Generic Protocol Extension for VXLAN
draft-quinn-vxlan-gpe-02.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 June 19, 2014.

Copyright Notice
Copyright (c) 2013 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 (http://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.

Table of Contents

1. Introduction .................................................. 3
2. VXLAN Without Protocol Extension ................................ 4
   3.1. VXLAN Header ............................................. 5
4. Backward Compatibility ........................................... 6
   4.1. VXLAN VTEP to VXLAN-gpe VTEP ................................ 6
   4.2. VXLAN-gpe VTEP to VXLAN VTEP ................................ 6
5. VXLAN-gpe and Encapsulated IP Header Fields ....................... 7
6. VXLAN-gpe Examples ............................................. 8
7. Security Considerations .......................................... 10
8. Acknowledgments .................................................. 11
9. IANA Considerations ............................................. 12
10. References ........................................................ 13
   10.1. Normative References ...................................... 13
   10.2. Informative References .................................... 13
Authors’ Addresses .................................................. 14
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 encapsulated in an 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. For IPv4/v6 payloads, this document also specifies expected behavior for handling certain inner IP header fields.

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 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.

```
+-----------------+-----------------+-----------------+-----------------+
| R | R | R | R | I | R | R | R | Reserved |
+-----------------+-----------------+-----------------+-----------------+
|                | VXLAN Network Identifier (VNI) | Reserved |
+-----------------+-----------------+-----------------+-----------------+
```

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.

VXLAN-gpe does not impact the UDP header; more specifically the destination port is 4789 as defined in [VXLAN].

```
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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|R|R|R|R|I|P|R|R|   Reserved    |   Protocol Type               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                VXLAN Network Identifier (VNI) |   Reserved    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 2: 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 is 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 VXLAN-gpe VTEP will set the P bit to 1 and the Protocol Type 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.
5. 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.
6. 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
7. 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.
8. Acknowledgments

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

Note that the contributors to this document are listed in alphabetical order according to their organizational affiliation.
9. IANA Considerations

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

10.1. Normative References


10.2. Informative References


Authors’ Addresses

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

Paul Quinn
Cisco Systems, Inc.
Email: paulq@cisco.com

Lucy Yong
Huawei USA
Email: lucy.yong@huawei.com
Xiaohu Xu  
Huawei Technologies  
Email: xuxiaohu@huawei.com

Michael Smith  
Insieme Networks  
Email: michsmit@insiemenetworks.com

Navindra Yadav  
Insieme Networks  
Email: nyadav@insiemenetworks.com

Uri Elzur  
Intel  
Email: uri.elzur@intel.com