Simple TWAMP (STAMP) Extensions for Segment Routing Networks

draft-gandhi-ippm-stamp-srpm-01

Rakesh Gandhi - Cisco Systems (rgandhi@cisco.com) - Presenter
Clarence Filsfils - Cisco Systems (cfilsfil@cisco.com)
Daniel Voyer - Bell Canada (daniel.voyer@bell.ca)
Mach(Guoyi) Chen - Huawei (mach.chen@huawei.com)
Bart Janssens - Colt (Bart.Janssens@colt.net)

Agenda

- Requirements and Scope
- Summary of Extensions
- Next Steps

Requirements and Scope

Requirements:

- Support in-band Delay Measurement
- Support stand-alone direct-mode Loss Measurement
- Support links and SR paths

Goals:

- Eliminate per session provisioning
- Stateless on session-reflector
- Support very high scale for number of sessions and faster detection interval

Scope:

- STAMP [RFC 8762]
- STAMP TLVs [draft-ietf-ippm-stamp-option-tlv]

Review Comments

1. Draft status:
   a) Draft defines extensions for RFC 8762 - STAMP
      - Updates RFC 8762 due to new field (control code) in the message
2. Extensions specific to SR?
3. Editorial
   a) Define Abbreviations (BSID, SRH, HMAC-SHA)
   b) Use Test packet, Session-Sender, Session-Reflector terms
   c) Show entire test packet with session-sender control code field
   d) Indicate new packet loss message is for direct-mode loss
   e) Move Receive Counter and other Response message fields to Section 4.1 from 3.2
      - Explain how the counters and sequence numbers are used to do loss measurement

STAMP - Session-Sender Control Code Field

In a Query: Session-Sender Control Code

0x0: Out-of-band Response Requested. This is the existing behavior.

0x1: In-band Response Requested. Indicates that this query has been sent over a bidirectional path and the probe response is required over the same path in reverse direction.

0x2: No Response Requested.

With this, the Session-Reflector node does not require any additional state for PM

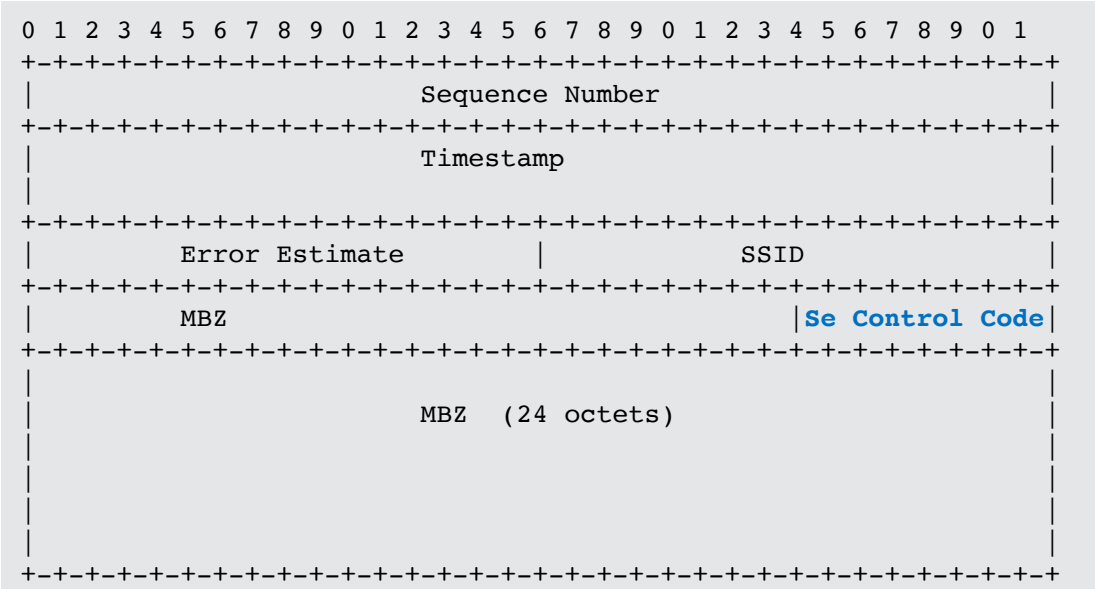


Figure: Session-Sender Control Code in STAMP DM Message

STAMP - Session-Sender Control Code Field

- Two-way measurement mode
  - Reflector needs to send reply on the same link (**in-band**) (symmetric delay on forward and reverse link)
- No way of knowing if one-way or two-way mode from the STAMP test packet
- Not scalable to configure for each (session id, source-address) on session-reflector (can have an order of 1K links)
  - Cannot always send reply on the same incoming interface as the STAMP test packet reply may need to be IP routed

STAMP - Return Path TLV

Return Path TLV (value TBA2):

Sub-TLVs Types:

- Type (value 1): Return Address. Target node address of the response; different than the Source Address in the query
- Type (value 2): SR-MPLS Label Stack of the Reverse SR Path
- Type (value 3): SR-MPLS Binding SID [draft-ietf-pce-binding-label-sid] of the Reverse SR Policy
- Type (value 4): SRv6 Segment List of the Reverse SR Path
- Type (value 5): SRv6 Binding SID [draft-ietf-pce-binding-label-sid] of the Reverse SR Policy

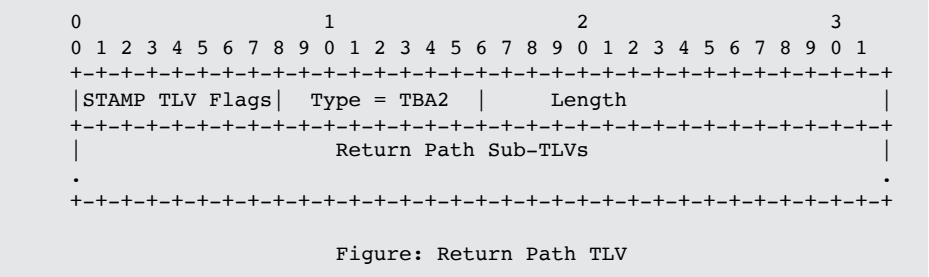


Figure: Return Path TLV

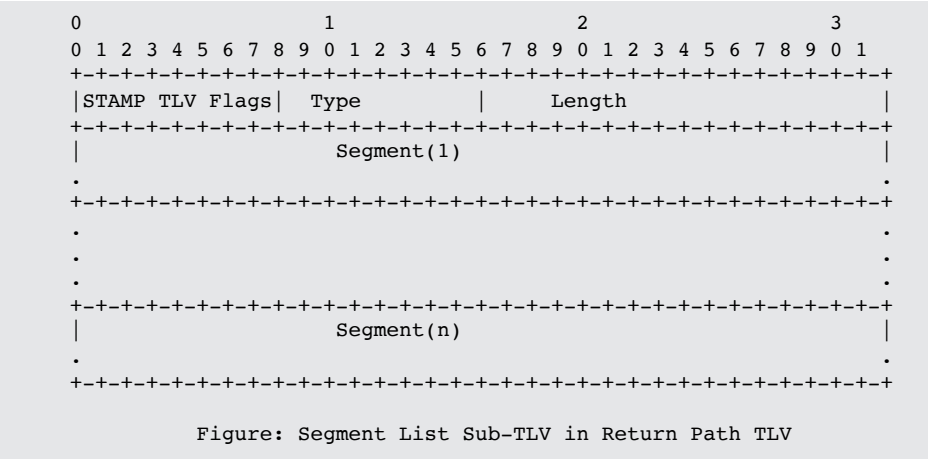


Figure: Segment List Sub-TLV in Return Path TLV

STAMP - Return Path TLV

- For Bidir SR Policy, reply test packet needs to be sent (**in-band)** on the reverse SR Policy
- Bidir SR Path (forward and reverse) dynamically computed using CSPF by the head-end node
  - Path can change often based on topology change, link/node failure in the network, etc.
- No signaling in SR, possible to use PCE
- Need per session state on session-reflector node to store reverse paths (each session-id, source-address) – order of 10Ks SR Policy (that can have active and standby candidate-paths and each can have multiple segment-lists)
- In SR, state is in the packet

STAMP - Destination Node Address TLV

Destination Node Address TLV (value TBA1):

- Indicates the address of the intended recipient node of the query message.
- The Session-Reflector node **MUST NOT** send response if it is not the intended destination node of the query.
- Useful when query is sent with 127/8 destination address (e.g. sweeping ECMP paths).

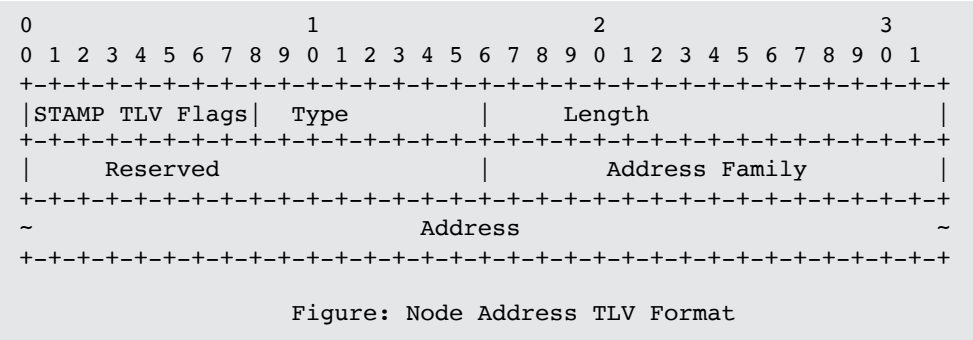
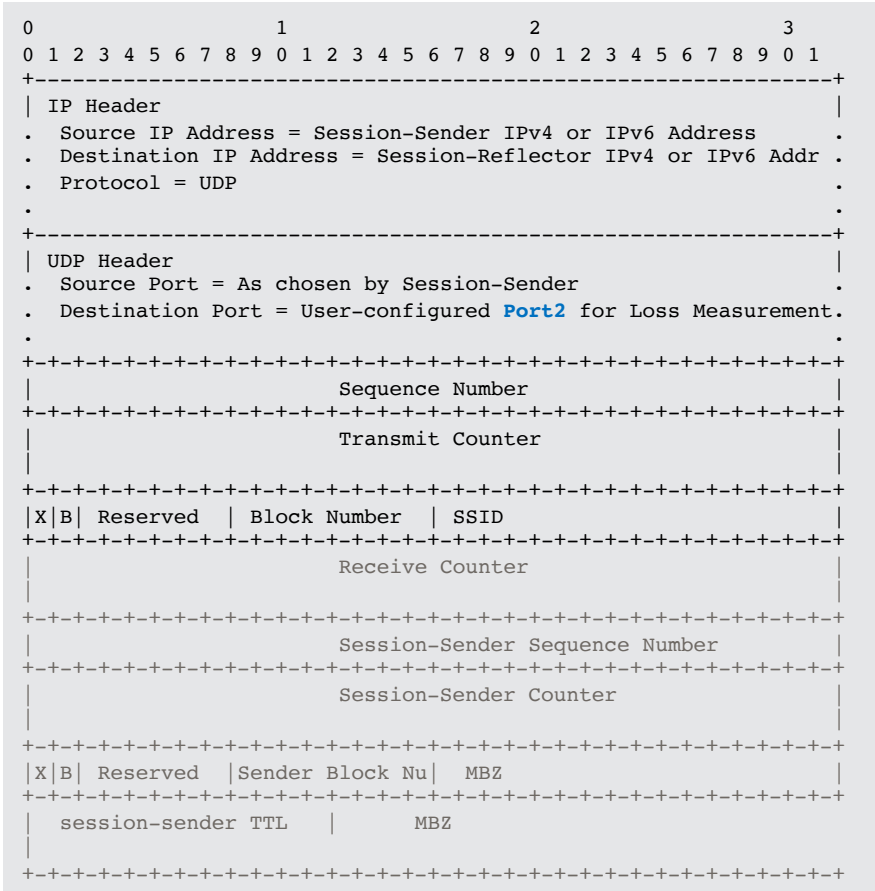


Figure: Node Address TLV Format

STAMP - Stand-alone Direct-mode LM Message Format

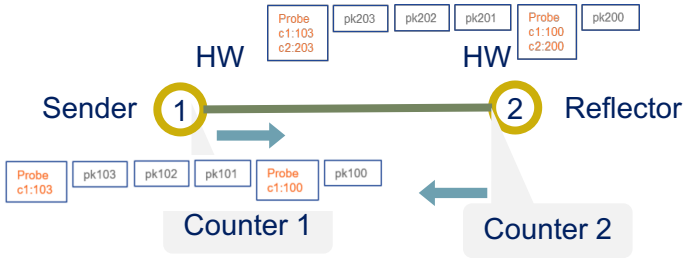
- Stand-alone Direct-mode Loss Measurement (LM) query and response messages defined
  - Hardware efficient counter-stamping
    - Well-known locations for transmit and receive traffic counters
  - Stand-alone LM message, not tied to DM
- Direct-mode LM message format is also defined for authenticated mode
- User-configured destination UDP **Port2** is used for identifying direct-mode LM probe packets
- Does not modify existing STAMP (which is for DM) procedure as different destination UDP port is used for direct-mode LM



STAMP - Stand-alone Direct-mode LM Message Format

- Sender:
  - Hardware needs to load the test packet in write-able memory which is limited
    - With LM TLV, counter may not be at fixed location
    - With LM TLV, counter also deeper into the test packet at offset (Eth 18, IPv6 40, UDP 8, STAMP 44, TLV Type 4, Total = 114 Byte)
    - Also need to include other Encaps / headers in offset
  - Hardware also not capable to write both TS and Counter in the same test packet
  - Hardware also not capable to recompute UDP checksum
- Reflector:
  - Some test packets received from one session-sender with base test packet and some with LM TLV, hence need to parse EVERY received test packet to check if direct-mode loss TLV is present before punting the packet
  - Hardware needs to punt with receive TS or receive Counter
  - Hardware also not capable to punt with both TS and Counter for the same test packet
- Separate UDP port + LM message format eliminate the complexity in Hardware
  - Counter at fixed location - offset (no TLV, Eth 18, IPv6 40, UDP 8, Seq 4, Total = 70 Byte)

Link Direct-mode Loss Measurement – Inline Counter-stamping in Hardware



- Advertise extended TE metrics – link loss percentage
  - RFC 8570 (IS-IS)
  - RFC 7471 (OSPF)
  - RFC 8571 (BGP-LS)

- One Way Packet Loss %
  = 100* (((C1(t) – C1(t-1)) – (C2(t) – C2(t-1)) / (C1(t) – C1(t-1))
  = 100* (((103 – 100) – (203 – 200)) / (103 – 100))
  = 0
- Hardware Counters – counter-stamping in hardware

```
 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        |     Length    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|A|  RESERVED   |                    Link Loss                  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Next Steps

- Welcome your comments and suggestions
- Requesting WG adoption

Thank you