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 Measurement Test Packet for Packet Loss
- Support links and SR paths

Goals:

- Eliminate per session provisioning on Session-Reflector
- No control-channel signaling for sessions
- Support hardware implementation - 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 test packet
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 packet is for direct measurement
   e) How packet loss calculated
   f) Details on direct measurement test packet usage
   g) Move Receive Counter and other Reply test packet 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

Session-Sender Control Code

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

0x1: In-band Reply Requested.
Indicates that this test packet has been sent over a bidirectional path and the reply is required over the same path in the reverse direction.

0x2: No Reply Requested.

```
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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                        Sequence Number                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                        Timestamp                              |
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|         Error Estimate        |             SSID              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|         MBZ                                   | Control Code  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
|                        MBZ  (24 octets)                       |
|                                                               |
|                                                               |
|                                                               |
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Session-Sender Control Code in Test Packet

STAMP - Session-Sender Control Code Field - Usage

- Two-way measurement mode for links
  - Reflector needs to send reply on the same link (**in-band**) (symmetric delay on forward and reverse link)
  - Link can be LAG (bundle) member
- No way of knowing if one-way or two-way mode from the received 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 for the reply; different than the Source Address in the test packet
- 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

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|STAMP TLV Flags|  Type = TBA2  |         Length                |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Return Path Sub-TLVs                      |
.                                                               .
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+

                     Figure: Return Path TLV
```

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|STAMP TLV Flags|  Type         |         Length                |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Segment(1)                                |
.                                                               .
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
.                                                               .
.                                                               .
.                                                               .
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Segment(n)                                |
.                                                               .
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+

          Figure: Segment List Sub-TLV in Return Path TLV
```

STAMP - Return Path TLV - Usage

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

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|STAMP TLV Flags|     Type      |           Length              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|      Reserved                 |       Address Family          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                           Address                             ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+

                Figure: Node Address TLV Format
```

Stand-alone Direct Measurement Test Packet for Packet Loss

- Stand-alone Direct Measurement test packet defined
  - Hardware efficient counter-stamping
    - Well-known locations for transmit and receive traffic counters
  - Block number of the counters for alternate marking method [RFC 8321]
  - Traffic class of the counters for per class packet loss
- Direct Measurement test packet is also defined for authenticated mode
- User-configured destination UDP **Port2** is used for identifying direct measurement test packets
- Does not modify the existing STAMP procedure as different destination UDP port is used for direct measurement test packets
- Sequence Numbers allow to detect test packet loss, and connectivity loss
- Flags
  - X set to 1 for 64-Bit Counter, set to 0 for 32-Bit Counter
  - B set to 1 for Byte Counter, set to 0 for Packet Counter
  - T set to 1 for Sender-DSCP scoped Counter

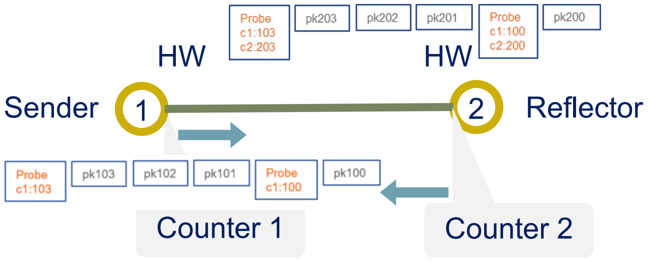
```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+---------------------------------------------------------------+
| IP Header                                                     |
.  Source IP Address = Session-Sender IPv4 or IPv6 Address      .
.  Destination IP Address = Session-Reflector IPv4 or IPv6 Addr .
.  Protocol = UDP                                               .
.                                                               .
+---------------------------------------------------------------+
| UDP Header                                                    |
.  Source Port = As chosen by Session-Sender                    .
.  Destination Port = User-configured Port2 direct measurement  .
.                                                               .
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                        Sequence Number                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                        Transmit Counter                       |
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|X|B|T| Send-DSCP | Block Number| SSID                          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                        Receive Counter                        |
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                        Session-Sender Sequence Number         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                        Session-Sender Counter                 |
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|X|B|T| Reserved  |Send-Block Nu|  MBZ                          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Ses-Sender TTL|  MBZ                                          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+

    Figure: Session-Reflector Direct Measurement Test Packet
```

Direct Measurement TLV vs. Direct Measurement Test Packet

STAMP Direct Measurement TLV	Direct Measurement Test Packet
1. Complex to implement counter collection in hardware for Sender TX, Reflector RX and Reflector TX to detect packet loss. – Session-reflector hardware to parse STAMP TLVs in receive packets to decide if receive counter to be punted to the control-plane – Reply test packet with transmit counter NOT at the same location – needed for hardware counter-stamping (like STAMP timestamp) – Hardware needs to write both timestamp and counter in the injected packet – not capable	1. Suitable for collecting data packet counters from hardware – inline counter-stamping (for P2P connections) – Reply test packet with transmit counter at the same location – important property for hardware counter-stamping (like STAMP timestamp)
2. Need STAMP TLV processing in hardware	2. No TLV processing in hardware
3. Counter not at fixed location due to TLVs	3. Counter at fixed location, well-known location for SRv6 network programming, needed for hardware implementation
4. Counter location deeper into the packet (Eth 18, IPv6 40, UDP 8, STAMP 44, TLV Type 4, Total = 114 Byte) – load into write-able memory	4. Counter location earlier into the packet (Eth 18, IPv6 40, UDP 8, Seq 4, Total = 70 Byte)
5. Direct Measurement TLV supports **32-bit packet** counters	5. Direct Measurement Test Packet supports **32-bit packet and byte** counters and **64-bit packet and byte** counters
6. For control-plane implementation, how can we measure packet loss without alternate marking method (block number)?	6. Direct Measurement Test Packet identifies the block number of the counters - used for alternate marking method (RFC 8321) for control-plane based packet loss
7. Does not support per-traffic class direct measurement (DSCP TLV processing not specified for Counters)	7. Per traffic-class counter collection (per traffic-class loss measurement) (e.g., drop best effort traffic)

Link Loss Direct 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
- Traffic 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                      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Alternate Marking Method for Packet Loss

- RFC 8321 - Alternate-Marking Method for Passive and Hybrid Performance Monitoring
- draft-ietf-mpls-sfl-framework - Synonymous Flow Label Framework
- Control plane-based packet loss with distributed forwarding LCs, using block number of the counters

```
   A: packet with A coloring
   B: packet with B coloring

             |           |           |           |           |
             |           |    Traffic Flow       |           |
    ---------------------------------------------------------------->
     BBBBBBB AAAAAAAAAAA BBBBBBBBBBB AAAAAAAAAAA BBBBBBBBBBB AAAAAAA
    ---------------------------------------------------------------->
       ...   |  Block 5  |  Block 4  |  Block 3  |  Block 2  |  Block 1
             |           |           |           |           |

                         Figure 2: Traffic Coloring
```

Packet Loss Calculation

```
        C1                  C2
         /                    \
+--------+     Test Packet     +--------+
|        | - - - - - - - - - ->|        |
|   R1   |=====================|   R3   |
|        |<- - - - - - - - - - |        |
+--------+     Test Reply      +--------+
         \                    /
          C4                 C3

Session-Sender               Session-Reflector

           Reference Topology
```

- Using the Counters from the n^{th} and $(n-1)^{th}$ test packets.
- Counters from the same color used in the computation.
- One-way receive packet loss[n-1, n] = (C2[n] - C2[n-1]) - (C1[n] - C1[n-1])
- Two-way receive packet loss[n-1, n] = (C4[n] - C4[n-1]) - (C3[n] - C3[n-1])
  + (C2[n] - C2[n-1]) - (C1[n] - C1[n-1])

Next Steps

- Welcome your comments and suggestions
- Requesting WG adoption

Thank you

STAMP Test Packet with Direct Measurement TLV

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                        Sequence Number                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          Timestamp                            |
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|         Error Estimate        |                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+                               |
|                                                               |
|                                                               |
+                                                               +
|                         MBZ (30 octets)                       |
+                                                               +
|                                                               |
|                                                               |
+                                                               +
|                                                               |
+                                                               +
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                                                               ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|STAMP TLV Flags|    Type       |         Length                |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|             Session-Sender Tx counter  (S_TxC)                |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|             Session-Reflector Rx counter  (R_RxC)             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|             Session-Reflector Tx counter  (R_TxC)             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+

           Figure: Session-Sender Test Packet 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                        Sequence Number                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          Timestamp                            |
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|         Error Estimate        |           MBZ                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          Receive Timestamp                    |
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                 Session-Sender Sequence Number                |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                  Session-Sender Timestamp                     |
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Session-Sender Error Estimate |           MBZ                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|Ses-Sender TTL |                  MBZ2                         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                                                               ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|STAMP TLV Flags|    Type       |         Length                |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|             Session-Sender Tx counter  (S_TxC)                |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|             Session-Reflector Rx counter  (R_RxC)             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|             Session-Reflector Tx counter  (R_TxC)             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+

           Figure: Session-Reflector Test Packet Format
```