Network Working Group
Internet-Draft
Intended status: Informational
Expires: April 4, 2014

P. Quinn
Cisco Systems, Inc.
P. Agarwal
Broadcom
R. Fernando
L. Kreeger
D. Lewis
F. Maino
Cisco Systems, Inc.
M. Smith
N. Yadav
Insieme Networks
October 1, 2013

Generic Protocol Extension for VXLAN
draft-quinn-vxlan-gpe-01.txt

Abstract

This draft describes a mechanism for adding multi-protocol support to Virtual eXtensible Local Area Network (VXLAN). Protocol identification is carried in the VXLAN header and is used to describe the encapsulated payload.

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 April 4, 2014.

Table of Contents

1. Introduction

Virtual eXtensible Local Area Network [VXLAN] defines an encapsulation format that encapsulates Ethernet frames in an outer UDP/IP transport. The VXLAN header does not specify the protocol being encapsulated and therefore is currently limited to encapsulating only Ethernet frame payloads. As data centers evolve, the need to carry other protocols in an encapsulated IP packet is required. Rather than defining yet another encapsulation, VXLAN can be extended to indicate the inner protocol, thus broadening the applicability of VXLAN.

This document describes extending VXLAN to support additional payload types beyond Ethernet frames. To support this capability, two elements of the existing VXLAN header are modified.

1. A reserved bit is allocated, and set in the VXLAN header.

2. A 16 bit Protocol Type field is present in the VXLAN header.

These two changes allow for the VXLAN header to support many different types of payloads, all the while maintaining backward compatibility with existing VXLAN deployments.

2. VXLAN Without Protocol Extension

As described in the introduction, the VXLAN header has no protocol identifier that indicates the type of payload being carried by VXLAN. Because of this, VXLAN is limited to an Ethernet payload.

The VXLAN header defines flags (some defined, some reserved), the VXLAN network identifier (VNI) field and several reserved bits. The flags provide flexibility to define how the reserved bits can be used to change the definition of the VXLAN header.

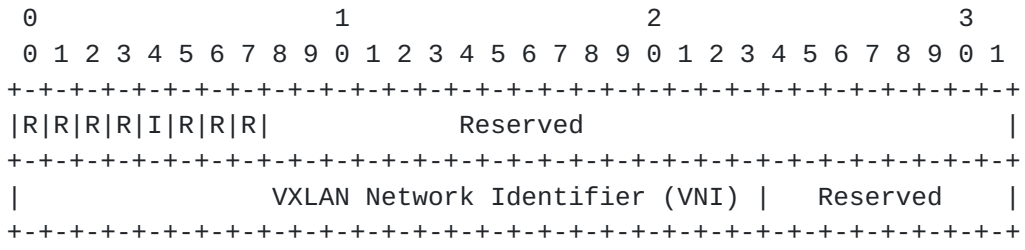

Figure 1: VXLAN Header

3. Generic Protocol Extension VXLAN (VXLAN-gpe)

3.1. VXLAN Header

This draft defines two changes to the VXLAN header in order to support multi-protocol encapsulation.

P Bit: Flag bit 5 is defined as the P bit. The P bit MUST be set to 1 to indicate the presence of the 16 bit protocol type field in the lower 16 bits of the first word.

P = 0 indicates that the payload MUST conform to VXLAN as defined in [VXLAN].

Flag bit 5 was chosen as the P bit because this flag bit is currently reserved in VXLAN.

Protocol Type Field: The lower 16 bits of the first word are used to carry a protocol type. This protocol type field contains the protocol, as defined in in [RFC1700] and in [ETYPES], of the encapsulated payload packet.

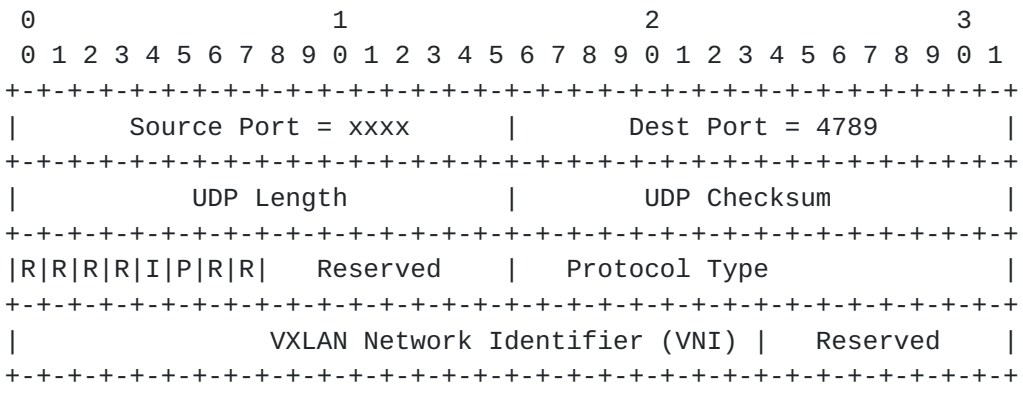

Figure 2: UDP + VXLAN-gpe

4. Backward Compatibility

In order to ensure compatibility with existing VXLAN deployments, P = 0 indicates that the encapsulated payload MUST be Ethernet.

4.1. VXLAN VTEP to VXLAN-gpe VTEP

If a packet is sent from a VXLAN VTEP to a VXLAN-gpe VTEP, the P bit MUST be set to 0, and the remaining fields remain as described in [VXLAN]. The encapsulated payload MUST be Ethernet.

4.2. VXLAN-gpe VTEP to VXLAN VTEP

A VXLAN-gpe VTEP MUST not encapsulate non-Ethernet frames to a VXLAN VTEP. When encapsulating Ethernet frames to a VXLAN VTEP the P bit will be set to 1 and the Protocol Type set to 0x6558. The VXLAN VTEP will ignore the P bit and the Protocol Type, and treat the packet as a VXLAN packet (i.e. the payload is Ethernet)

A method for determining the capabilities of a VXLAN VTEP (gpe or non-gpe) is out of the scope of this draft.

4.3. IP Type of Service/Traffic Class

When a VXLAN-gpe VTEP performs IPv4 encapsulation, the inner IPv4 Type of Service field MAY be copied from the encapsulated packet to the Type of Service or Traffic Class field in the outer IPv4 or IPv6 header respectively.

Similarly, when a VXLAN-gpe VTEP performs IPv6 encapsulation, the inner IPv6 Traffic Class field MAY be copied from the encapsulated packet to the Type of Service or Traffic Class field in the outer IPv4 or IPv6 header respectively.

5. VXLAN-gpe Examples

This section provides three examples of protocols encapsulated using the Generic Protocol Extension for VXLAN described in this document.

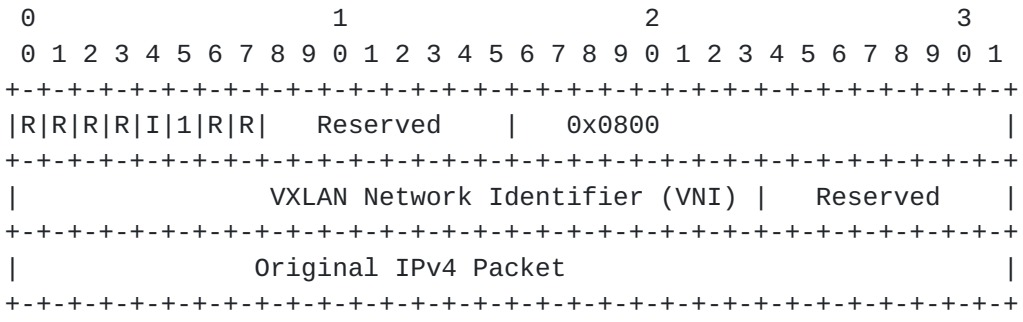

Figure 3: IPv4 and VXLAN-gpe

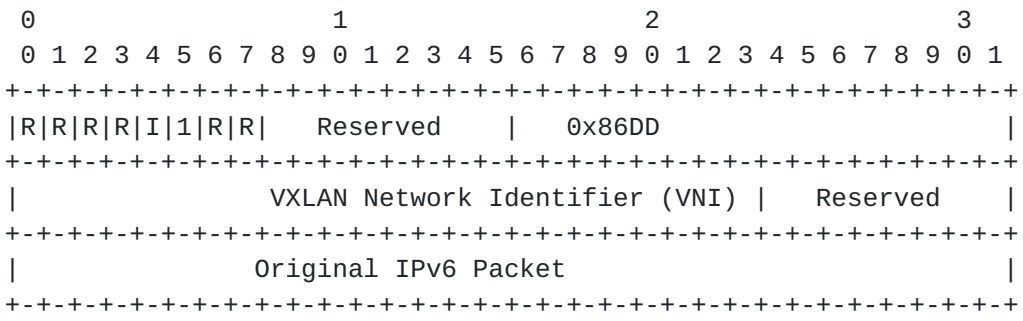

Figure 4: IPv6 and VXLAN-gpe

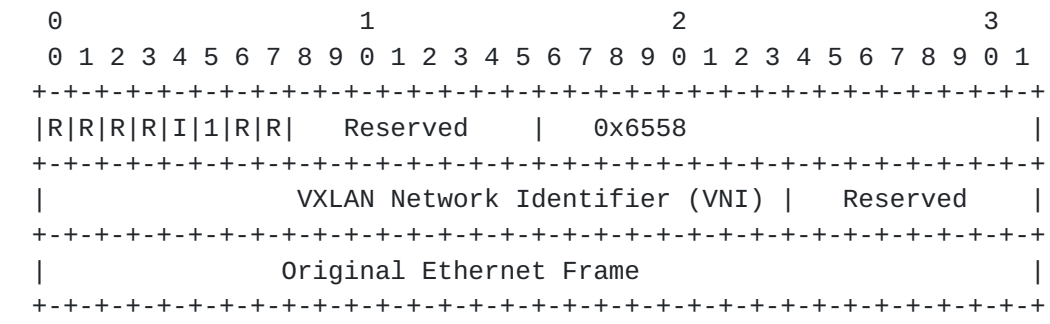

Figure 5: Ethernet and VXLAN-gpe

6. Security Considerations

VXLAN's security is focused on issues around L2 encapsulation into L3. With VXLAN-gpe, issues such as spoofing, flooding, and traffic redirection are dependent on the particular protocol payload encapsulated.

7. Acknowledgments

A special thank you goes to Dino Farinacci for his guidance and detailed review.

8. IANA Considerations

This document creates no new requirements on IANA namespaces [RFC5226].

9. References

9.1. Normative References

[RFC0768] Postel, J., "User Datagram Protocol", STD 6, RFC 768, August 1980.

[RFC0791] Postel, J., "Internet Protocol", STD 5, RFC 791, September 1981.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC5226] Narten, T. and H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 5226, May 2008.

9.2. Informative References

[ETYPES] The IEEE Registration Authority, "IEEE 802 Numbers", 2012, <http://www.iana.org/assignments/ieee-802-numbers/ieee-802-numbers.xml>.

[RFC1700] Reynolds, J. and J. Postel, "Assigned Numbers", RFC 1700, October 1994.

[VXLAN] Dutt, D., Mahalingam, M., Duda, K., Agarwal, P., Kreeger, L., Sridhar, T., Bursell, M., and C. Wright, "VXLAN: A Framework for Overlaying Virtualized Layer 2 Networks over Layer 3 Networks", 2013.

Authors' Addresses

Paul Quinn
Cisco Systems, Inc.

Email: paulq@cisco.com

Puneet Agarwal
Broadcom

Email: pagarwal@broadcom.com

Rex Fernando
Cisco Systems, Inc.

Email: rex@cisco.com

Larry Kreeger
Cisco Systems, Inc.

Email: kreeger@cisco.com

Darrel Lewis
Cisco Systems, Inc.

Email: darlewis@cisco.com

Fabio Maino
Cisco Systems, Inc.

Email: kreeger@cisco.com

Michael Smith
Insieme Networks

Email: michsmit@insiemenetworks.com

Navindra Yadav
Insieme Networks

Email: nyadav@insiemenetworks.com