VDP Extensions

Yizhou Li (liyizhou@huawei.com)

Motivations

- VDP is a candidate control plane protocol to be used for signaling between the hypervisor and the associated external NVEs (Network Virtualization Edge) to distribute the virtual machine networking state to the external NVE(s) for further handling in IETF NVO3 context.
- draft-ietf-nvo3-hpvr2nve-cp-req-02 lists the requirements of the protocols and compares VDP with the requirements.
- draft-pt-nvo3-vdp-vm2nve-gap-analysis-00 describes the gap between current VDP and the earlier version of requirements.
- VDP extensions are required to fulfill the requirements.
- Either IETF or IEEE could do the extensions. IEEE 802.1 may be a good place to steer the work as VDP was defined here?

Split-NVE structure

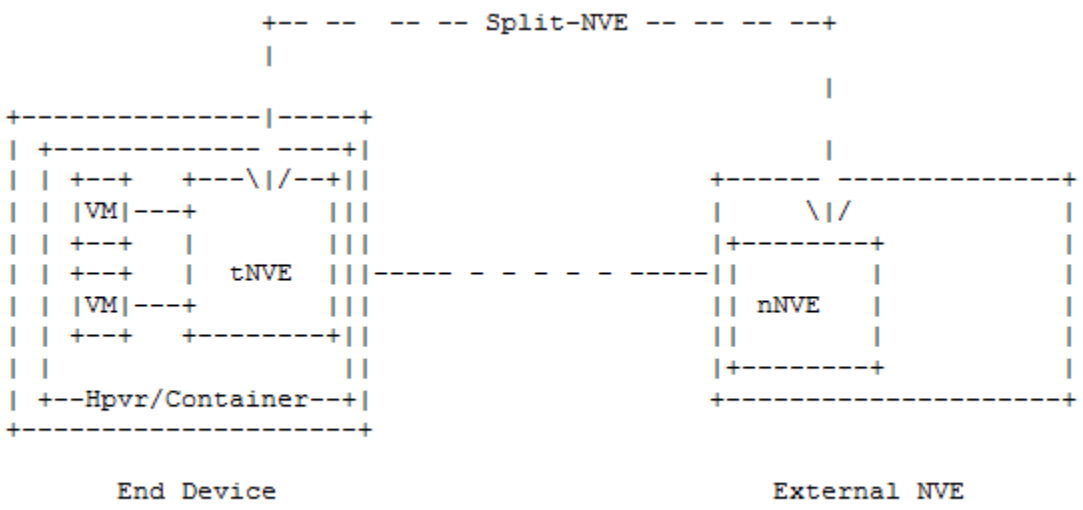


Figure 1 Split-NVE structure

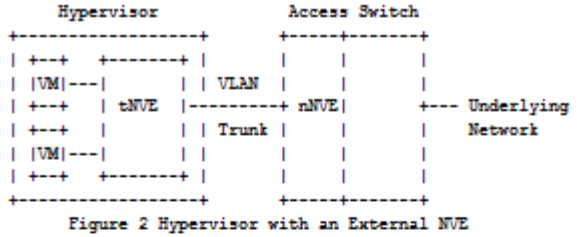


Figure 2 Hypervisor with an External NVE

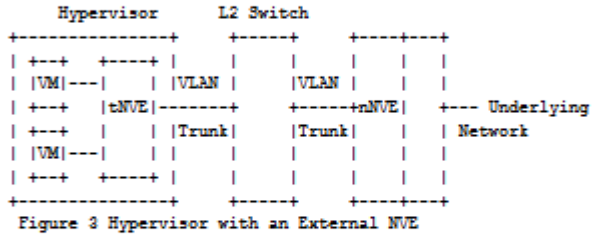


Figure 3 Hypervisor with an External NVE across an Ethernet Access Switch

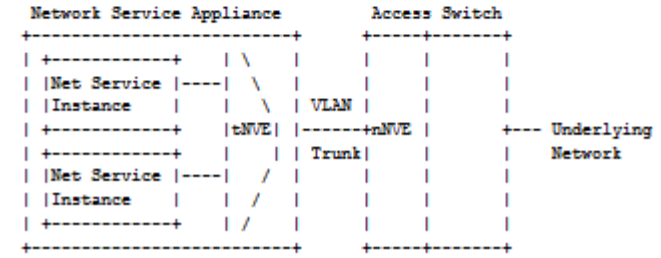


Figure 4 Physical Network Service Appliance with an External NVE

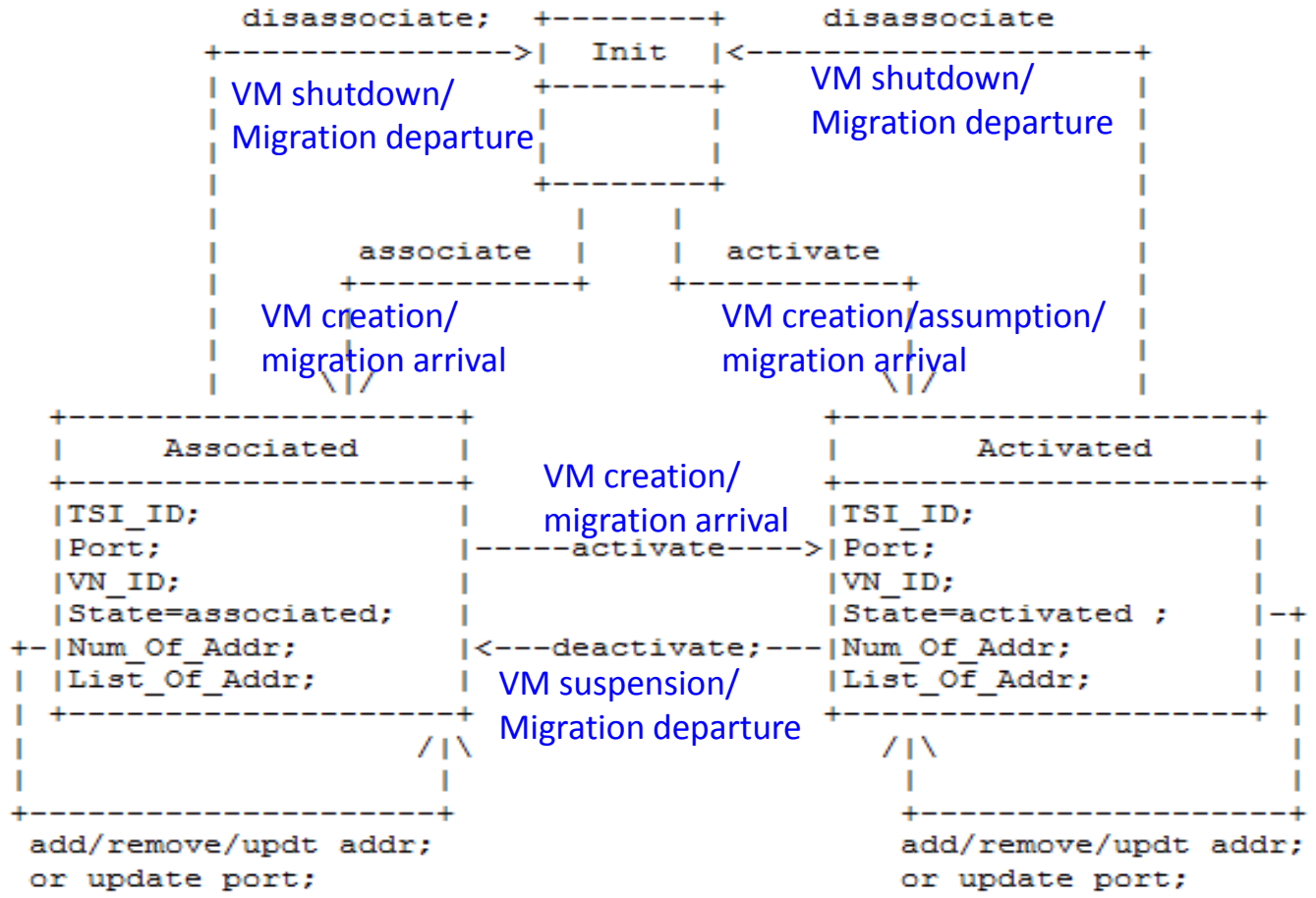
Split-NVE: a type of NVE that the functionalities of it are split across an end device supporting virtualization and an external network device.

tNVE: the portion of Split-NVE functionalities located on the end device supporting virtualization.

nNVE: the portion of Split-NVE functionalities located on the network device which is directly or indirectly connects to the end device holding the corresponding tNVE.

External NVE: the physical network device holding nNVE

State Transition of a TSI Instance on an External NVE



TSI: Tenant System Identifier Interface to a Virtual Network as presented to a Tenant System. To the Tenant System, the TSI is like a NIC.

Requirements – 1

#	req	VDP Extension required?
1	support a bridged network connecting End Devices to External NVE.	Yes. Must be able to send to a specific unicast MAC and should be able to send to a non-reserved well known multicast address other than the nearest customer bridge address
2	support multiple End Devices sharing the same External NVE via the same physical port across a bridged network.	
3	MAY support an End Device using multiple external NVEs simultaneously, but only one external NVE for each VN	
4	MAY support an End Device using multiple external NVEs simultaneously for the same VN	
5	allow the End Device initiating a request to its associated External NVE to be connected/disconnected to a given VN.	No. VN is indicated by GroupID
6	allow an External NVE initiating a request to its connected End Devices to be disconnected to a given VN	No. Bridge may send De-Associate

Requirements – 2

#	req	VDP Extension required?	
7	When a TS attaches to a VN, the protocol MUST allow for an End Device and its external NVE to negotiate a locally-significant tag for carrying traffic associated with a specific VN (e.g., 802.1Q tags).	No. VID==0 in request and bridge returns the assigned value in response Or specify GroupID in request and get VID assigned in returning response. Multiple VLANs per group is allowed	
8	allow an End Device initiating a request to associate/disassociate and/or activate/deactivate address(es) of a TSI instance to a VN on an NVE port.	requirements	VDP equivalence
		associate/disassociate activate/deactivate	pre-asso/de-associate associate/de-associate
		No. Needs clarification to allow transition from associate->pre-assoc	
9	allow the External NVE initiating a request to disassociate and/or deactivate address(es) of a TSI instance to a VN on an NVE port.	No. VDP bridge may initiate de-associate	

Requirements – 3

#	req	VDP Extension required?
10	allow an End Device initiating a request to add, remove or update address(es) associated with a TSI instance on the external NVE. Addresses can be expressed in different formats, for example, MAC, IP or pair of IP and MAC.	Yes. Needs extension for IPv4/IPv6 address association. Add a new “filter info format” type
11	MUST allow the External NVE to authenticate the End Device connected.	Out-band mechanism is preferred, e.g. MACSec, 802.1x. Not necessarily be VDP extension.
12	be able to run over L2 links between the End Device and its External NVE.	No. L2 protocol naturally
13	SHOULD support the End Device indicating if an associate or activate request from it results from a VM hot migration event.	Yes. New bits for migration indication in new “filter info format” type (Current VDP has M bit for migrated VM on destination hypervisor and S bit for that on source hypervisor. It is indistinguishable when M/S is 0 between no guidance and events not caused by migration where NVE may act differently.)

Summary

- Extensions include:
  - Allow ECP destination address to be:
    - Specific unicast destination MAC, or
    - A non-reserved well known multicast address other than the nearest customer bridge addres
  - Add new filter info format type for IPv4/v6 address association
  - Put bits for migration indication into new filter info format
  - Clarify assoc->pre-assoc : state machine in current VDP allows it but not intentionally

Current Status

- IETF NVO3 has sent a liaison to IEEE 802 to start the protocol extension per requirement draft (https://datatracker.ietf.org/liaison/1401/)

Purpose: For action
Deadline: 2015-07-19 Action Needed

The IETF NVO3 Working Group has been developing the requirements for a Control Plane Protocol between server Hypervisors and Network Virtualization Edge (NVE) devices in virtualized overlay networks. The current draft can be found at http://datatracker.ietf.org/doc/draft-ietf-nvo3-hpvr2nve-cp-req/.

This requirements document suggests extending the VDP (VSI Discovery and Configuration Protocol) protocol specified by IEEE Std 802.1Qbg as a solution. We would particularly welcome IEEE 802.1 Working Group's review of Section 5 of the draft. That section compares the conceptually similar terms in NVO3 and the VDP context. It also summarizes the potential technical extension work required for VDP to be used as the control plane protocol between the hypervisor and NVE.

In view of the progress of this work, we would like to suggest IEEE 802.1 Working Group to use that draft as a base requirement reference for VDP extensions in the aforementioned context. Please note that the status of this draft in the IETF, a "Working Group draft", indicates that the Working Group considers it an appropriate starting point but it is still subject to change based on NVO3 WG consensus. While a determination has not yet been made that there is technical consensus on all the details of the draft, there is consensus on basing the Hypervisor to NVE protocol on VDP with appropriate extensions.