Signaling control/forward plane information between network virtualization edges (NVEs)
draft-wu-nvo3-nve2nve-00

Abstract

This document discusses how to provide control plane and forward plane information to the NVE associated with the tenant system for enabling interconnect between Tenant Systems that belong to specific tenant network.

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 August 22, 2013.

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. Conventions used in this document ................. 4
3. Solution Overview ..................................... 5
4. Mapping tables at the NVE ............................. 7
5. Key functions for signaling control/forwarding info to NVEs . 8
   5.1. Create and Update tenant Virtual Network(VN) .... 8
   5.2. Associate the NVE and tenant system with VN context ... 8
   5.3. Populate mapping tables at the local NVE ........... 9
   5.4. Distribute the mapping table to remote NVEs in the VN .. 9
   5.5. The mapping table update at the NVE when VM moves or connection fails ......................... 9
   5.6. The VN context re-association at the NVE when VM moves .. 10
6. IANA Considerations .................................... 11
7. Security Considerations ................................. 12
8. References ............................................... 13
   8.1. Normative References ............................. 13
   8.2. Informative References ........................... 13
Author’s Address ........................................... 14
1. Introduction

In [I.D-ietf-nvo3-framework], one control component is defined to provide the capability for Address advertisement and tunnel mapping. In [I.D-fw-nvo3-server2vcenter], the control interface between NVE and interconnection functionality is defined to provide the capability:

- Enforce the network policy for each VM in the path from the NVE Edge associated with VM to the Tenant End System.
- Populate forwarding table in the path from the NVE Edge associated with VM to the Tenant End System in the data center.
- Populate mapping table in each NVE Edge that is in the virtual network across data centers under the control of the Director.

However, there is no relevant work to discuss how those capability can be realized at the NVEs. This document goes into details to discuss how to provide control plane and forward plane information to the NVE associated with the tenant system for enabling interconnect between Tenant Systems that belong to specific tenant network.
2. 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 RFC2119 [RFC2119].

Site :

If multiple tenant systems connect to the VN through one NVE, the collection of these tenant systems and the NVE associated with these tenant systems are referred to as a site or virtualization network subnet.
3. Solution Overview

This document addresses how to provide control plane and forward plane information to the NVE associated with the tenant system for enabling interconnect between Tenant Systems that belong to specific tenant network.

Figure 1 shows the example architecture for interconnection between tenant systems. This example architecture assumes that:

- One tenant system or a NVE may belong to one tenant VN or several tenant VNs, e.g., VMa and NVE Edge4 belong to both VN2 and VN3.

- If one tenant system belongs to multiple tenant VNs, it may connect to each tenant VN by being attached to one NVE or multiple NVEs, e.g., VM1 connect to VN1 by being attached to NVE Edge 1.

- One site may belong to one tenant VN or several tenant VN, e.g., Site 2 belong to both VN2 and VN3.

- If one tenant system in one VN want to communicate with one tenant system in another VN, the interconnection functionality should get involved to setup tunnel between the interconnection functionality and the NVEs associated with the tenant system.
Figure 1: Figure 1.
4. Mapping tables at the NVE

Every NVE pair (local NVE and remote NVE) associated with the tenant system MUST maintain a mapping table entry for each currently attached tenant system. Each mapping table entry conceptually contains the following fields:

- The tunnel interface identifier (tunnel-if-id) of the tunnel between the remote NVE and the local NVE where the tenant system is currently attached. The tunnel interface identifier is acquired during the tunnel creation.

- The MAC address of the attached tenant end system. This MAC address is obtained from auto-discovery protocol between Tenant System and its local NVE.

- The IP address of the attached tenant system. This IP address is obtained from auto-discovery protocol between Tenant System and its local NVE.

- The IP address of the local NVE associated with the tenant system.

- The Identifier of VN context-VNID. This Identifier is obtained from auto-discovery protocol between Tenant System and its local NVE.
5. Key functions for signaling control/forwarding info to NVEs

5.1. Create and Update tenant Virtual Network (VN)

The tenant virtualization network (VN) is a collection of tenant systems, Network Virtualization Edges (NVE)(s) and end systems that are interconnected with each other. The tenant VN also consists of a set of sites where each can send traffic directly to the other.

In order to create or update a tenant VN, when a Tenant System is attached to a local NVE, the tenant system should inform the attached local NVE which VN the tenant system belong to.

- If the tenant system are the first participant in the VN through the local NVE, the tenant system and associated local NVE should be firstly added to the VN and the mapping table between the tenant system and the local NVE should be setup.

- If both the tenant system and the local NVE are not on the VN, the tenant system and associated local should be firstly added to the VN and then the mapping table associated with this tenant system should be setup at the local NVE and distributed to the other remote NVEs that belong to the same VN.

- If the local NVE is on the same tenant VN as the tenant system associated with the local NVE, only the tenant system needs to be added into the VN, i.e., the local NVE only needs to distribute mapping table to the other remote NVEs that belong to the same tenant VN.

- If the local NVE is not on the same tenant VN as the tenant system associated with that local NVE, the local NVE should firstly be added into the VN and then distributes the new mapping table to the other remote NVEs that belong to the same tenant VN.

- If one tenant system is the last participant connecting to the VN through local NVE, when this tenant system leave the VN, both the local NVE associated with this tenant system and mapping table associated with this tenant system should be removed from the VN.

5.2. Associate the NVE and tenant system with VN context

The VN context includes a set of configuration attributes defining access and tunnel policies and (L2 and/or L3) forwarding functions. When a Tenant System is attached to a local NVE, a VN network instance should be allocated to the local NVE. The tenant system should be associated with the specific VN context using virtual Network Instance (VNI). The tenant system should also inform the
attached local NVE which VN context the tenant system belong to. Therefore the VN context can be bound with the data path from the tenant system to the local NVE and the tunnel from local NVE associated with the tenant system and all the remote NVEs that belong to the same VN as the local NVE. For the data path from the tenant system and the local NVE, the network policy can be installed on the underlying switched network and forwarding tables also can be populated to each network elements in the underlying network based on the specific VNI associated with the tenant system. For the tunnel from local NVE to the remote NVEs, the traffic engineering information can be applied to each tunnel based on VNI associated with the tenant system.

5.3. Populate mapping tables at the local NVE

In some cases, two tenant systems may be attached to the same local NVE. In order to allow the NVE to locally route traffic between two tenant systems that are attached to the same NVE, the mapping table that maps a final destination address to the proper tunnel should be populated at the local NVE.

In some cases, two tenant systems may connect to the different VNs through the same interconnection functionality, in order to allow two tenant systems communication between two VNs, the mapping table that maps a final destination address to the proper tunnel should be populated in both NVE associated with two communicated tenant system and the interconnection functionality associated corresponding NVE.

5.4. Distribute the mapping table to remote NVEs in the VN

When the packet sent from one tenant system arrives at the ingress NVE associated with that tenant system, in order to determine which tunnel the packet needs to be sent to, the mapping table that maps a final destination address to the proper tunnel should also be distributed to all the remote NVEs in the VN using a control plane protocol or dynamic data plane learning. The mapping table may be advertised directly to other remote NVEs that belong to the same VN or firstly advertised to the centralized controller that maintain global view of NVEs that belong to the same VN and then let the centralized controller distribute the mapping tables to all the relevant remote NVEs that belong to the same VN.

5.5. The mapping table update at the NVE when VM moves or connection fails

In some cases, one tenant system may be detached from one NVE and move to another NVE. In such cases, the mapping table should be removed from the NVE to which the tenant system was previously
attached and the new mapping table should be created at the new NVE to which the tenant system is currently attached. Such mapping table should be updated at each remote NVE associated with the tenant system and the new NVE.

In some cases, one tenant system may fail to connect to the VN through the NVE. In such cases, the mapping table should be removed from the NVE to which the tenant system is currently attached. In addition, the mapping table should be updated at each remote NVE in the same VN through which the tenant system is communicating with the destination tenant system.

5.6. The VN context re-association at the NVE when VM moves

In some cases, one tenant system may be detached from one NVE and move to another NVE. In such cases, the VN context should be moved from the NVE to which the tenant system was previously attached to the new NVE to which the tenant system is currently attached. In order to achieve this, the per tenant system VN context can be maintained at the centralized database and be retrieved at the new place based on the VN Identifier (VNID).
6. IANA Considerations

This document has no actions for IANA.
7. Security Considerations

TBC.
8. References

8.1. Normative References

[I.D-ietf-nvo3-framework]


8.2. Informative References

[I.D-fw-nvo3-server2vcenter]
Author’s Address

Qin Wu
Huawei
101 Software Avenue, Yuhua District
Nanjing, Jiangsu 210012
China

Email: bill.wu@huawei.com