Overlay OAM Design Team Report

Greg Mirsky gregory.mirsky@ericsson.com
Carlos Pignataro cpignata@cisco.com
Nagendra Kumar naikumar@cisco.com
Deepak Kumar dekumar@cisco.com
Erik Nordmark nordmark@acm.org
David Mozes davidm@mellanox.com
Mach Chen mach.chen@huawei.com
Yizhou Li liyizhou@huawei.com

IETF-96 July, 2016

Motivation

- Existing work in NVO3, SFC and BIER WGs on OAM Framework, Requirements, and Solutions
- Look at the OAM “puzzle” holistically, prevent divergence
- E.g.:
  - Proposed adaptation, extension of existing OAM protocols (BFD in VXLAN)
  - Proposed new mechanism (Transcending Traceroute)
- Common OAM presentation and discussion at IETF-94
- OOAM DT report to RTGWG at IETF-95

OOAM DT Charter

https://trac.tools.ietf.org/area/rtg/trac/wiki/RtgOoamDT

This Design Team is chartered to first produce a brief **gap analysis** and **requirements** document to focus its work on protocol extensions. This should be published by March 2016. With that basis, this Design Team is chartered to rapidly propose extensions to existing IETF OAM protocols such as those discussed in [RFC 7276] and new ones to support the requirements for OAM from NVO3, BIER, and SFC. The Design Team will produce an initial proposal by IETF 95. It is expected that the initial proposal will provide guidance to additional people who will be interested in working on the details and gaps.

The Design Team will consider the preliminary OAM requirements from NVO3, BIER, and SFC. The Design Team should align with the LIME WG's work on common YANG models of OAM.

Overlay OAM Requirements

draft-ooamdt-rtgwg-ooam-requirement-01

Structure

- Fault Management
  - Proactive FM
  - On-demand FM
- Performance Management
  - Active PM
  - Passive PM
- Alarm Indication Signal (Suppression)
- Resiliency
- Security Considerations

Requirements

- OOAM independent from a transport layer
- Any node implicitly serves as MEP
- SDN-azation of Overlay OAM
- Proactive and on-demand OAM created equal
- Unidirectional Overlay OAM (CC and PM) optimization as services (multicast, SFC) are unidirectional
- OAM is about what is going in the transport layer and thus it must be in-band , i.e. fate sharing with data traffic
- Bi-directional OAM is important too, e.g. CC-CV and out-of-band Fault Management Signal
- Path MTU Discovery

Fault Management

- Proactive
  - Continuity Check
  - Remote Defect Indication
  - Connectivity Verification
- On-demand
  - LoC defect localization
  - path tracing through overlay network
  - verification of mapping between overlay network and client layer services
  - fault localization of Loss of Continuity check at transport layer
  - ECMP discovery and verification
  - proxy ping/traceroute
- Fault Management Signals like Alarm Indication Signal to suppress client layer alarms when server layer fault detected
- Overlay network survivability may use protection switching and restoration

Performance Measurements

- Passive and Active Performance Measurement OAM are complimentary instruments in OOAM toolbox
  - One-way active and passive
  - Two-way active
- Support calculation of performance metrics:
  - packet delay
  - packet delay variation
  - packet loss
  - goodput (delivered throughput)
- Definition of Terms at: https://tools.ietf.org/html/rfc7276
- Interpretation of Passive PM is different from RFC 7799:
  - a measurement method that should not modify the actual data packet processing behavior on underlay and overlay network

Security Considerations

OAM requirements for various Overlay encapsulations may have security implications. For example, if proactive Fault Management (FM) is required, the security implication is that a passive eavesdropper can know when the session is down. Or, proactive FM may be used either to launch DoS or to highjack session and impact state, e.g. cause protection switchover. These security implications are natural results of the requirements, and do not depend on the particular implementation. Whether existing security mechanisms of existing protocols proposed to be re-used in OAM for overlay networks are adequate or require enhancements is for further study. New OAM protocols for overlay networks must consider their security mechanism to on per-solution basis.

Overlay OAM: Gap Analysis

draft-ooamdt-rtgwg-oam-gap-analysis-02

Gap Analysis Goals

- Today, we can ping/traceroute/BFD the underlay; that does not tell us much about the VNI/SFP/overlay.
- Two dimensions:
  1. <u>Operators</u>: Functionally adjacent to long-existing operational practice (format on the wire is less important)
  2. <u>Implementers</u>: Similar across different Encaps (reuse encodings?)

Gap Analysis Detail

- Done:
  - Identification of existing **OAM Protocols**
  - Possible feature Gaps within each OAM protocol
  - Applicability of OAM Protocols to different Overlays
  - The encapsulation of an overlay network uses one of methods discussed in draft-ietf-rtgwg-dt-encap to distinctly identify the payload as OAM, i.e. non-user, packet
  - All Overlay OAM protocols share the common Overlay OAM Header
- To be Done:
  - Encapsulation-specific requirements of OAM Protocol (extensions to the underlay encap)

Available OAM tools

Fault Management:

- proactive continuity check:
  - Bidirectional Forwarding Detection (BFD) for point-to-point as defined in [RFC5880], [RFC5882], [RFC5883], [RFC5884], [RFC5885], [RFC6428] and [RFC7726];
  - BFD for multipoint network as defined in [I-D.ietf-bfd-multipoint] and [I-D.ietf-bfd-multipoint-active-tail];
  - S-BFD as defined in [I-D.ietf-bfd-seamless-base] and [I-D.ietf-bfd-seamless-ip];
- on-demand continuity check and connectivity verification:
  - MPLS Echo Request/Reply, a.k.a. LSP Ping, as defined in [RFC4379] and its numerous extensions;
  - LSP Self-ping, as defined in [RFC7746];
  - [I-D.kumarzheng-bier-ping] is a good example of generic troubleshooting and defect localization tool that can be extended and suited for more specific requirements of the particular type of an overlay network.

Performance Measurement OAM

- Active:
  - Loss and Delay Measurement in MPLS networks, RFC 6374
  - One/Two-way Active Performance Measurement Protocol(s), RFC 4656/RFC 5357
- Passive:
  - Alternate Marking Method
- Conclusions:
  - RFC 6374 can be used as foundation of active PM OAM in overlay networks. The YANG data model of the packet loss and delay measurement based on RFC 6374 can improve control and increase operational value of active performance measurement in overlay networks.
  - Alternate Marking Method being proposed as Passive OAM in BIER and can be used in NVO3 and SFC, given supported by overlay network encapsulation

Example 1:

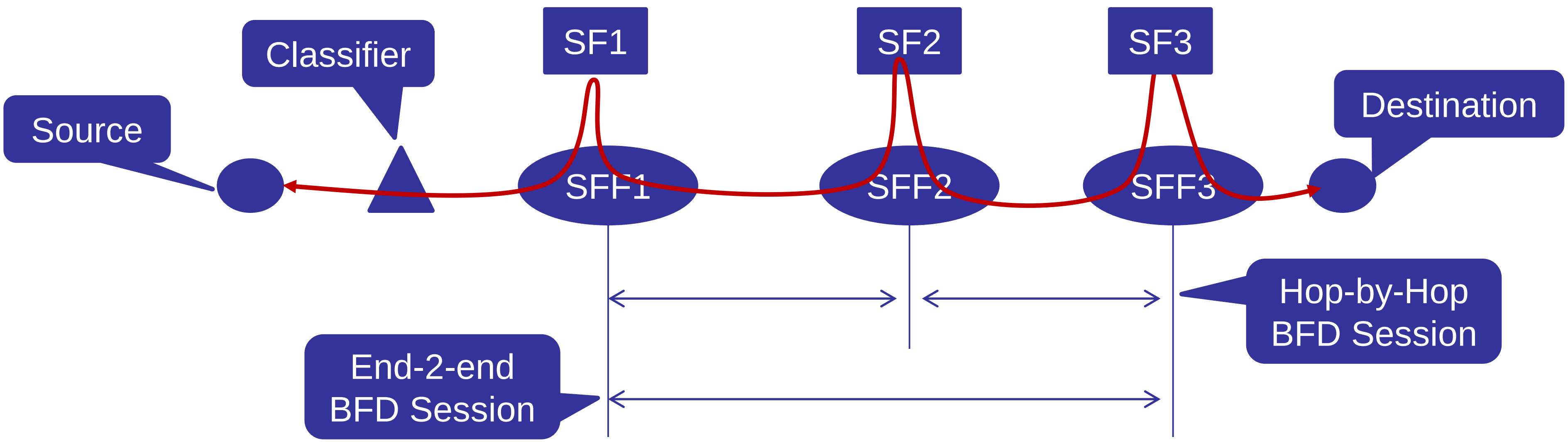
BFD for SFC

based on draft-ooamdt-rtgwg-gap-analysis-02

OAM for SFC Scope

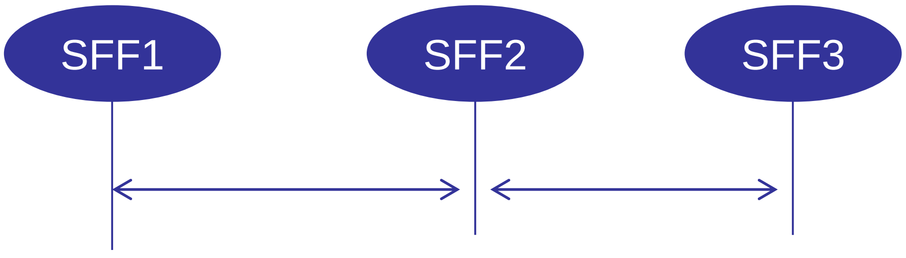
- Continuity Check on the SFP
- Verify that the SFF has the attachment point to talk to the SF
- Testing connectivity, **NOT** SF Functionality

Use Case



- Where to deploy the BFD sessions?
  - Between SFFs (the major case)
    - Hop-by-hop BFD session (e.g., SFF1<->SFF2, SFF2<->SFF3)
    - End-2-end BFD session (e.g., SFF1<->SFF3)
  - Other possibilities ?

Hop-by-hop Case



- An SFF should have capability to determine whether a packet should be delivered to an SF or terminated.
- Encapsulation dependent

End-2-End Case

- An SFF (e.g., SFF2) should have capability to determine whether a packet should be delivered to an SF or the next hop SFF.
- Encapsulation dependent

Control Plane

- BFD session bootstrapping
  - In-band signaling
  - Out-of-band channel
  - Centralized controller

Encapsulations

- BFD with IP/UDP encapsulation
  - Same as RFC5881 and 5884
  - The source/destination addresses and UDP port are derived from the IP/UDP header
- BFD without IP/UDP encapsulation
  - Add source and destination addresses field
  - UDP port is not necessary, the “Next Protocol” and/or “type” fields can be used to indicate a BFD packet
- BFD with embedded Src/Dst Info
  - Source and destination address are embedded in the BFD control packet
  - Similar to RFC6428, e.g., Source MEP ID TLV

Example 2:

SFC Trace

based on draft-ooamdt-rtgwg-gap-analysis-02

SFC Traceroute

```
sff_client.py --remote-sff-ip 10.0.1.41 --remote-sff-port
4789 --sfp-id 22 --sfp-index 255 --trace-req --num-trace-
hops 3

Sending Trace packet to Service Path and Service Index:
(22, 255)
Trace response...
Service-hop: 0. Service Type: dpi, Service Name: SF1,
Address of Reporting SFF: ('10.0.1.41', 4789)
Service-hop: 1. Service Type: firewall, Service Name: SF4,
Address of Reporting SFF: ('10.0.1.42', 4789)
Service-hop: 2. Service Type: napt44, Service Name: SF5,
Address of Reporting SFF: ('10.0.1.43', 4789)
Trace end
```

In-band Telemetry Probe

(Yes, we need this too for)

What is this ?

- At some moment in time we would like to know the exact network state of the data path traffic
  - Example: ECMP next hop
  - Real time control feedback loop
    - Like ECN, XCP, RCP or utilization aware routing (CONGA)
  - Real time of network event detections
  - OAM
- We would like to get this info without control plan intervention.

How it works

- Traffic source (Application, NIC,TOR, etc.) will embed a request inside the data packet generate special probe packet
- Destination nodes, Sink, receive the instructions and possibly report the collected results of those instructions to an application or a controller
- Allowing the traffic Sink to monitor the exact state of the network
- The request and response have to be send over an Overlay network. Documented use-cases for some overlays already exist:
  - NSH
  - Geneve
  - Vxlan GPE

Telemetry Next steps

- Update the draft to cover following aspects of the gaps
  - Use Case
  - Control plane
  - Data plane/Encapsulations

Performance Measurement OAM

- Active:
  - Loss and Delay Measurement in MPLS networks, RFC 6374
  - One/Two-way Active Performance Measurement Protocol(s), RFC 4656/RFC 5357
- Passive:
  - Alternate Marking Method
- Conclusions:
  - RFC 6374 can be used as foundation of active PM OAM in overlay networks. The YANG data model of the packet loss and delay measurement based on RFC 6374 can improve control and increase operational value of active performance measurement in overlay networks.
  - Alternate Marking Method being proposed as Passive OAM in BIER and can be used in NVO3 and SFC, given supported by overlay network encapsulation

OAM Header for use in Overlay Networks

draft-ooamdt-rtgwg-ooam-header-00

Overlay OAM Header

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| V | Msg Type  |     Flags     |           Length              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                   OOAM Control Packet                         ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

where:

- V - two bits long field indicates the current version of the Overlay OAM Header. The current value is 0
- Msg Type - six bits long field identifies OAM protocol, e.g. Ping or BFD
- Flags - eight bits long field carries bit flags that define optional capability and thus processing of the OOAM control packet, e.g. optional timestamping
- Length - two octets long field that is length of the OOAM control packet in octets

Timestamp

The idea comes from work on Residence Time Measurement, interest in measuring Delay and Delay Variation in addition to other Active OAM measurements

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  QTF  |  RTF  |                  Reserved                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          Timestamp 1                          |
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                                                               ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          Timestamp 4                          |
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

where:

- QTF - Querier timestamp format, e.g. NTP or IEEE-1588v2
- RTF - Responder timestamp format
- Timestamp 1-4 - 64-bit timestamp values

On-demand CC/CV for Overlay Networks

draft-ooamdt-rtgwg-demand-cc-cv-00

Overlay OAM Ping format

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|         Version Number        |         Global Flags          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Message Type  |   Reply mode  |  Return Code  | Return S.code |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                        Sender's Handle                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                         Sequence Number                       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                              TLVs                             ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Overlay OAM Ping format (cont.)

Where:

- the Version reflects the current version
- the Global Flags is a bit vector field
- The Message Type filed reflects the type of the packet. Value TBA2 identifies Echo Request and TBA3 - Echo Reply
- the Reply Mode defines the type of the return path requested by the Sender of the Echo Request
- Return Codes and Subcodes can be used to inform the sender about result of processing its request
- the Sender's Handle is filled in by the sender, and returned unchanged by the receiver in the echo reply
- The Sequence Number is assigned by the sender and can be, for example, used to detect missed replies
- TLVs (Type-Length-Value tuples) have the two octets long Type field, two octets long Length field that is length of the Value field in octets

To be addressed

- Sender ID to be used for out-of-band, i.e. IP network, Echo Reply
- Source MEP ID (OOAM Domain ID + MEP ID) for Connectivity Verification
- Specification of Return Path Control Channel
- and more …

Conclusion

- We need your review and comments!
- We are ready to start the protocol work –> What’s missing from the Requirements or Gap Analysis?