NVO3 Architecture Updates and Open Issues

draft-ietf-nvo3-arch-02

Thomas Narten, Lawrence Kreeger, David Black,
Jon Hudson, Marc Lasserre

Recharter's Affect on Architecture?

- The new charter text specifically moves BGP based solutions to the BESS WG, and focuses NVO3 on a "xxx logically centralized" NVA architecture.
- The current architecture document already reflects this.
- Question: Does the charter change also rule out MPLS with Locally Significant VN Context in encapsulations?
  - There is a mention of MPLS-over-GRE and 20 bit labels used as VN Context IDs; However there is no mention of locally significant VN context in the document.
  - It is possible to use MPLS-over-GRE encapsulation with a globally significant label as the VN Context ID.

Section 3.4 - VM Orchestration Systems

- The following text was added to address concerns about NVO3 applying beyond VMs on Hypervisors:

"Note also, that although this section uses the term "VM" and "hypervisor" throughout, the same issues apply to other virtualization approaches, including Linux Containers (LXC), BSD Jails, Network Service Appliances as discussed in Section 5.1, etc.. From an NVO3 perspective, it should be assumed that where the document uses the term "VM" and "hypervisor", the intention is that the discussion also applies to other systems, where, e.g., the host operating system plays the role of the hypervisor in supporting virtualization, and a container plays the equivalent role as a VM."

Section 11 - NVO3 Data Plane Encapsulation

- Removed the paragraph arguing against the NVO3 WG working on new encapsulations
- However, text still states that the most important requirement is a sufficiently large VN Context ID

Open Areas

- The document has several areas beginning with "[Note:" which are looking for WG feedback.
- These Notes have been in the document for a long time without any WG feedback being expressed.
- The following slides propose how the architecture team may resolve these areas if there is no further feedback from the WG.

Open Area: Section 7.3 NVA External Interface

Options:

A. Single IP Address for entire NVA (failover within NVA hidden from NVEs)
B. Multiple IP addresses for NVA, all equivalent (NVEs fail over to another IP if one becomes unresponsive)
C. Multiple IP addresses for NVA, each IP addresses given a priority with interactions distributed across IPs of equal priority (affinity determined by NVE)
D. Multiple IP addresses for NVA, IP addresses used partitioned on a per VN basis
E. Multiple IP addresses for NVA, each response from NVA may direct NVE to use a different IP address for the request next time

Proposed Resolution: Leave all the options open in the architecture and leave it to solutions to decide (remove the "[Note:" comments).

Open Area: Section 8.1 NVE-NVA Interaction Models

A. Should the architecture support a model where all NVEs interact with the NVA, or
B. Should the architecture support a hybrid approach to support legacy VM orchestration systems that control some NVEs, while other NVEs use the NVA.
   - This requires the VM orchestration system to interact with the NVA. A protocol would be needed between the two. Would we extend the NVE-to-NVA protocol to support the VM orchestration acting as a proxy for all the NVEs that it supports? Would a different protocol be better suited?

Proposed Resolution: Do not support a hybrid solution and NVA to VM orchestration system protocol.

Gateways

- There has been mailing list discussion on “gateways”
- It was noted that some of the discussion may have not been productive due to different views of what is meant by “gateway”. Due to the overloading of this term, we propose a finer grain taxonomy for gateways.

Proposed Gateway Taxonomy

1. L2 Gateways (Bridging)
   a) L2 VN to Legacy L2 (e.g. VLANs, L2 VPNs)
   b) L2 VN to L2 VN: for the purposes of creating Closed User Groups within a bridged domain
2. L3 Gateways (Only IP packets flow through gateway)
   a) L3 VN to Legacy L2
   b) L3 VN to Legacy L3 (e.g. local to DC, to WAN, to L3 VPN)
   c) L3 VN to L2 VN
   d) L2 VN to L2 VN
   e) L3 VN to L3 VN: for the purposes of creating Closed User Groups within an IP routed domain

- All the above gateways can be centralized.
- All VN to VN (L2-L2, L3-L2, L3-L3) gateways are candidates for distribution to the NVEs. VN to legacy network gateways are not (if the gateways are routers, they can use traditional router methods for multi-homing).
- We could add sections to the architecture covering each of the above types of gateways, including both the centralized and distributed implementations.