NVo3 Control Plane Protocol Using IS-IS

draft-xu-nvo3-isis-cp-02

Xiaohu Xu (Huawei)
Himanshu Shah (Ciena Corp)
Yongbing Fan (China Telecom)

IETF96, Berlin

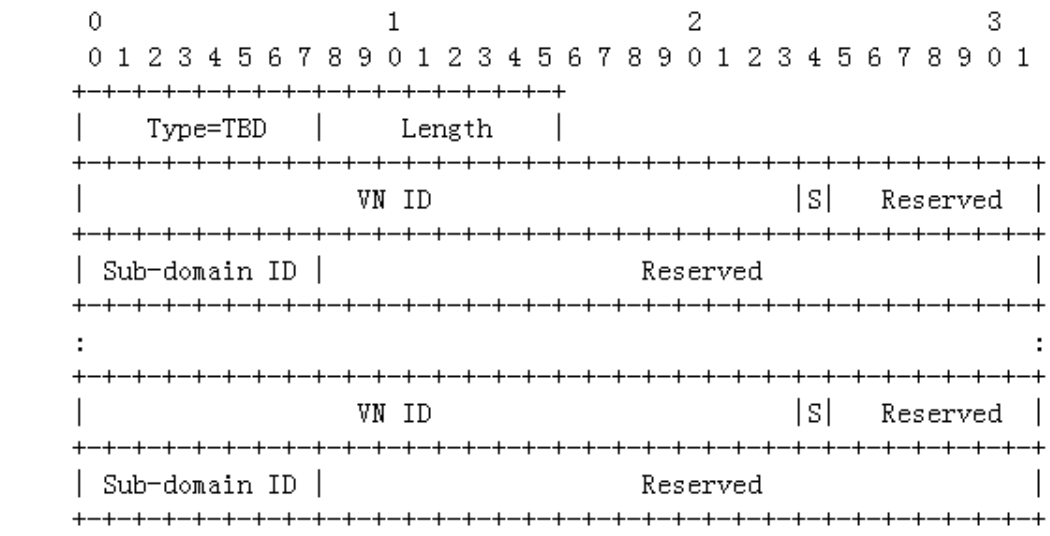
Motivation

- 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.
- Although BGP or SDN controller could still be used as the control plane protocol in campus network environment, both of them seem a bit heavyweight, especially for small and even medium sized campus networks.
- 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).

VN Membership Auto-discovery

- By propagating the VN membership info among 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 (see next page) of the following TLVs originated by that NVE:
  - TLV-135 (IPv4 extended reachability)
  - TLV-236 (IPv6 reachability)
- When the above TLV is propagated across level boundaries, the VN Membership Info sub-TLV contained in that TLV SHOULD be kept.

VN Membership Info Sub-TLV



- **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.

Tunnel Capability Advertisement

- This IS-IS based control plane protocol for NVo3 could support any specific NVo3 data encapsulation formats such as VXLAN.
- 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].

MAC Address Learning

- MAC addresses of local CE hosts would still be learnt by NVEs as normal bridges.
- Two options for learning MAC addresses of remote CE hosts:
  - data-plane based MAC learning
  - 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.

Control-plane based MAC Learning

- NVEs advertise local 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 an 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.

Next Steps

- **Comments?**