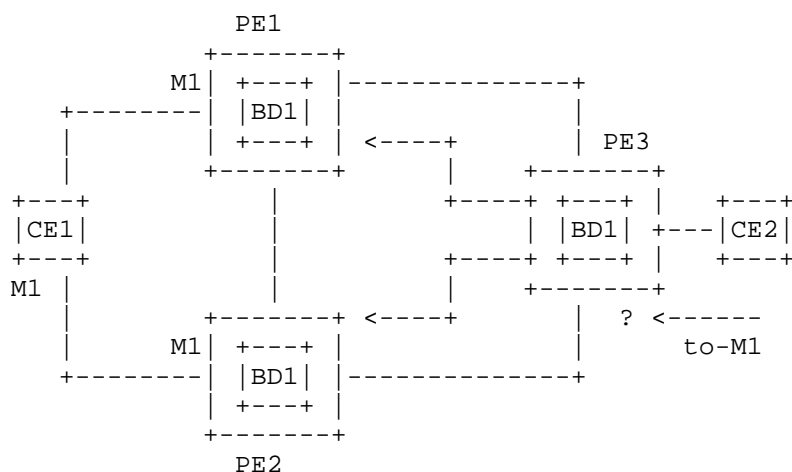


Figure 1: Packet Duplication

Figure 2 illustrates a similar example under different conditions. Without the enhancement described in this document, consider a scenario where CE2 sends a unicast frame destined for MAC M1. It is possible that PE3's FIB still contains an entry for M1 pointing to PE1, while PE1 has already removed M1 from its own FIB (due to MAC aging or other conditions). In this case, if PE1 is not the Designated Forwarder (DF) for the Ethernet Segment, it will discard the frame, disrupting communication between CE2 and CE1. This represents another transient condition that can negatively affect applications.

Using the procedures in this document:

- \* PE3 does not set the Ingress Replication BUM flag for the VXLAN packet that encapsulates the frames to M1.
- \* PE1 receives the VXLAN packet with the flag unset and therefore treats it as unicast traffic. This allows PE1 to safely forward the frame to CE1, and to any other local attachment circuits, if present. The communication between CE2 and CE1 is no longer disrupted.

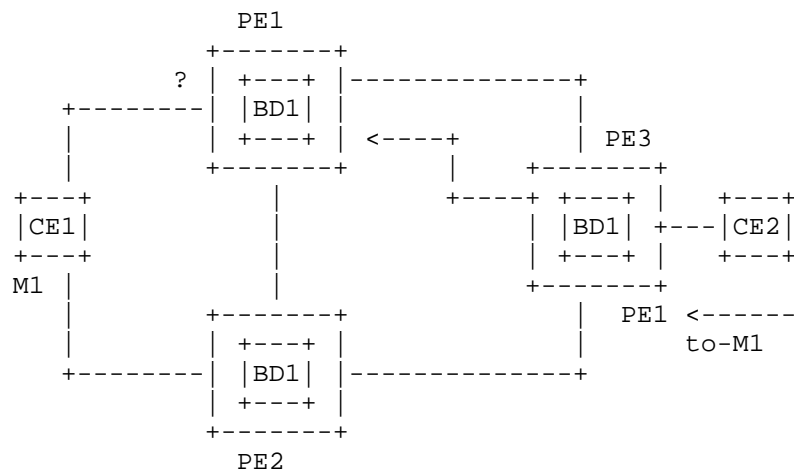


Figure 2: Packet Discard

4. Security Considerations

The use of the Ingress Replication BUM flag in VXLAN packets is not expected to introduce any new security risks in existing EVPN VXLAN networks. On the contrary, this flag provides a clear indication to receiving PEs whether the traffic was originally classified as BUM or unicast. Since the setting of this bit for BUM-encapsulated packets is automatic and not controlled by configuration commands, an attacker with access to an EVPN VXLAN PE cannot exploit or modify it as a means of attack.

5. IANA Considerations

This document requests a new Flag in the registry called "VxLAN Flags", to indicate "Ingress Replication BUM".

Bit	Name	Reference
6	Ingress Replication BUM	This document

Figure 3

6. Contributors

7. Acknowledgements

The authors would like to acknowledge the authors of [RFC8365] and [I-D.ietf-nvo3-vxlan-gpe] for their initial discussions about the use of an Ingress Replication BUM flag in VXLAN-GPE.

## 8. References

### 8.1. Normative References

- [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>>.
- [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>>.
- [RFC8365] Sajassi, A., Ed., Drake, J., Ed., Bitar, N., Shekhar, R., Uttaro, J., and W. Henderickx, "A Network Virtualization Overlay Solution Using Ethernet VPN (EVPN)", RFC 8365, DOI 10.17487/RFC8365, March 2018, <<https://www.rfc-editor.org/info/rfc8365>>.
- [RFC9574] Rabadan, J., Ed., Sathappan, S., Lin, W., Katiyar, M., and A. Sajassi, "Optimized Ingress Replication Solution for Ethernet VPNs (EVPNs)", RFC 9574, DOI 10.17487/RFC9574, May 2024, <<https://www.rfc-editor.org/info/rfc9574>>.
- [I-D.ietf-nvo3-rfc7348bis] Mahalingam, M., Dutt, D., Kreeger, L., Sridhar, T., and A. Sajassi, "Virtual eXtensible Local Area Network (VXLAN): A Framework for Overlaying Virtualized Layer 2 Networks over Layer 3 Networks", Work in Progress, Internet-Draft, draft-ietf-nvo3-rfc7348bis-01, 23 July 2025, <<https://datatracker.ietf.org/doc/html/draft-ietf-nvo3-rfc7348bis-01>>.

### 8.2. Informative References

- [I-D.ietf-nvo3-vxlan-gpe] Maino, F., Kreeger, L., and U. Elzur, "Generic Protocol Extension for VXLAN (VXLAN-GPE)", Work in Progress, Internet-Draft, draft-ietf-nvo3-vxlan-gpe-13, 4 November 2023, <<https://datatracker.ietf.org/doc/html/draft-ietf-nvo3-vxlan-gpe-13>>.

### Authors' Addresses

Jorge Rabadan (editor)  
Nokia  
520 Almanor Avenue  
Sunnyvale, CA 94085  
United States of America  
Email: [jorge.rabadan@nokia.com](mailto:jorge.rabadan@nokia.com)

Ali Sajassi  
Cisco  
Email: [sajassi@cisco.com](mailto:sajassi@cisco.com)