

VDP Extensions

Yizhou Li
Pat Thaler
Eric Gray

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

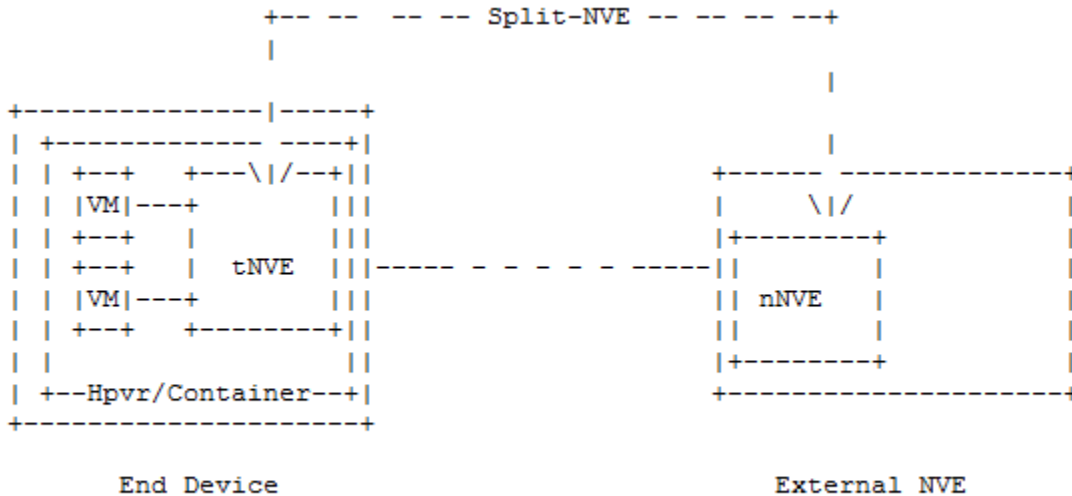


Figure 1 Split-NVE structure

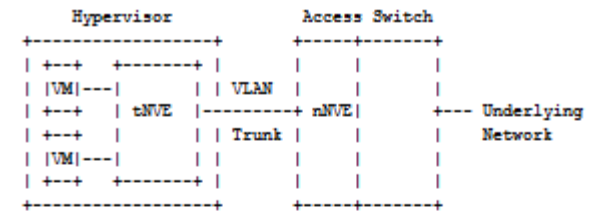


Figure 2 Hypervisor with an External NVE

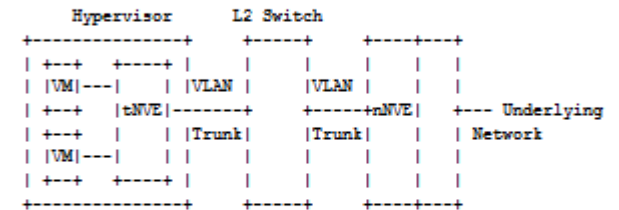


Figure 3 Hypervisor with an External NVE across an Ethernet Access Switch

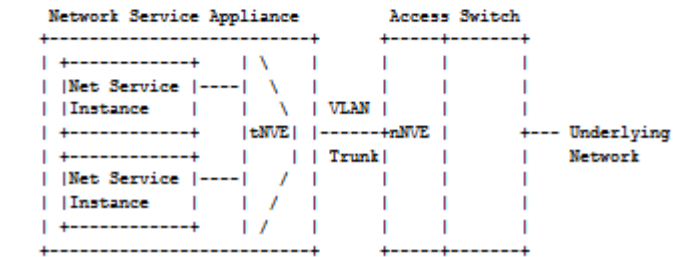


Figure 4 Physical Network Service Appliance with an External NVE

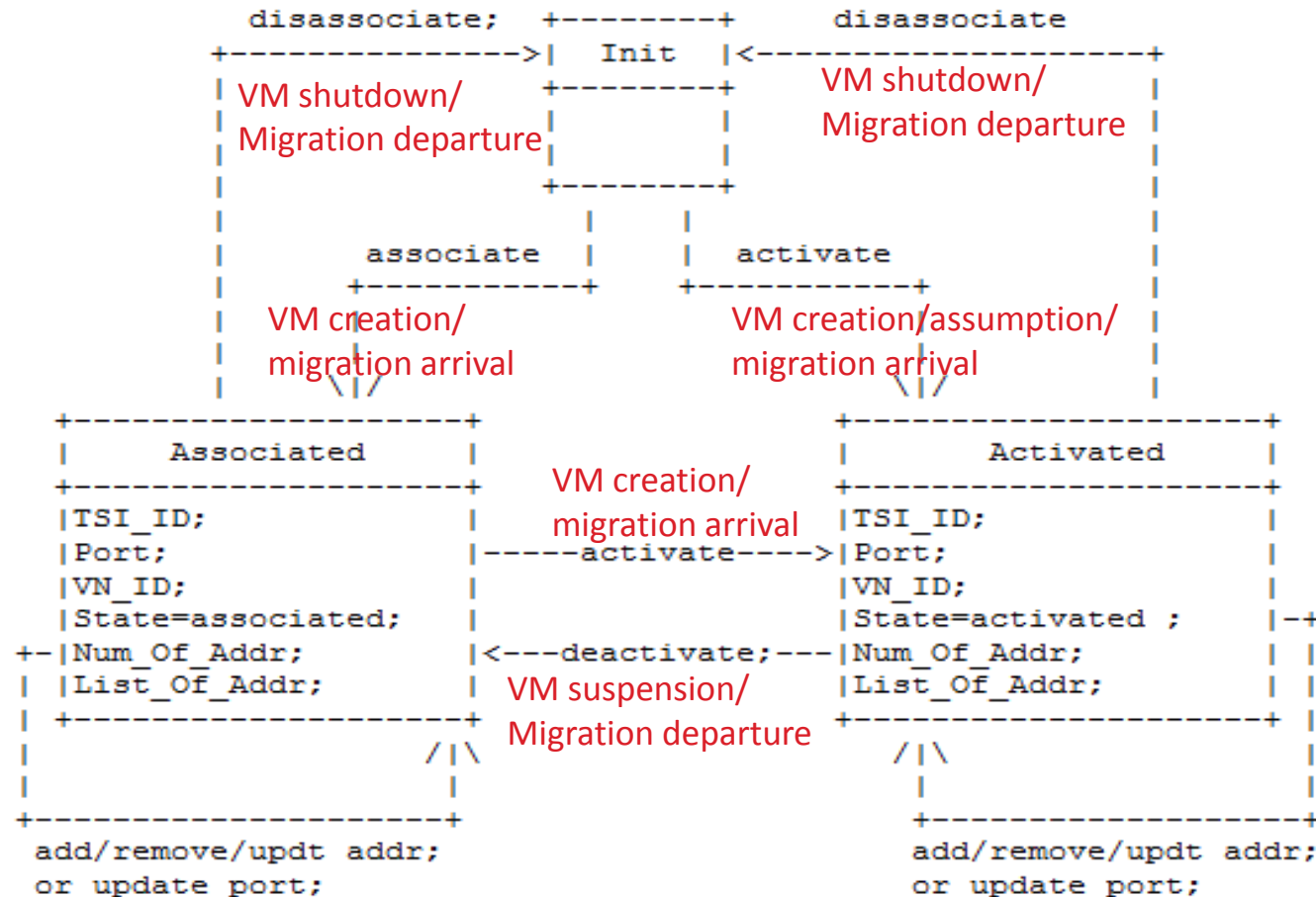
Split-NVE: a type of NVE that the functionalities of it are split across an end device supporting virtualization and an external network device.

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 | extension? |
|---|----------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------|
| 1 | support a bridged network connecting End Devices to External NVE. | Needs extension. Dest MAC can be a specific unicast MAC besides Nearest Customer Bridge group MAC |
| 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 | Needs clarification and extension for link aggregation support. For req-4, (pre-)associate status needs to be synchronized on all NVE ports. |
| 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. | 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 | Bridge sends De-Associate |

Requirements – 2

| # | req | | |
|---|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------|-------------------------------------------------|
| 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). | VID==0 in request and bridge returns the assigned value in response | |
| 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 |
| | | Needs extension to allow 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. | VDP bridge initiates de-associate | |

Requirements – 3

| # | req | |
|----|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 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. | Needs extension for IPv4/IPv6 address |
| 11 | MUST allow the External NVE to authenticate the End Device connected. | Needs extension for authentication |
| 12 | be able to run over L2 links between the End Device and its External NVE. | 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. | 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. Needs extension to clearly define them. |

Summary

- ▶ Extensions include:
 - **Specific destination MAC** other than nearest customer bridge group
 - Multiple uplinks considerations:
 - Direct connection: **field** to indicate which NVEs are for which VNs
 - Indirect connection: If intermediate bridge has no knowledge, how does traffic get to distribute to NVEs? (needs more NVO3 discussion)
 - Authentication: **TLV** for integrity check.
 - Direct connection: using existing mechanism like .1x or macsec.
 - Indirect connection : IP sec? needs ietf work
 - IP address binding: **add new filter info format type**
 - Clearer migration indicator: **put bits into new filter info format**
 - **Clarify assoc→pre-assoc** : state machine allows it but not intentionally

Next Steps?

- ▶ Interest to participant in this work?
- ▶ Liaison from IETF:
 - Small amendments on VDP spec
 - IEEE takes the work to define the extensions based on IETF requirement document
or
 - IETF asks for code point allocation of new filter info format and possible TLV types.