NVO3 Overlay P2MP Ping
draft-xia-nvo3-overlay-p2mp-ping-00

Liang Xia, Weiguo Hao, Greg Mirsky

July 2014   Toronto

Goal

- Background:
  - NVO3 Overlay P2P Ping mechanism drafts: draft-jain-nvo3-overlay-oam-01, draft-kumar-nvo3-overlay-ping-01;
  - Currently, lack of P2MP Ping solution contents;

- Goal: try to provide NVO3 Overlay P2MP Ping solution based on MPLS P2MP LSP Ping solution [RFC6425], and by extending the above drafts.

NVO3 Multicast Service Analysis

- 4 methods:
  1. No multicast support;
  2. Replication at the source NVE;
  3. Replication at a centralized multicast service node;
  4. IP multicast in the underlay.

NVO3 P2MP Ping Requirements

- The NVO3 P2MP ping must have the ability to:
  - Support the connectivity verification from an arbitrary NVE to either a specific set of NVEs or all NVEs overlaid on the underlay Multicast Distribution Tree (MDT);
  - Support the fault localization of the underlay MDT;
  - Support the path tracing of the underlay MDT exercised by any given overlay path.

Overview of NVO3 P2MP Ping

- This draft refers to [RFC6425] and describes the following extensions:
  - The packet format extension to the NVO3 overlay P2MP Echo Request/Reply messages;
  - The extensive mechanisms of NVO3 overlay P2MP ping operations;
  - Different operation mechanisms for two scenarios - with or without underlay multicast support;
  - Special considerations for traceroute function;
  - Support for the hierarchical NVE scenario.

Packet Format

- NVO3 overlay P2MP ping packet is an IPv4 or IPv6 UDP packet with UDP Port 3503, and remains the same basic structure as mentioned in Section 3 of [RFC4379];

- Other extensions: references Section 4 of [NVO3OVERLAYPING] as the extended specification to the payload of inner UDP of NVO3 Echo Request/Reply packet format, which include: a new "N" flag; extended Message Type, Reply Mode, Return Code and Return Subcode; new defined Target Object TLV; DDM Extension; P2MP specific TLVs.

P2MP NVO3 Overlay Path's Identification

- Different conditions:
  - Without underlay multicast support: split into several NVO3 P2P sessions between ingress node and egress nodes;
  - With underlay multicast support:
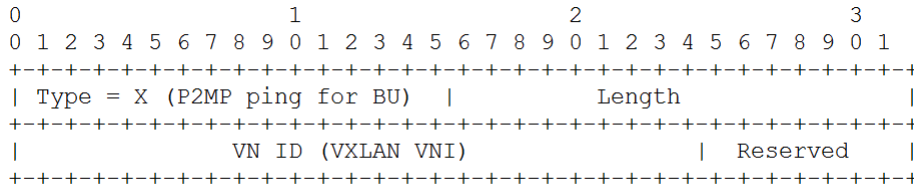
Broadcast and Unknown Unicast Packets Case

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Type = X (P2MP ping for BU)  |            Length              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|           VN ID (VXLAN VNI)             |    Reserved         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
          Figure 1: L2 VN ID sub-TLV for VXLAN
```

L3 Multicast Packets Case
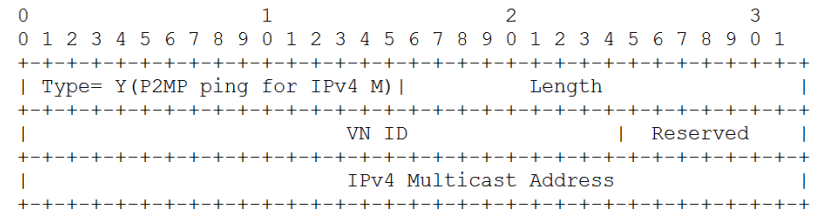
```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Type= Y(P2MP ping for IPv4 M)|            Length              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                VN ID                   |    Reserved         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                  IPv4 Multicast Address                       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      Figure 2: VN multicast sub-TLV for IPv4 overlay network


 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Type= Z(P2MP ping for IPv6 M)|            Length              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                VN ID                   |    Reserved         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                  IPv6 Multicast Address                       |
|                                                              |
|                                                              |
|                                                              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      Figure 3: VN multicast sub-TLV for IPv6 overlay network
```
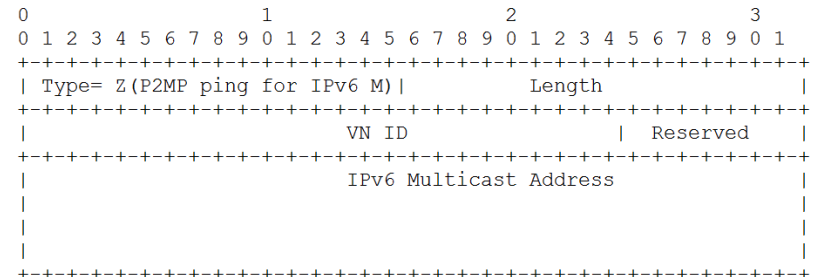
Other TLVs and Flags

- P2MP Responder Identifier TLV: limit the scope of responses
  - IPv4/IPv6 Egress Address P2MP Responder sub-TLVs: cannot be used for some constraints
  - IPv4/IPv6 Node Address P2MP Responder sub-TLVs: can be directly used
- The Echo Jitter TLV, The Respond Only If TTL Expired Flag: can be directly used

Theory of Operation

- Different conditions to ping mode:
  - Without underlay multicast support: same with P2P ping, need ingress node congestion control;
  - With underlay multicast support:
    - Ingress node: encapsulating ping packet and sending;
    - Transit node, branch node, bud node, egress node: transfer ping packet and/or respond to it, through checking multicast address, Target Object TLV, other TLVs against control plane.

- Considerations for traceroute mode:
  - Ingress replication: no problem;
  - Underlay multicast support: need specific mechanism for check end;

Next Step

- Comments and suggestions?

- Coordination with NVO3 Overlay P2P Ping drafts ☺

Thanks!

Liang Xia