Network working group L. Yong
Internet Draft X. Xu
Category: Standard Track Huawei

Expires: April 2014 October 17, 2013

NVGRE and VXLAN Encapsulation Extension for L3 Overlay
draft-yong-l3vpn-nvgre-vxlan-encap-03

Status of this Memo

This Internet-Draft is submitted to IETF in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

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."

The list of current Internet-Drafts can be accessed at http://www.ietf.org/ietf/1id-abstracts.txt.

The list of Internet-Draft Shadow Directories can be accessed at http://www.ietf.org/shadow.html.

This Internet-Draft will expire on April 17, 2014.

Copyright Notice

Abstract

Both NVGRE and VXLAN encapsulations were originally designed for L2 overlay only. This draft proposes the enhancement on both to support L3 overlay as well. The proposed method completely decouples the L3 overlay from the L2 overlay in terms of encoding schema and data processing.

Conventions used in this document

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC-2119 [RFC2119].

Table of Contents

1. Introduction

Network Virtualization Overlay [NVO3FRWK] explicitly states that both L2 and L3 overlays are needed in practice. However both NVGRE encapsulation [NVGRE] and VXLAN encapsulation [VXLAN] were originally designed for L2 overlay only.

This document proposes enhancements to NVGRE and VXLAN encapsulations to allow the same data encapsulation semantics for both L2 overlay and L3 overlay. The benefits of this approach are generalizing the data encapsulation semantics for overlay technologies, maintaining L3 overlay natively, and decoupling it from L2 overlay completely.

2. NVGRE Encapsulation Extension for L3 Overlay

NVGER [NVGRE] leverages the GRE protocol [RFC2890] and specifies that the protocol type field in the GRE header MUST be filled with the value of 0x6558, which means for Transparent Ethernet.

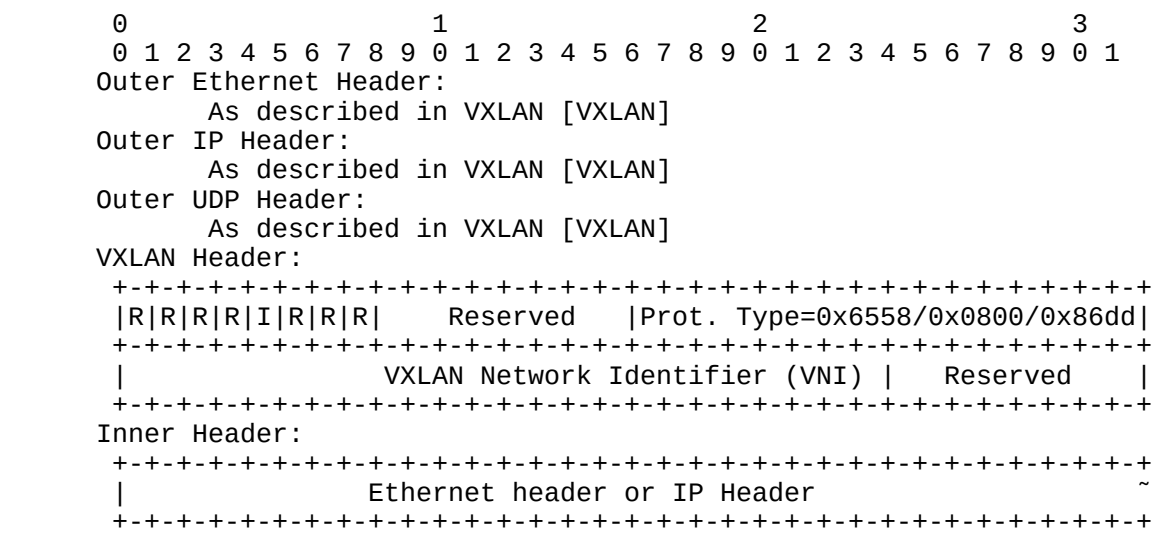
This document proposes the protocol type field to be filled with the value of 0x6558, 0x0800(IPv4), or 0x86dd(IPv6). The value of 0x0800 and 0x86dd means that the payload is IP. The value 0x6558 MUST be used if the inner header is an Ethernet header. When NVGRE encapsulation is used for L3 overlay, it MUST use the value of 0x0800 or 0x86dd in the protocol type field and MUST encode an IPv4 or IPv6 header as the inner header. Other fields in the outer header and the GRE header remain the same.

To support backward compatibility, when the remote tunnel end point only support the NVGRE described in [NVGRE], the tunnel end point that supports NVGRE described in this document MUST only encapsulate L2 packets. This capability can be either manually configured or be dynamically informed. How tunnel end points inform each other the encapsulation capabilities is beyond the scope of this document. Note that a tunnel may have more than two end points.

3. VXLAN Encapsulation Extension for L3 Overlay

This document proposes adding a protocol type field in the VXLAN header as shown below. It takes 16 bits from the reserved 24 bits as the protocol type field. The remained 8 reserved bits MUST be filled with zero. For L2 overlay encapsulation, the protocol type field MUST be filled with the value of 0x6558 and inner header MUST be an Ethernet header. For L3 overlay encapsulation, the protocol type

field MUST be filled with the value of 0x0800(IPv4) or 0x86dd(IPv6), and inner header MUST be an IPv4 or IPv6 header. Other fields in the outer header and VXLAN header remain the same.



To be backward compatible with the existing VXLAN encapsulation [VXLAN], the value 0x0000 in the Protocol Type field MUST be treated as Ethernet payload too. When the end points of a tunnel support different VXLAN formats, i.e. one, say A, supports old VXLAN format and another, say B, supports the new format described in this document, B MUST only encapsulate L2 packets and set value 0x0000 in the protocol type field. This capability can be either manually configured at B or be dynamically informed. How tunnel end points inform each other the encapsulation capabilities is beyond the scope of this document. Note that a tunnel may have more than two end points.

Having protocol type field in the VXLAN header enables other overlay payload type beside L2 and L3 overlays. The application for other payload type is for future study.

4. Security Considerations

The mechanism proposed in this document does not add any additional security concern beside what has been described in the NVGRE [NVGRE] and VXLAN [VXLAN].

5. IANA Considerations

The document does not require any IANA action.

6. References

6.1. Normative References

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

[RFC2890] Dommety, G., "Key and Sequence Number Extension to GRE", RFC2890, September 2000

6.2. Informative References

[NVO3FRWK] Lasserre, M., et al, "Framework for DC Network Virtualization", draft-ietf-nvo3-framework-03.txt, work in progress.

[NVGRE] Sridharan, M., et al, "NVGRE: Network Virtualization using Generic Routing Encapsulation", draft-sridharan-virtualization-nvgre-03, work in progress

[VXLAN] Mahalingam, M., Dutt, D., etc, "VXLAN: A Framework for Overlaying Virtualized Layer 2 Networks over Layer 3 Networks", draft-mahalingam-dutt-dcops-vxlan-05.txt, work in progress

Authors' Addresses

Lucy Yong
Huawei Technologies, USA

Phone: 918-808-1918
Email: lucy.yong@huawei.com

Xiaohu Xu
Huawei Technologies,
Beijing, China

Phone: +86-10-60610041
Email: xuxiaohu@huawei.com