

Network Working Group  
Internet-Draft  
Intended status: Standards Track  
Expires: January 2, 2017

X. Xu  
Huawei  
H. Shah  
Ciena Corp  
Y. Fan  
China Telecom  
July 1, 2016

NVo3 Control Plane Protocol Using IS-IS  
draft-xu-nvo3-isis-cp-02

Abstract

This document describes the use of IS-IS as a light-weight control plane protocol for Network Virtualization over L3 (NVo3) overlay networks. This light-weight control plane protocol is intended for small and even medium sized enterprise campus networks where the NVo3 technology is to be used.

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 2, 2017.

Copyright Notice

Copyright (c) 2016 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 . . . . .	2
1.1. Requirements Language . . . . .	3
2. Terminology . . . . .	3
3. VN Membership Auto-discovery . . . . .	3
3.1. VN Membership Info Sub-TLV . . . . .	3
4. Tunnel Encapsulation Capability Advertisement . . . . .	4
5. MAC Address Learning . . . . .	5
5.1. Control-plane based MAC Learning for Remote CE Hosts . . . . .	5
6. IANA Considerations . . . . .	5
7. Security Considerations . . . . .	5
8. Acknowledgements . . . . .	6
9. References . . . . .	6
9.1. Normative References . . . . .	6
9.2. Informative References . . . . .	6
Authors' Addresses . . . . .	7

## 1. Introduction

[RFC7364] discusses the need of an overlay-based network virtualization approach, referred to as Network Virtualization over Layer3 (NVo3), for providing multi-tenancy capabilities in large data centers networks and outlines the needs for a control plane protocol to facilitate running these NVo3 overlay networks. [RFC7365] provides a framework for NVo3 overlay networks and meanwhile describes the needs for a control plane protocol to provide the following capabilities such as auto-provisioning/service discovery, address mapping advertisement and tunnel management.

Due to the success of the NVo3 technology in data center networks, more and more enterprises are considering the deployment of this technology in their campus networks so as to replace the old spanning tree protocols. Although BGP or Software Defined Network (SDN) controller could still be used as the control plane protocol in campus networks, both of them seem a bit heavyweight, especially for small and even medium sized campus networks.

IS-IS protocol [IS-IS] is a much proven and well-known routing protocol which has been widely deployed in campus networks for many years. Due to its extendibility, IS-IS protocol now is not only used for propagating IP reachability information in Layer3 networks (see [RFC1195]), but also used for propagating MAC reachability information in Layer2 networks or Layer2 overlay networks [RFC6165].

By using IS-IS as a lightweight control plane protocol for NVo3 overlay networks, the network provisioning is greatly simplified ((e.g., only a single protocol to be deployed)), which is much significant to campus networks.

This IS-IS based NVo3 control plane protocol could support any specific NVo3 data encapsulation formats such as VXLAN [RFC7348], VXLAN-GPE [I-D.ietf-nvo3-vxlan-gpe] , and NVGRE [RFC7637].

### 1.1. Requirements Language

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].

## 2. Terminology

This memo makes use of the terms defined in [RFC7365] and [I-D.ietf-bier-architecture].

## 3. VN Membership Auto-discovery

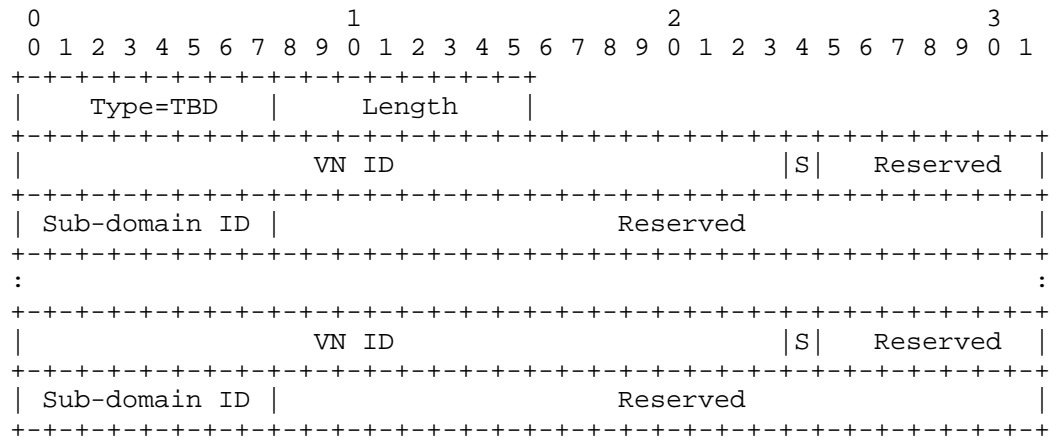
By propagating the VN membership info among Network Virtualization Edges (NVEs), NVEs belonging to the same VN instance could discover one another automatically. The VN membership info is carried in a VN Membership Info sub-TLV (as shown in Section 3.1) of the following TLVs originated by that NVE:

1. TLV-135 (IPv4) defined in [RFC5305].
2. TLV-236 (IPv6) defined in [RFC5308]

When the above TLV is propagated across level boundaries, the VN Membership Info sub-TLV contained in that TLV SHOULD be kept.

### 3.1. VN Membership Info Sub-TLV

The VN Membership Info sub-TLV has the following format:



Type: TBD;

Length: Variable;

VN ID: This field is filled with a 24-bit globally significant VN ID for a particular attached VN instance.

S-Flag: This field indicates the existence of the Sub-domain ID field. When the S-Flag is set, the Sub-domain ID field MUST be filled with a valid sub-domain ID. Otherwise, it SHOULD be set to zero.

Sub-domain ID: This field is filled with a 8-bit BIER sub-domain ID to which the VN has been associated [I-D.ietf-bier-architecture]. The field is only useful in the case where the Broadcast, Unknown-unicast and Multicast (BUM) packets within a VN are transported across the underlay by using the BIER forwarding mode.

#### 4. Tunnel Encapsulation Capability Advertisement

To reach a consensus on what specific tunnel encapsulation format to be used between ingress and egress NVE pairs automatically, egress NVEs SHOULD advertise their own tunnel encapsulation capabilities by using the Encapsulation Capability sub-TLV as defined in [I-D.xu-isis-encapsulation-cap]

## 5. MAC Address Learning

MAC addresses of local CE hosts would still be learnt by NVEs as normal bridges. As for learning MAC addresses of remote CE hosts, there are two options: 1) data-plane based MAC learning and 2) control-plane based MAC learning. If unknown unicast flood suppression is strongly required even at the cost of consuming more forwarding table resources, the control-plane based MAC learning option could be considered. Otherwise, the data-plane based MAC learning option is RECOMMENDED.

### 5.1. Control-plane based MAC Learning for Remote CE Hosts

In the control-plane based MAC address learning mechanism, MAC reachability information of a given VN instance would be exchanged across NVEs of that VN instance via IS-IS as well. Upon learning MAC addresses of their local TES's somehow, NVEs SHOULD immediately advertise these MAC addresses to remote NVEs of the same VN instance by using the MAC-Reachability TLV as defined in [RFC6165]. One or more MAC-Reachability TLVs are carried in a LSP which in turn is encapsulated with an Ethernet header. The source MAC address is the originating NVE's MAC address whereas the destination MAC address is a to-be-defined multicast MAC address specifically identifying all NVEs. Such Ethernet frames containing IS-IS LSPs are forwarded towards remote NVEs as if they were customer multicast Ethernet frames. Egress NVEs receiving the above frames SHOULD intercept them and accordingly process them. The routable IP address of the NVE originating these MAC routes could be derived either from the "IP Interface Address" field contained in the corresponding LSPs (Note that the IP address here SHOULD be identical to the routable IP address associated with the VN membership Info) or from the tunnel source IP address of the NVo3 encapsulated packet containing such MAC routes. Since these LSPs are fully transparent to core routers of the underlying networks (i.e., non-NVE routers), there is no impact on the control plane of core routers at all.

## 6. IANA Considerations

The type code for VN Membership Info sub-TLV is required to be allocated by IANA.

## 7. Security Considerations

This document doesn't introduce additional security risk to IS-IS, nor does it provide any additional security feature for IS-IS.

## 8. Acknowledgements

TBD

## 9. References

### 9.1. 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, <<http://www.rfc-editor.org/info/rfc2119>>.
- [RFC4971] Vasseur, JP., Ed., Shen, N., Ed., and R. Aggarwal, Ed., "Intermediate System to Intermediate System (IS-IS) Extensions for Advertising Router Information", RFC 4971, DOI 10.17487/RFC4971, July 2007, <<http://www.rfc-editor.org/info/rfc4971>>.
- [RFC5305] Li, T. and H. Smit, "IS-IS Extensions for Traffic Engineering", RFC 5305, DOI 10.17487/RFC5305, October 2008, <<http://www.rfc-editor.org/info/rfc5305>>.
- [RFC5308] Hopps, C., "Routing IPv6 with IS-IS", RFC 5308, DOI 10.17487/RFC5308, October 2008, <<http://www.rfc-editor.org/info/rfc5308>>.

### 9.2. Informative References

- [I-D.ietf-bier-architecture] Wijnands, I., Rosen, E., Dolganow, A., Przygienda, T., and S. Aldrin, "Multicast using Bit Index Explicit Replication", draft-ietf-bier-architecture-03 (work in progress), January 2016.
- [I-D.ietf-nvo3-vxlan-gpe] Kreeger, L. and U. Elzur, "Generic Protocol Extension for VXLAN", draft-ietf-nvo3-vxlan-gpe-02 (work in progress), April 2016.
- [I-D.xu-isis-encapsulation-cap] Xu, X., Decraene, B., Raszuk, R., Chunduri, U., Contreras, L., and L. Jalil, "Advertising Tunnelling Capability in IS-IS", draft-xu-isis-encapsulation-cap-06 (work in progress), November 2015.

- [IS-IS] "ISO/IEC 10589, "Intermediate System to Intermediate System Intra-Domain Routing Exchange Protocol for use in Conjunction with the Protocol for Providing the Connectionless-mode Network Service (ISO 8473)", 2005."
- [RFC1195] Callon, R., "Use of OSI IS-IS for routing in TCP/IP and dual environments", RFC 1195, DOI 10.17487/RFC1195, December 1990, <<http://www.rfc-editor.org/info/rfc1195>>.
- [RFC6165] Banerjee, A. and D. Ward, "Extensions to IS-IS for Layer-2 Systems", RFC 6165, DOI 10.17487/RFC6165, April 2011, <<http://www.rfc-editor.org/info/rfc6165>>.
- [RFC7348] Mahalingam, M., Dutt, D., Duda, K., Agarwal, P., Kreeger, L., Sridhar, T., Bursell, M., and C. Wright, "Virtual eXtensible Local Area Network (VXLAN): A Framework for Overlaying Virtualized Layer 2 Networks over Layer 3 Networks", RFC 7348, DOI 10.17487/RFC7348, August 2014, <<http://www.rfc-editor.org/info/rfc7348>>.
- [RFC7364] Narten, T., Ed., Gray, E., Ed., Black, D., Fang, L., Kreeger, L., and M. Napierala, "Problem Statement: Overlays for Network Virtualization", RFC 7364, DOI 10.17487/RFC7364, October 2014, <<http://www.rfc-editor.org/info/rfc7364>>.
- [RFC7365] Lasserre, M., Balus, F., Morin, T., Bitar, N., and Y. Rekhter, "Framework for Data Center (DC) Network Virtualization", RFC 7365, DOI 10.17487/RFC7365, October 2014, <<http://www.rfc-editor.org/info/rfc7365>>.
- [RFC7637] Garg, P., Ed. and Y. Wang, Ed., "NVGRE: Network Virtualization Using Generic Routing Encapsulation", RFC 7637, DOI 10.17487/RFC7637, September 2015, <<http://www.rfc-editor.org/info/rfc7637>>.

## Authors' Addresses

Xiaohu Xu  
Huawei

Email: [xuxiaohu@huawei.com](mailto:xuxiaohu@huawei.com)

Himanshu Shah  
Ciena Corp

Email: [hshah@ciena.com](mailto:hshah@ciena.com)

Yongbing Fan  
China Telecom

Email: fanyb@gsta.com