Network Working Group
Internet-Draft
Intended status: Experimental
Expires: January 4, 2015

P. Quinn
Cisco Systems, Inc.
P. Agarwal
Broadcom
R. Fernando
L. Kreeger
D. Lewis
F. Maino
M. Smith
N. Yadav
Cisco Systems, Inc.
L. Yong
Huawei USA
X. Xu
Huawei Technologies
U. Elzur
Intel
P. Garg
Microsoft

July 3, 2014

# Generic Protocol Extension for VXLAN
## draft-quinn-vxlan-gpe-03.txt

Abstract

This draft describes extending Virtual eXtensible Local Area Network (VXLAN), via changes to the VXLAN header, with three new capabilities: support for multi-protocol encapsulation, operations, administration and management (OAM) signaling and explicit versioning.

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 January 4, 2015.

## Copyright Notice

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. As data centers evolve, the need to carry other protocols encapsulated in an IP packet is required, as well as the need to provide increased visibility and diagnostic capabilities within the overlay. The VXLAN header does not specify the protocol being encapsulated and therefore is currently limited to encapsulating only Ethernet frame payload, nor does it provide the ability to define OAM protocols. Rather than defining yet another encapsulation, VXLAN is extended to provide protocol typing and OAM capabilities.

This document describes extending VXLAN via the following changes:

Next Protocol Bit (P bit): A reserved flag bit is allocated, and set in the VXLAN-gpe header to indicate that a next protocol field is present.

OAM Flag Bit (O bit): A reserved flag bit is allocated, and set in the VXLAN-gpe header, to indicate that the packet is an OAM packet.

Version: Two reserved bits are allocated, and set in the VXLAN-gpe header, to indicate VXLAN-gpe protocol version.

Next Protocol: A 8 bit next protocol field is present in the VXLAN-gpe header.

## 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. Furthermore, the VXLAN header has no mechanism to signal OAM packets.

The VXLAN header defines bits 0-7 as 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.

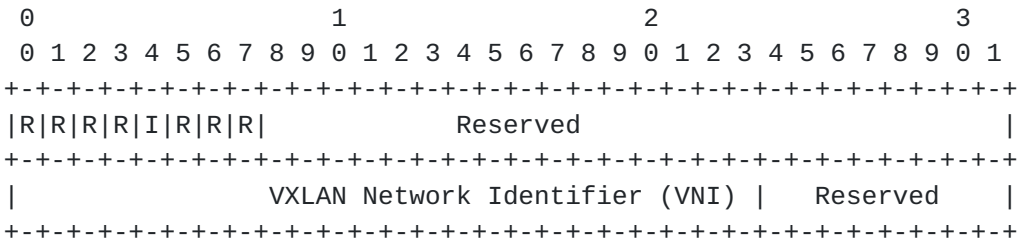


Figure 1: VXLAN Header

## 3. Generic Protocol Extension VXLAN (VXLAN-gpe)

### 3.1. Multi Protocol Support

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

P Bit: Flag bit 5 is defined as the Next Protocol bit. The P bit MUST be set to 1 to indicate the presence of the 8 bit next protocol field.

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.

Next Protocol Field: The lower 8 bits of the first word are used to carry a next protocol. This next protocol field contains the protocol of the encapsulated payload packet. A new protocol registry will be requested from IANA.

This draft defines the following Next Protocol values:

0x1 : IPv4
0x2 : IPv6
0x3 : Ethernet
0x4 : Network Service Header [NSH]

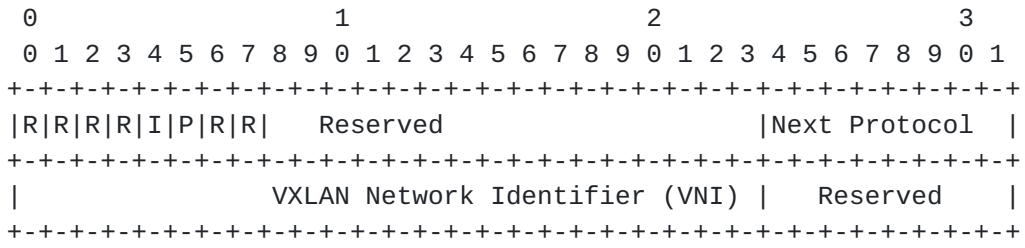


Figure 2: VXLAN-gpe Next Protocol

### 3.2. OAM Support

Flag bit 7 is defined as the O bit. When the O bit is set to 1, the packet is an OAM packet and OAM processing MUST occur. The OAM protocol details are out of scope for this document. As with the P-bit, bit 7 is currently a reserved flag in VXLAN.

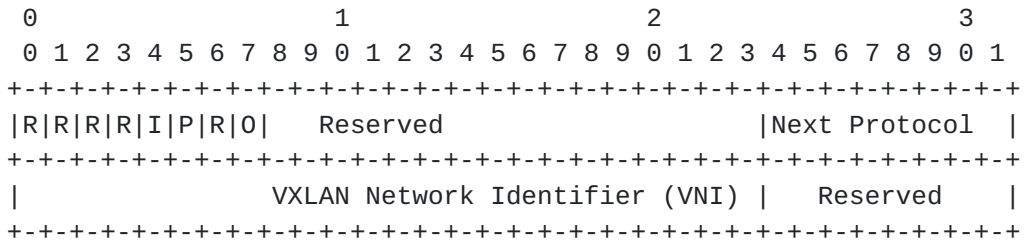


Figure 3: VXLAN-gpe OAM Bit

### 3.3. Version Bits

VXLAN-gpe bits 8 and 9 are defined as version bits. These bits are reserved in VXLAN. The version field is used to ensure backward compatibility going forward with future VXLAN-gpe updates.

The initial version for VXLAN-gpe is 0.

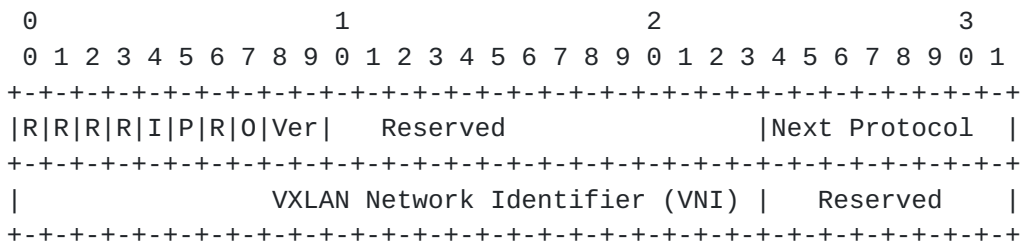


Figure 4: VXLAN-gpe Version Bits

## 4. Backward Compatibility

### 4.1. VXLAN VTEP to VXLAN-gpe VTEP

As per VXLAN, reserved bits 5 and 7, VXLAN-gpe P and O-bits respectively must be set to zero. The remaining reserved bits must be zero, including the VXLAN-gpe version field, bits 8 and 9. 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 VXLAN-gpe VTEP will set the P bit to 0, the Next Protocol to 0 and use UDP destination port 4789. A VXLAN-gpe VTEP MUST also set O = 0 and Ver = 0 when encapsulating Ethernet frames to VXLAN VTEP. The receiving VXLAN VTEP will threat this packet as a VXLAN packet.

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

### 4.3. VXLAN-gpe UDP Ports

VXLAN-gpe uses a new UDP destination port (to be assigned by IANA) when sending traffic to VXLAN-gpe VTEPs.

### 4.4. VXLAN-gpe and Encapsulated IP Header Fields

When encapsulating and decapsulating IPv4 and IPv6 packets, certain fields, such as IPv4 Time to Live (TTL) from the inner IP header need to be considered. VXLAN-gpe IP encapsulation and decapsulation utilizes the techniques described in [RFC6830], section 5.3.

## 5. VXLAN-gpe Examples

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

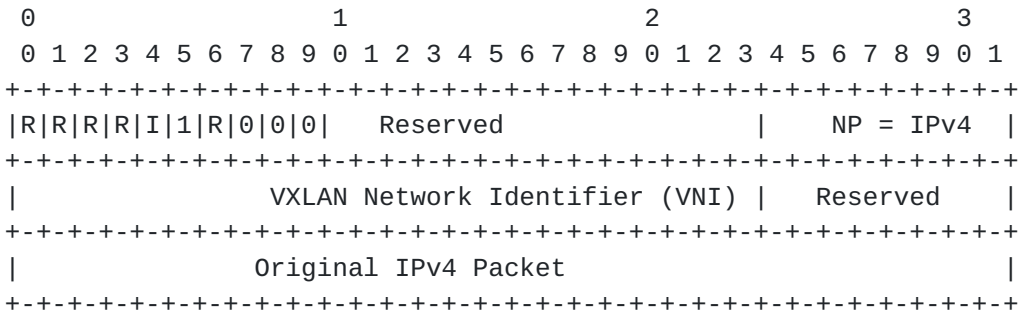


Figure 5: IPv4 and VXLAN-gpe

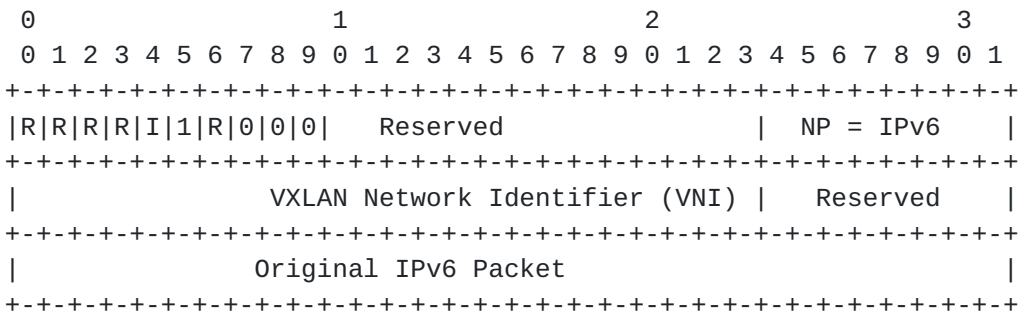


Figure 6: IPv6 and VXLAN-gpe

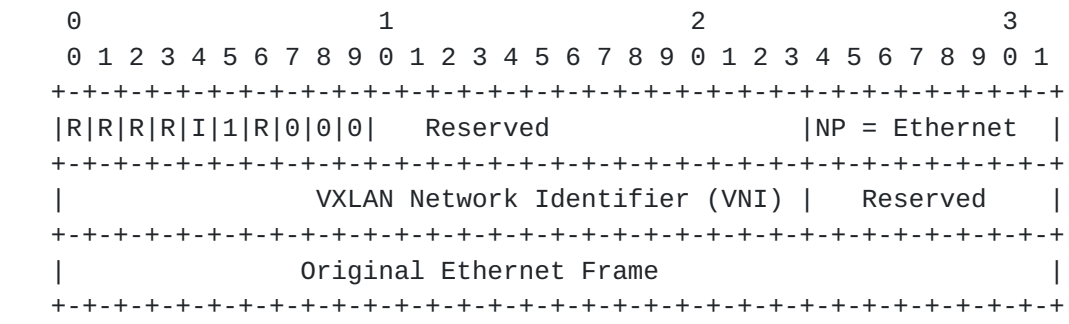


Figure 7: Ethernet and VXLAN-gpe

## 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

### 8.1. UDP Port

A new UDP port will be requested from IANA.

### 8.2. VXLAN-gpe Next Protocol

IANA is requested to set up a registry of "Next Protocol". These are 8-bit values. Next Protocol values 0, 1, 2, 3 and 4 are defined in this draft. New values are assigned via Standards Action [RFC5226].

| Next Protocol | Description | Reference |
|---|---|---|
| 0 | Reserved | This document |
| 1 | IPv4 | This document |
| 2 | IPv6 | This document |
| 3 | Ethernet | This document |
| 4 | NSH | This document |
| 5..253 | Unassigned | |

Table 1

### 8.3. VXLAN-gpe Reserved Bits

There are ten bits at the beginning of the VXLAN-gpe header. New bits are assigned via Standards Action [RFC5226].

Bits 0-3 - Reserved
Bit 4 - Instance ID (I bit)
Bit 5 - Next Protocol (P bit)
Bit 6 - Reserved
Bit 7 - OAM (O bit)
Bits 8-9 - Version

# 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

[NSH] Quinn, P. and et al. , "Network Service Header", 2014.

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

[RFC6830] Farinacci, D., Fuller, V., Meyer, D., and D. Lewis, "The Locator/ID Separation Protocol (LISP)", RFC 6830, January 2013.

[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
Cisco Systems, Inc.

Email: michsmit@cisco.com

Navindra Yadav
Cisco Systems, Inc.

Email: nyadav@cisco.com

Lucy Yong
Huawei USA

Email: lucy.yong@huawei.com

Xiaohu Xu
Huawei Technologies

Email: xuxiaohu@huawei.com

Uri Elzur
Intel

Email: uri.elzur@intel.com

Pankaj Garg
Microsoft

Email: Garg.Pankaj@microsoft.com