

Generic Overlay OAM and Datapath Failure Detection

Diego Garcia del Rio (Nuage Networks)

Pradeep Jain, Kanwar Singh

Wim Henderickx

Anil Lohiya, Ravi Shekhar

Vinay Bannai

Kris Price

Nuage Networks

Alcatel-Lucent

Juniper Networks

Paypal

Microsoft

Motivation of Generic Overlay OAM Application

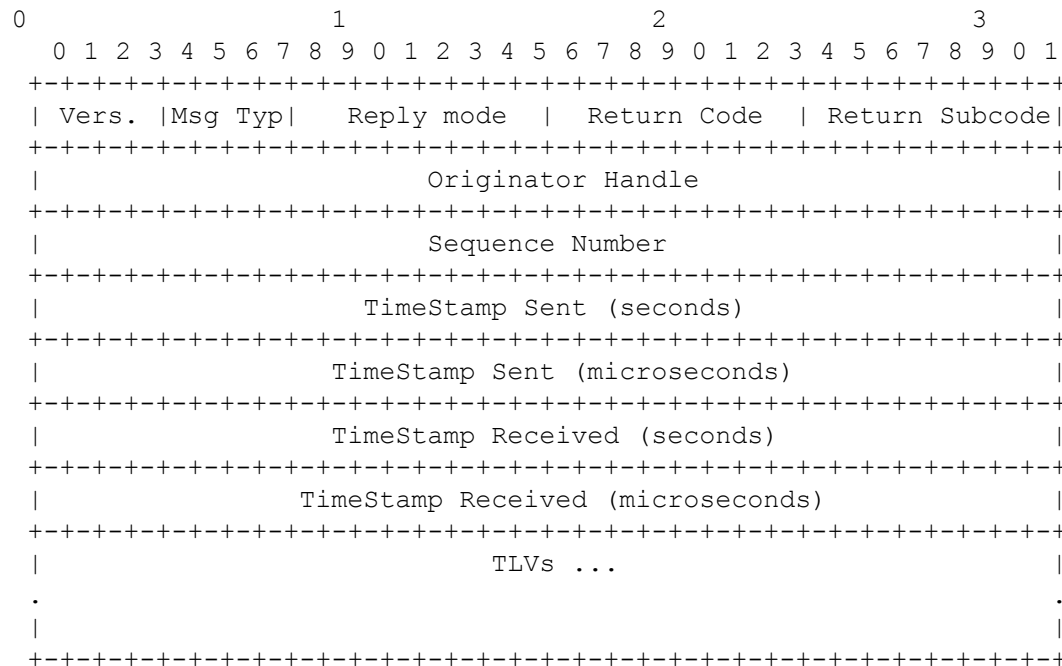
- Existing IP/Ping Trace does not work well for Overlays.
- Consistency of Control and Data Plane Programming.
- Continuity Check .
- Fault Verification.
- Fault Isolation.
- Performance.
 - Packet Delay Statistics.
 - Packet Loss Measurements.

Requirements for Overlay OAM Framework

- Originating Overlay End Point should send the OAM Frame (Echo Request) following the same Data-Path as that of End-System's traffic.
- Terminating Overlay End Point should be able to differentiate the OAM Frames from the End System Data, and send it to OAM Application for Reply.
- Should be able to achieve OAM for applications using L2 and L3 Overlays.
- Should give capability to Trace the Path taken in the Underlay for a given Overlay Segment.
- Should work for Overlay Technologies as VxLAN, NVGRE, MPLSoGRE, MPLSoUDP.

Generic Overlay OAM Packet

- UDP Packet with well defined Port XXXX.
- Version number identifies subsequent element format.
- TLV Format for identifying various types of Overlay OAM Request/Reply (e.g. VxLAN, NVGRE, MPLSoGRE, MPLSoUDP).
- Optional Sub-TLVs to identify End-System information.



Overlay Ping - Originating Overlay End Point Procedures

- Outer Header Should follow Encapsulation similar to Overlay Segment
 - VxLAN / NVGRE : Set Router Alert Bit ([draft-singh-nvo3-vxlan-router-alert](#) / [draft-singh-nvo3-nvgre-router-alert](#)) – indicating Control Packet
 - VXLAN-GPE or GeNeVE: Relying on the OAM BIT of the header
 - GUE – Setting the C bit
 - MPLSoGRE/MPLSoUDP : Add Router Alert Label following Overlay Segment Label.
- Inner Header Encapsulation
 - L2 Overlay
 - Echo Request should have inner Ethernet Header, followed by IP and UDP Header
 - Inner DMAC: 00-00-5E-90-XX-XX (to be assigned by IANA)
 - Inner DST-IP: In Range of 127/8 (for IPv4) or In Range 0:0:0:0:0:FFFF:127/104 (for IPv6)
 - Inner DST-UDP Port: XXXX (assigned by IANA for Overlay OAM).
 - Generic OAM Frame
 - L3 Overlay
 - Echo Request's Encapsulation is same as above, except Inner Ethernet Header.

Overlay Ping - Terminating Overlay End Point Procedures

- Identify Packet as Control Packet from Outer Header
 - VxLAN / NVGRE : Router Alert Bit ([draft-singh-nvo3-vxlan-router-alert](#) / [draft-singh-nvo3-nvgre-router-alert](#))
 - MPLSoGRE/MPLSoUDP : Router Alert Label.
- Identify Packet as Overlay OAM Packet and sent it to OAM Application using Inner Header Encapsulation
 - L2 Overlay
 - Inner DMAC: 00-00-5E-90-XX-XX (to be assigned by IANA)
 - Inner DST-IP: In Range of 127/8 (for IPv4) or In Range 0:0:0:0:0:FFFF:127/104 (for IPv6)
 - Inner DST-UDP Port: XXXX (assigned by IANA for Overlay OAM).
 - L3 Overlay
 - Inner DST-IP: In Range of 127/8 (for IPv4) or In Range 0:0:0:0:0:FFFF:127/104 (for IPv6)
 - Inner DST-UDP Port: XXXX (assigned by IANA for Overlay OAM).
- Validate Control/Data Plane and Send Echo Reply with Return Code

Overlay Ping Construct

- Two Type of Packets
 - Echo Request
 - Echo Reply
- Reply Modes
 - Do not reply
 - Reply via IPv4 UDP Packet
 - Reply via Overlay Segment
- Return Codes
 - No return code
 - Malformed Echo Request Received
 - Overlay Segment Not Present
 - Overlay Segment Not Operational
 - Return-Code-OK

Path Trace Procedures

- Echo Request is prepared via same Encapsulation as that of Overlay Tunnel.
- Echo Request packets are sent with Incremental TTL values 1, 2 ..n of Outer Header to probe the Path taken by the Overlay.
 - For each TTL value, on UDP-based encapsulations, the inner and outer source UDP ports are changed (incremented) a configurable number of times as to potentially exercise multiple ECMP paths.
 - Alternatively, the user can request a specific udp source port if a particular ECMP path is suspect and the user has knowledge on which source-ports hash to the particular link.
 - MPLSoGRE encapsulation could include a hash label with varying
- TTL Expiry creates an Exception, packet is sent to OAM Application based on inner UDP Port.
- Each Transit Node as well as Overlay End Point receiving such an exception should send Echo Reply.

Procedures End-System Ping

- **Originating Overlay End Point**
 - Echo Request similar to Overlay Ping along with :-
 - TLV Identifying Overlay Segment.
 - Sub-TLV identifying Type of End-System Ping (e.g MAC, IP, MAC/IP or 'raw header' of End-System).
 - Echo Request follows same Data Path as that of End-System Data.
- **Terminating Overlay End Point**
 - Terminating Overlay End Point validates End-System(s) information and send Echo Reply with proper Return Code for each End-System.
 - The overlay endpoint can use the provided sub-TLV for the End-System to probe the different forwarding tables and ACL tables to determine the fate of the packet.
- **Return Code**
 - End-System Present
 - End-System Not Present
- **Sub-return code**
 - Cannot determine action
 - End system action forward
 - End system action flood
 - End-System action dropped by rules
 - End-System action dropped by other

Next Step

- Looking forward to Comment/Feedback.
- Request the document to be accepted as WG document