IPPM
[IP Performance Measurement]

IETF 107 Virtual / Wednesday 1 April 2020

Tommy Pauly and Ian Swett, Co-Chairs
Bill Cerveny and Brian Trammell, Outgoing Co-Chairs

Note Well

- This is a reminder of IETF policies in effect on various topics such as patents or code of conduct. It is only meant to point you in the right direction. Exceptions may apply. The IETF's patent policy and the definition of an IETF "contribution" and "participation" are set forth in BCP 79; please read it carefully.
- As a reminder:
  - By participating in the IETF, you agree to follow IETF processes and policies.
  - If you are aware that any IETF contribution is covered by patents or patent applications that are owned or controlled by you or your sponsor, you must disclose that fact, or not participate in the discussion.
  - As a participant in or attendee to any IETF activity you acknowledge that written, audio, video, and photographic records of meetings may be made public.
  - Personal information that you provide to IETF will be handled in accordance with the IETF Privacy Statement.
  - As a participant or attendee, you agree to work respectfully with other participants; please contact the ombudsteam if you have questions or concerns about this.
- Definitive information is in the documents listed below and other IETF BCPs.
  For advice, please talk to WG chairs or ADs:
  - BCP 9 (Internet Standards Process)
  - BCP 25 (Working Group processes)
  - BCP 25 (Anti-Harassment Procedures)
  - BCP 54 (Code of Conduct)
  - BCP 78 (Copyright)
  - BCP 79 (Patents, Participation)
  - https://www.ietf.org/privacy-policy/ (Privacy Policy)

Meeting Management

WebEx Meeting: https://ietf.webex.com/meet/ippm

Jabber Room: ippm@jabber.ietf.org

Add your name to the “Blue Sheet” in Etherpad:

https://etherpad.ietf.org:9009/p/notes-ietf-107-ippm

To speak, use WebEx chat and comment “q+” and “q-” to Everyone to enter and leave the queue

IPPM Agenda / IETF 107 Virtual

Wednesday 1 April 2020 — 15:00-16:30 UTC

Working Group Documents

Time	Length	What	Who
15:00	10m	Welcome, Note Well, Agenda, Status	Chairs
15:10	15m	draft-morton-ippm-capacity-metric-method	A. Morton
15:25	5m	dratf-ietf-ippm-stamp-option-tlv	G. Mirsky
15:30	30m	IOAM update	F. Brockners

IPPM Agenda / IETF 107 Virtual

Wednesday 1 April 2020 — 15:00-16:30 UTC

Other work

Time	Length	What	Who
16:00	10m	draft-geib-ippm-connectivity-monitoring	R. Geib
16:10	10m	draft-song-ippm-postcard-based-telemetry	H. Song
16:20	5m	draft-cfb-ippm-spinbit-measurements	B. Fabio
16:25	5m	draft-gandhi-spring-twamp-srpm	R. Gandhi

IPPM Chair Updates

Thanks to Brian Trammell and Bill Cerveny for eight years of chairing IPPM!

IPPM Document Updates

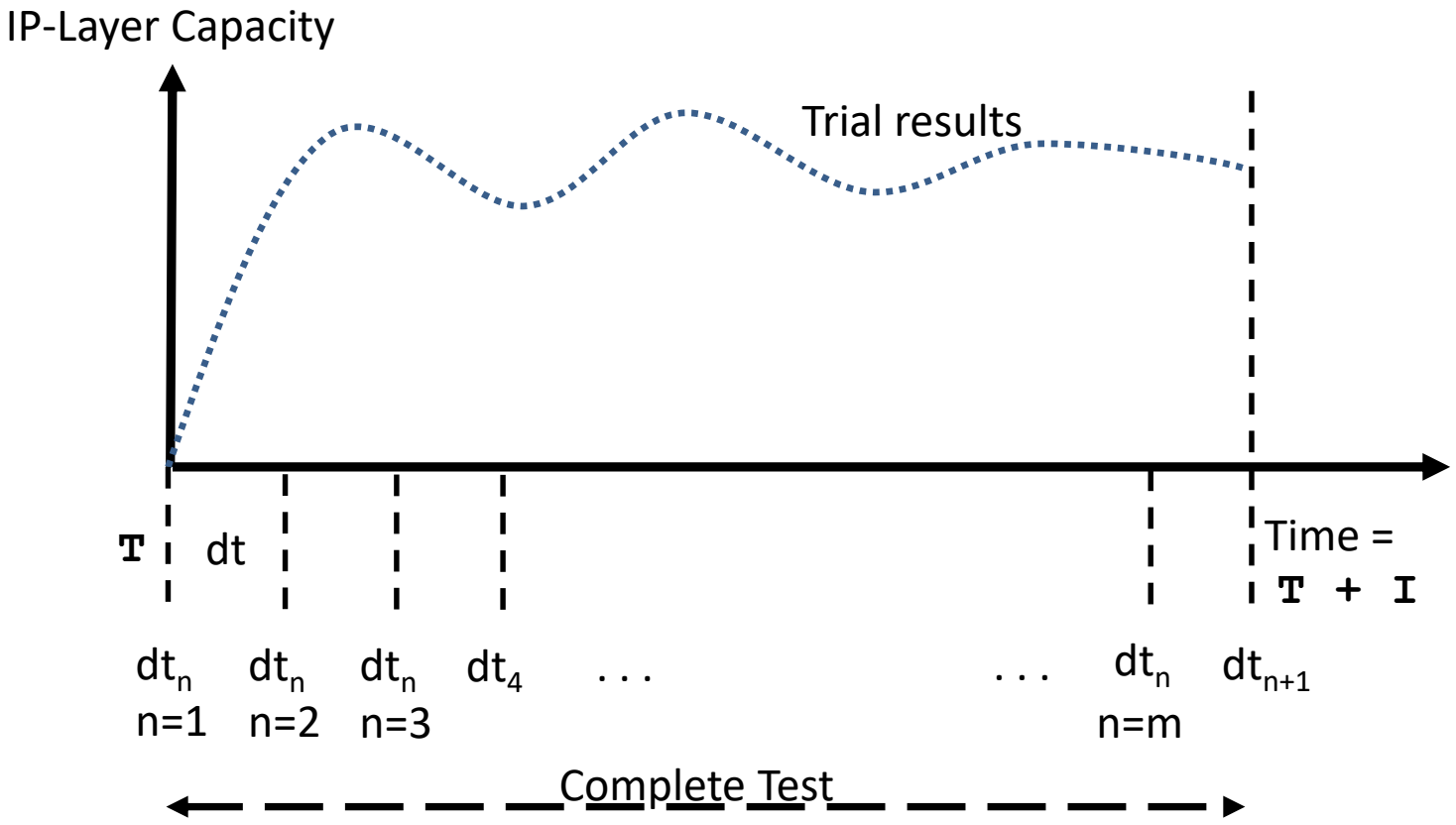
- STAMP published as RFC 8762
- Metric Registry and Initial Registry in RFC Editor queue
- Multipoint Alt Mark in RFC Editor queue
- draft-ietf-ippm-ioam-data went through WG Last Call and received updates
- draft-ietf-ippm-capacity-metric-method and draft-ietf-ippm-ioam-direct-export newly uploaded as WG documents

Metrics and Methods for IP Capacity

draft-ietf-ippm-capacity-metric-method-01

A. Morton, R. Geib, L. Ciavattone

Receiver Rate Measurement



Define the Metric

- Maximum IP-Layer Capacity (incl headers + UDP payload)
- One of many metrics that could be defined
- Def. in Words and an Equation (with variables explained)

$$\text{Maximum\_C}(T,I,\mathbf{PM}) = \frac{\max_{[T,T+I]} \left(n0[\mathbf{dtn},\mathbf{dtn+1}] \right)}{dt}$$

where:

T < ------ Measurement Interval ------> T+I

| | | | | | | | | | |

dtn=1 2 3 4 5 6 7 8 9 10 n+1

m=10

sub-intervals

IPPM Draft Status

- Many-many comments and reviews have resulted in a very complete draft.
  - New Reviews from ETSI STQ MOBILE
  - Four New Members of ITU-T SG12 (testing co's)
  - Testing from various volunteers
- Key topics updated/added in 01:
  - Measurement Considerations
  - Reporting Formats

8.3 Meas. Considerations (new)

Conditions which might be encountered during measurement, **where packet losses may occur independently** from send rate:

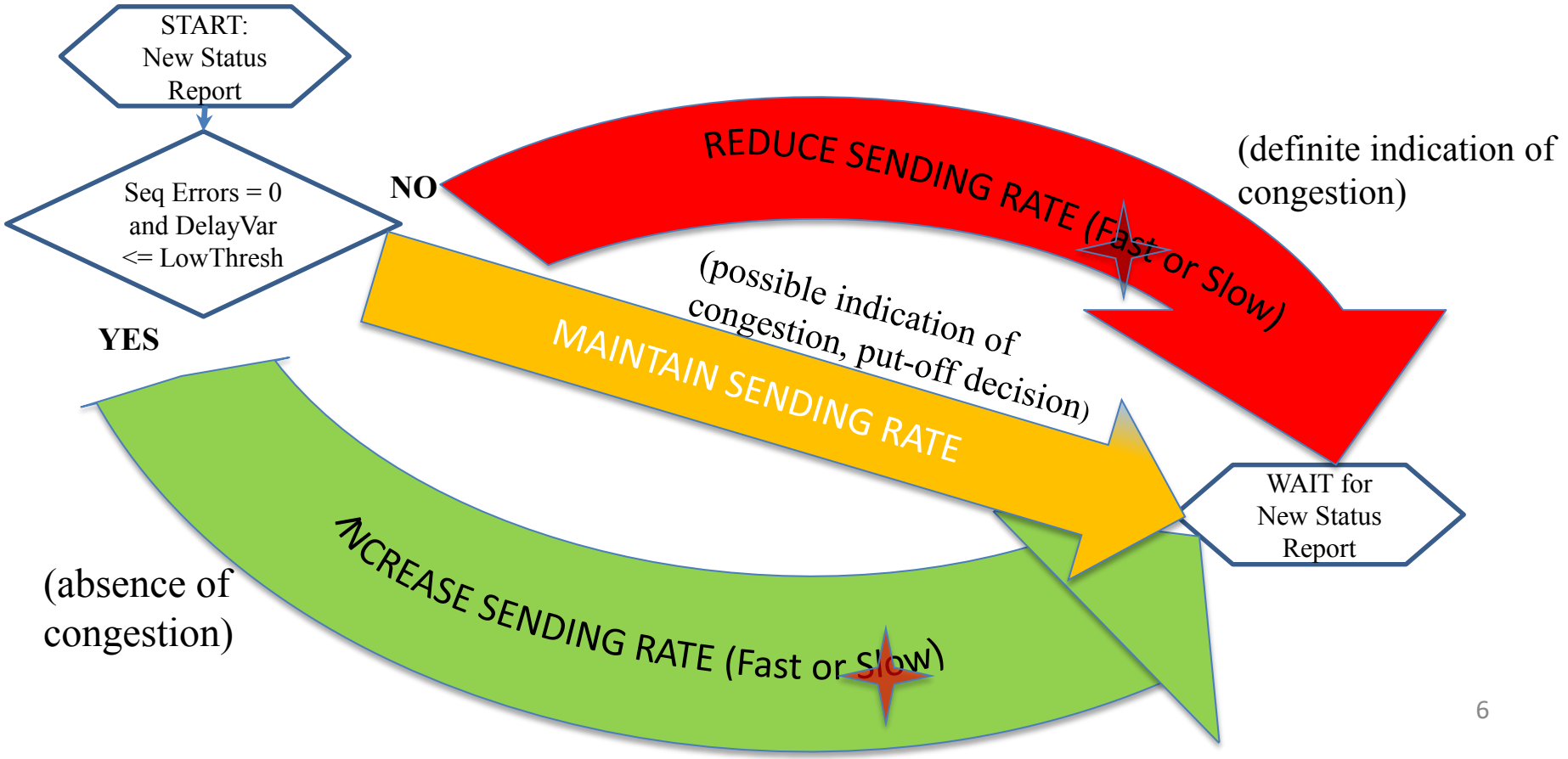
1. Congestion of an interconnection or backbone interface may appear as packet losses distributed over time in the test stream, due to much higher rate interfaces in the backbone.
2. Packet loss due to use of Random Early Detection (RED) or other active queue management.
3. There may be only small delay variation independent of sending rate under these conditions, too. THIS IS A "TELL"
4. Persistent competing traffic on measurement paths that include shared media may cause random packet losses in the test stream.

It is possible to mitigate these conditions… but try locating measurement points as close as possible, first!

8.3 Meas. Considerations (new)

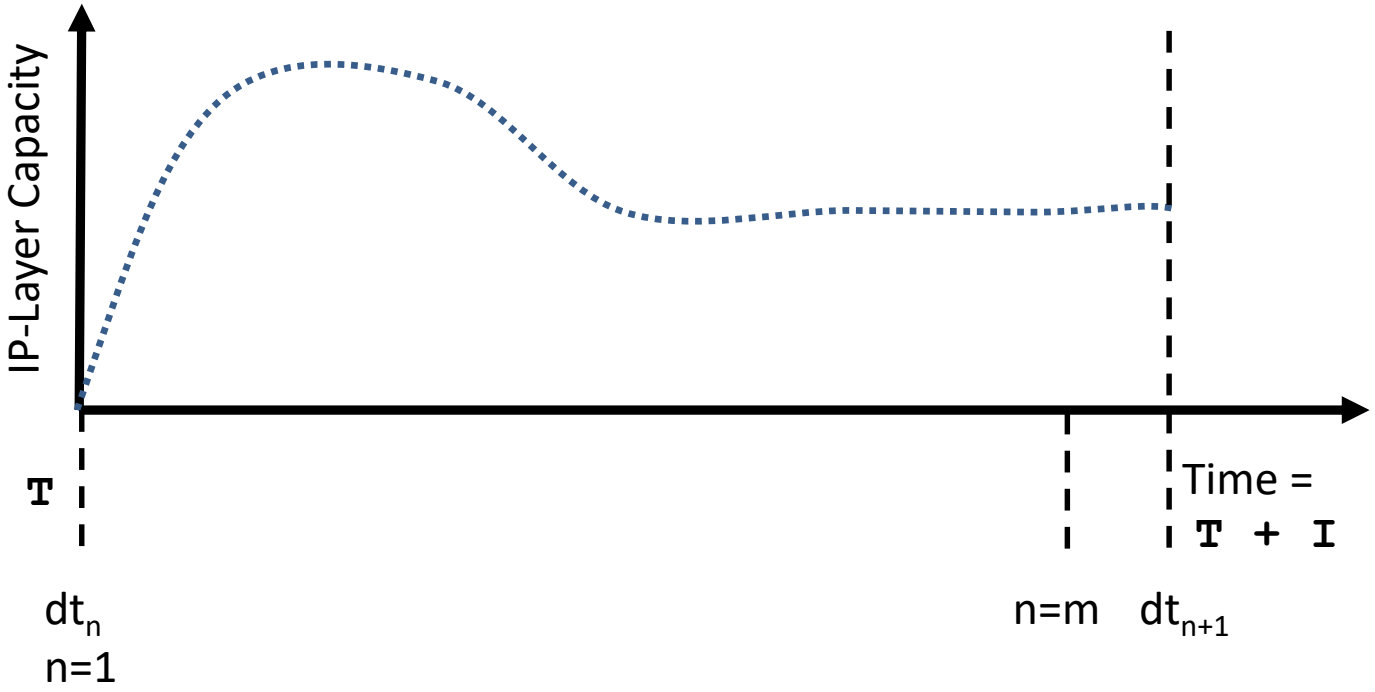
where packet losses occur independently from send rate:

Mitigate using parameters of search alg. described in Section 8.1 (tuning specific parameters, **more flexibility than typical CCA**).



Results Reporting Considerations

- "Turbo-mode" concept (Matt Mathis' testing)
- Report separate results for repeatable modes



- Other modes may be encountered (repeatable?)
- Radio constellations, Cellular modes, weather

9. Reporting Format Elements (Others?)

The Singleton IP-Layer Capacity results SHOULD be accompanied by the context under which they were measured.

- timestamps

  (especially the time when the maximum was observed in dtn)
- source and destination (by IP or other meaningful ID)
- other inner parameters of the measurement (Section 4)
- outer parameters, such as “performed in motion" or other factors belonging to the context of the measurement
- result validity (indicating cases where the process was somehow interrupted or the attempt failed)
- a field where unusual circumstances could be documented
- a field for "ignore/mask out" purposes in further processing

Standards High-Level Status: IP-Layer Capacity Metric and Meas.

- ITU-T Study Group 12 - **Approved**
  - Question 17 on Packet Network Performance the Metric and Method of Measurement to Rec. **Y.1540 - 2019 (Annexes A and B)**
    - Considerable background (test results; research) in Appendices X thru XIII
- ETSI TC Speech and Multimedia Transmission Quality (STQ)
  - **Approved** the Metric in **TS 103 222 Part 2** on High Speed Internet KPIs
  - Reference to Rec Y.1540 for all other material
- Broadband Forum (BBF) – **Project Approved**: **WT-471**
  - Standardize the identical Metric and Methods with additional details on Measurement Points and Information Model for control and reporting. *First Ballot in May, 2020, next meeting in June.*
- IETF IP Performance Measurements (IPPM) Working Group
  - **Internet Draft** **Adopted** by WG, adding Metric details, Measurement Considerations, and Results presentation formats

Next Steps

- Post-WG Adoption work:
  - Harmonization: Keep-up with parallel efforts to ensure IPPM's expertise incorporated elsewhere
  - Reach Consensus soon, start protocol support
- Additional Volunteers for Review
  - Trigger more reviews with a WG Last Call?

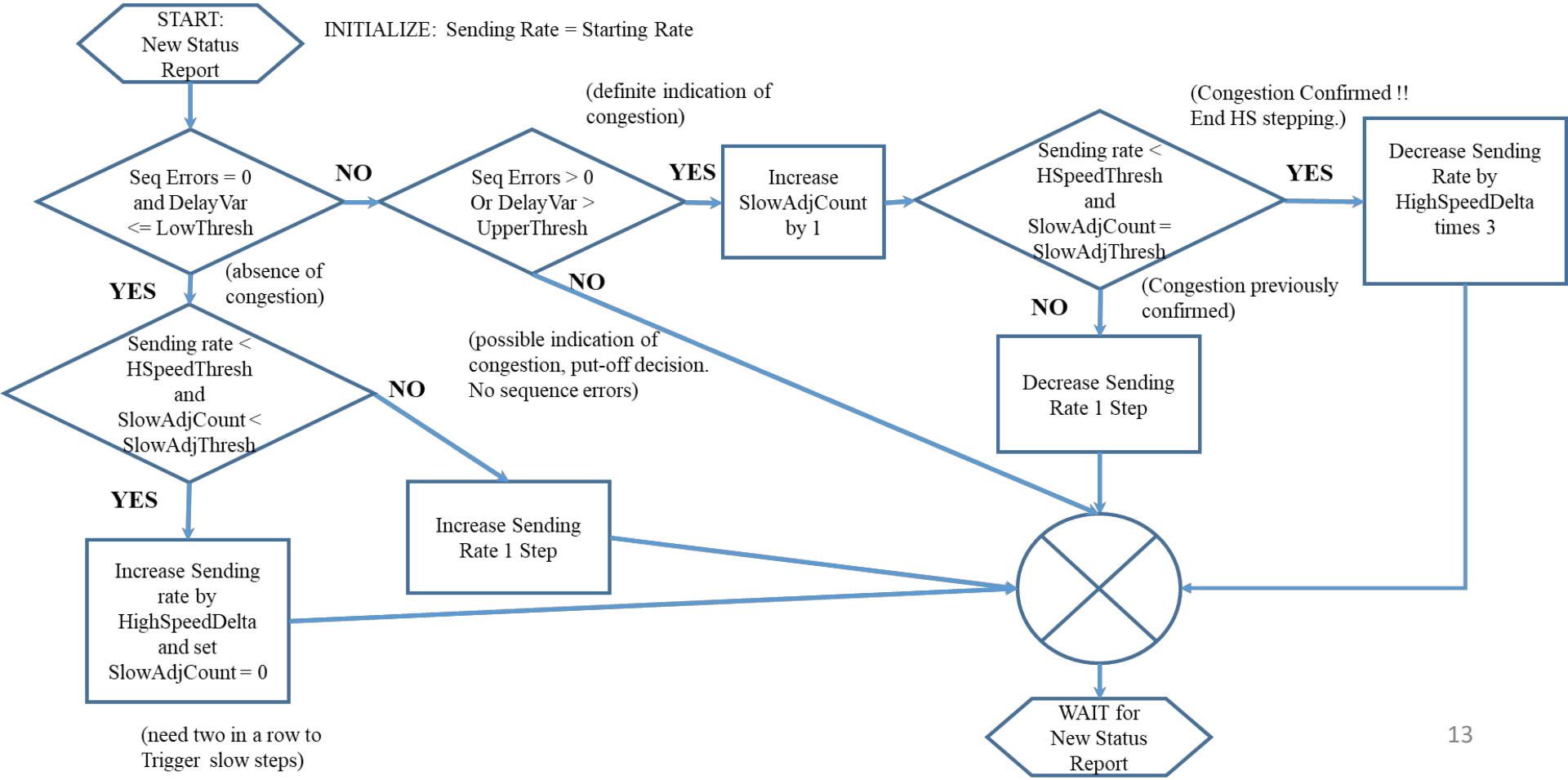
REFERENCES

- Hackfest 106 Slides: Test Results
- Hackfest 105 Slides: Test Results
- Liaisons from ITU-T SG 12 and ETSI TC STQ – see email for links, or
- https://datatracker.ietf.org/liaison/1645/
- https://datatracker.ietf.org/liaison/1643/
- https://datatracker.ietf.org/liaison/1634/
- https://datatracker.ietf.org/liaison/1632/
- More Test results in the Liaison attachments

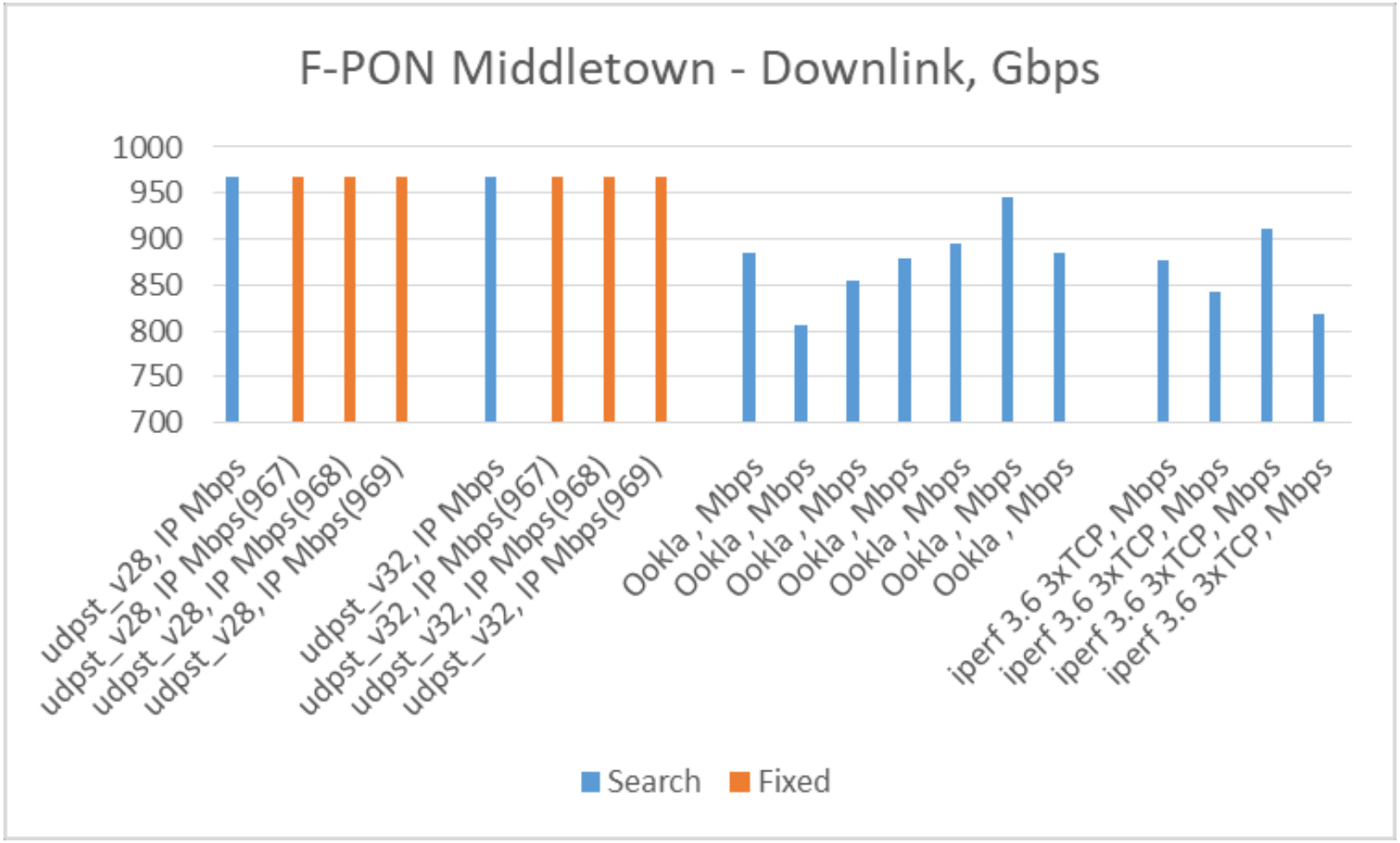
BACKUP

Define the Method

- "PM" is short-hand for the performance constraints on the Load Rate Adjustment Alg.:



Recent Test Results



Udpst and Ookla Web Sockets Clients

Udpst and Ookla Web Sockets Servers

UDP-Speedtest

Middletown, NJ

IETF – Sept List Points raised (and addressed) -1

- @@@@ A clear take-away is that reporting must account for **bimodal** features, if/when measured.
- ++++ Covered in Section 6.6, reporting the Metric
- @@@@ Also, that wide-spread measurements will encounter wide-spread behaviors - testing should continue + expect some evolution.
- ++++ Covered in the Methods of Measurement Section
- @@@@ IMO, many of the above challenges fall on the measurement methodology: allow for traffic & time to initiate an on-demand access.
- @@@@ Also, results depend on the sending stream characteristics; we've known this for a long time, still need to keep it front of mind.
- ++++ both above covered in Methods of Measurement, Considerations.
- @@@@ Max IP-Layer Capacity and RFC 3148 BTC (goodput) **are different** metrics. Max IP-layer Capacity is like the theoretical goal for goodput.
- ++++ Section 1, Intro
- @@@@ This is a big one: when the path we measure is state-full based on many factors, the Parameter "Time of day" when a test starts is not enough info. We need to know the time from the beginning of a measured flow, and how the flow is constructed including how much traffic has already been sent on that flow, because state-change may be based on time or bytes sent or both. See RFC 7312.
- ++++ included in Measurement Considerations

IETF – Sept List Points raised (and addressed) -2

- @@@@ The **Singleton and Statistic** formulations of IPPM's framework RFC 2330 are still valuable in this context, possibly combined with results criteria ("stable" for X singletons, non-arbitrary threshold needed to define "stable").
- ++++ The Singleton, Sample and Statistic for IP Capacity are implemented.
- ---- "stable" needs more discussion, or may be resolved by Qualification below.
- @@@@ Measurements depend on the access network and the use case. Here, the use case is to assess the maximum capacity of the access network, with specific performance criteria used in the measurement.
- ++++ Covered in the Intro.
- @@@@ Goals made clearer in the next draft, if possible.
- ++++ Covered in the Intro.
- @@@@ A qualification measurement for the search result is a subsequent measurement, sending at a **fixed 99.x %** of the Max IP-layer Capacity for I, or an indefinite period. The same Max Capacity Metric is applied, and the Qualification for the result is a sample without packet loss or a growing minimum delay trend in subsequent singletons (or each dt of the measurement interval, I). Samples exhibiting losses or increasing queue occupation require a repeated search and/or test at reduced fixed sender rate for qualification.
- Here, as with any Active Capacity test, the test duration must be kept short. 10 second tests for each direction of transmission are common today. In combination with a fast search method and user-network coordination, the concerns raised in [RFC 6815] are alleviated.
- +++ covered in the method of measurement section, subsection on Measurement Qualification and Verification

https://tools.ietf.org/html/draft-morton-ippm-capacity-metric-method-00

October List Discussion: Matt, Rüdiger, acm (1)

- Summary: Matt is saying (? Subject to confirmation)
  - @@@@@ RTT is a good singleton measurement interval (dt) to avoid "bursts & silence"
  - Use windowed Max of max_rate from BBR (but see our measurements)
- Rüdiger: "Len and acm meas. results show convergence to an LTE receiver bandwidth meas. with limited queuing and no drops."
  - Defaults of dt = 1 second, Δt = 10 sec
  - udpst tool sends feedback measurement at regular intervals = 50 ms
- acm thinking: sub-second rate meas. are more susceptible to the cases described by Matt, and by Joachim Fabini (time-slot service with full link capacity play-out of the queue: LTE, others).
- acm: But no assessment of loss with BBR, QUIC encrypt & aggregates

++++ We've added the defaults above with parameters when they appear, and more discussion in section 8.2

++++ Considerations for testing with parallel flows (sec 8).

++++ Default for the Sending rate measurement interval (sec 7, 0.05 s)

https://tools.ietf.org/html/draft-morton-ippm-capacity-metric-method-01

October List Discussion: Matt and Rüdiger

- It is fairly normal to see packets arrive in back to back packet trains, separated by periods of silence. Half- Duplex, Pkt Aggregation, …
- MM: simplistic meas. of LTE receive rates often see modes at 1Gb/s.
- BBRv2 uses rate measurement per RTT:
  - rtt_sample = delta(timestamp) # 1 RTT
  - rate_sample = delta(total data ACKed)/rtt_sample # one RTT's worth of data
- Effectively: Capacity(t, Δt, n, <no PM>) = n0[dtn-1,dtn]/(dt = RTTn)
- min_rtt and max_rate (used by BBR congestion control) are the windowed (?) max and min of rtt_sample and rate_sample above
- MM: I predict that max of BBR's max_rate will be a more robust and more accurate measure of the short duration maximum rate than anything you can do with UDP (except perhaps QUIC, BBR over UDP).

Simple Two-way Active Measurement Protocol (STAMP) Extensions

draft-ietf-ippm-stamp-option-tlv

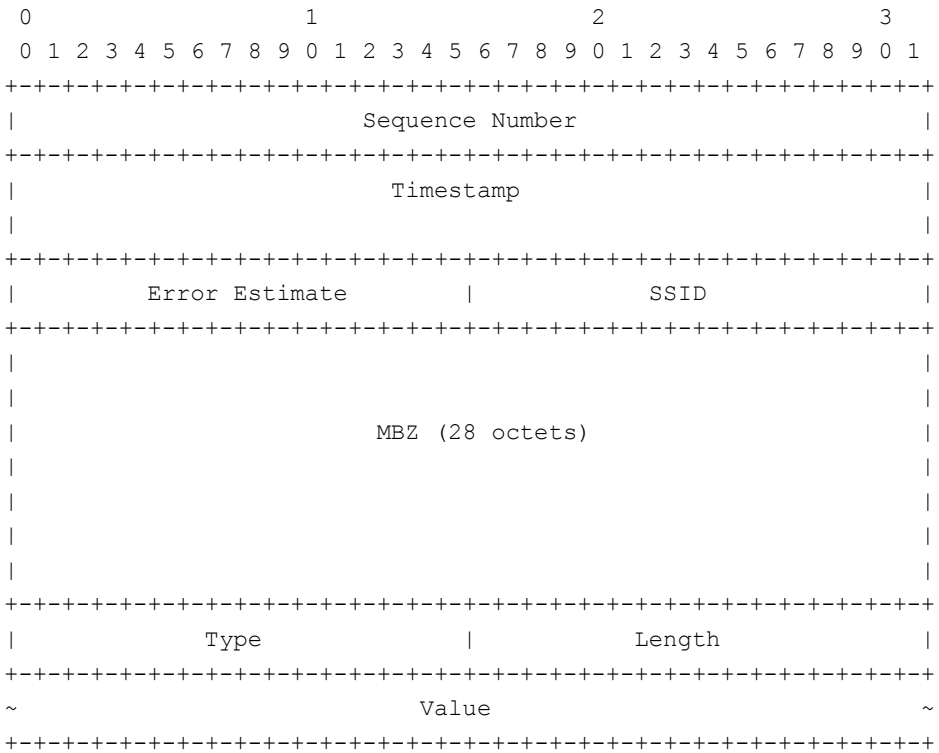
Greg Mirsky gregimirsky@gmail.com
Henrik Nydell hnydell@accedian.com
Ernesto Ruffini eruffini@outsys.org
Richard Foote, footer.foote@nokia.com
Xiao Min xiao.min2@zte.com.cn
Adi Masputra adi@apple.com

IPPM WG Interim meeting, April 2020

Update

- Defined STAMP Session Identifier (SSID)
- Added HMAC TLV
- Clarify STAMP test packet processing
- Location TLV - more space for the Destination Port and the Source Port fields
- Follow-up TLV – re-named the field as Follow-up Timestamp

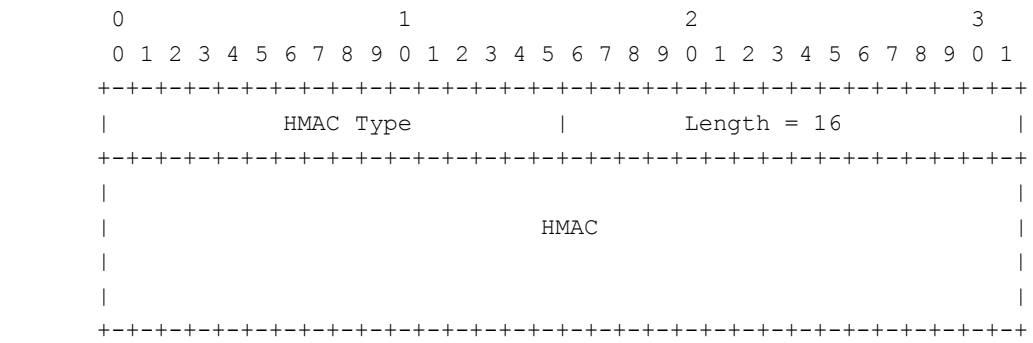
STAMP Session Identifier



- A STAMP Session is identified using 4-tuple (source and destination IP addresses, source and destination UDP port numbers).
- A STAMP Session-Sender MAY generate locally unique STAMP Session Identifier (SSID).
- SSID is two octets long non-zero unsigned integer. A Session-Sender MAY use SSID to identify a STAMP test session.
- If SSID is used, it MUST be present in each test packet of the given test session.
- An implementation of STAMP Session-Reflector that supports this specification SHOULD identify a STAMP Session using the SSID in combination with elements of the usual 4-tuple.
- A conforming implementation of STAMP Session-Reflector MUST copy the SSID value from the received test packet and put it into the reflected packet.

HMACTLV

- The STAMP authenticated mode protects the integrity of data collected in STAMP base packet.
- STAMP extensions are designed to provide valuable information about the condition of a network, and protecting the integrity of that data is also essential.
- The keyed Hashed Message Authentication Code (HMAC) TLV MUST be included in a STAMP test packet in the authenticated mode, excluding when the only TLV present is Extra Padding TLV.



- The HMAC TLV MUST follow all TLVs included in a STAMP test packet, except for the Extra Padding TLV.
- The HMAC TLV MAY be used to protect the integrity of STAMP extensions in STAMP unauthenticated mode.
- HMAC is calculated, as HMAC-SHA-256, over text as the concatenation of all preceding TLVs. The digest then MUST be truncated to 128 bits and written into the HMAC field.
- If HMAC verification by the Session-Reflector fails, then an ICMP Parameter Problem message MUST be generated (with consideration of limiting the rate of error messages). The Code value MUST be set to 0 and the Pointer identifying HMAC Type.
- Both Session-Sender and Session-Reflector SHOULD log the notification that HMAC verification of STAMP TLVs failed. The packet that failed HMAC verification MUST be dropped.

STAMP TLV Processing

- A system that has received a STAMP test packet with extension TLVs MUST validate each fixed-size TLV by verifying that the value in the Length field equals the value defined for the particular type.
- If the values are not equal, the processing of extension TLVs MUST be stopped and the event logged (logging SHOULD be throttled).
- If the system is the Session-Reflector in that test, it MUST send (transmission of ICMP Error messages SHOULD be throttled) the ICMP Parameter Problem message with Code set to 0 and the Pointer referring to the Length field of the TLV.

Next steps

- Comments are welcome
- Ready for the WGLC

IOAM Update

IPPM Interim Meeting,
April 1, 15:00-16:30 UTC

IOAM related WG documents

- draft-ietf-ippm-ioam-data-09
- draft-ietf-ippm-ioam-ipv6-options-01
- draft-ietf-ippm-ioam-flags-01
- draft-ietf-ippm-ioam-direct-export-00

IOAM Data Fields: draft-ietf-ippm-ioam-data

WGLC on draft-ietf-ippm-ioam-data-08

Many comments received. Thanks to all reviewers!

- Issue #149: draft-ietf-ippm-ioam-data-08 WGLC#1 comments: Editorial
- Issue #150: draft-ietf-ippm-ioam-data-08 WGLC#1 comments: Nomenclature
- Issue #151: draft-ietf-ippm-ioam-data-08 WGLC#1 comments: Clarifications to SHOULD statements
- Issue #152: draft-ietf-ippm-ioam-data-08 WGLC#1 comments: Add further detail/clarifications to existing definitions
- Issue #153: draft-ietf-ippm-ioam-data-08 WGLC#1 comments: Changes to existing definitions
- Issue #154: draft-ietf-ippm-ioam-data-08 WGLC#1 comments: Suggestions for additional data fields
- Issue #155: draft-ietf-ippm-ioam-data-08 WGLC#1 comments: Security related
- Issue #156: draft-ietf-ippm-ioam-data-08 WGLC#1 comments: Needs clarification on timestamp insertion in E2E option
- Issue #157: draft-ietf-ippm-ioam-data-08 WGLC#1 comments: Editorial on Pre-allocated and Incremental Trace Options
- Issue #158: draft-ietf-ippm-ioam-data-08 WGLC#1 comments: Trace-Flags Registry clarification
- Issue #159: draft-ietf-ippm-ioam-data-08 WGLC#1 comments: Editorial DEX leftovers

Draft-ietf-ippm-ioam-data: Updates from -08 to -09

Editorial changes (Issue #149, Issue #157) - Haoyu Song, Greg Mirsky, Mickey Spiegel

- Clarified use of RemainingLen in opaque snapshot
- Removed sloppy language
- Removed reference to expired draft
- Explicit statements that IOAM-Trace-Type bits determine which data fields are included in each node data element, IOAM transit nodes must not modify fields in the fixed header, reserved “must be zero” fields need to be set and also ignored

Nomenclature (Issue #150) - Greg Mirsky

- Clarified that, despite the name “in-situ” not all IOAM functions require piggybacking meta-data onto live customer traffic.

Draft-ietf-ippm-ioam-data: Updates from -08 to -09

Use of SHOULD statements (Issue #151) - Greg Mirsky

- Reworded sentences which used uppercase SHOULD, despite RFC2119 style “SHOULD” was not intended.
  (e.g. “It SHOULD be possible to enable IOAM on a selected set of traffic” -> “Using IOAM on a selected set of traffic could be useful in deployments where …”)

Draft-ietf-ippm-ioam-data: Updates from -08 to -09

Expand/clarify existing definitions (Issue #152) - Greg Mirsky

- Hop_lim = 0xFF if the encap protocol does not carry TTL/Hop-Limit
- Field length for Trace-Type 0 = 4 Bytes
- Reference to “nodes supporting functionality defined in draft-ietf-ippm-ioam-data” instead of introducing terms like “IOAM capable node”.

Draft-ietf-ippm-ioam-data: Updates from -08 to -09

Clarify existing definitions (Issue #153) - Greg Mirsky, Barak Gnafi

- Unit type for “buffer occupancy”: Field may be implementation specific. Unit may be interpreted within the context of a namespace.
  The authors acknowledge that in some operational cases there is a need for the units to be consistent across a packet path through the network, hence recommend the implementations to use standard unit such as Bytes.

Draft-ietf-ippm-ioam-data: Updates from -08 to -09

Suggestions for additional data fields (Issue #154) - Greg Mirsky, Barak Gnafi

(Higher resolution timestamps, interface sent/receive rate, byte count on port)

- No updates on the document to allow the base document to be finished up. New data fields are expected to be covered by new drafts.

Security related (Issue #155) - Greg Mirsky, Tal Mizrahi

- Section 8 was extended to cover additional security aspects, incl. malicious change to IOAM data, mitigation to leaking IOAM data from network domain that employs IOAM, security considerations related to specific IOAM encapsulations

Draft-ietf-ippm-ioam-data: Updates from -08 to -09

Timestamp insertion in E2E option type (Issue #156) - Mickey Spiegel

- Within the IOAM encapsulating node, the time that the timestamp is retrieved can depend on the implementation. Some possibilities are: 1) the time at which the packet was received by the node, 2) the time at which the packet was transmitted by the node, 3) when a tunnel encapsulation is used, the point at which the packet is encapsulated into the tunnel. Each implementation should document when the E2E timestamp that is going to be put in the packet is retrieved.

Draft-ietf-ippm-ioam-data: Updates from -08 to -09

Trace-Flags Registry clarification (Issue #158) - Mickey Spiegel

- Section 7.4 now states: “Bit 1 - 3 are available for assignment via RFC Required process as per [RFC8126]” - which was missing prior.

Direct export references leftovers (Issue #159) - Mickey Spiegel

- All references to direct export are removed. Direct export is covered in draft-ietf-ippm-ioam-direct-export-00

draft-ietf-ippm-ioam-data-09 - Next Steps

- draft-ietf-ippm-ioam-data-09 should include all WGLC;
  Since the WGLC finished, no further comments have been received.

- Issue another WGLC once everyone had a chance to review
  draft-ietf-ippm-ioam-data-09, e.g. by mid May?

IOAM IPv6 Options
draft-ietf-ippm-ioam-ipv6-options-00

draft-ietf-ippm-ioam-ipv6-options: Updates from -00 to -01

Updates from -00 to -01

- Minor editorial updates only (author email address change)

Early allocation 2 IPv6 Option Types

- IPPM WG chairs initiated the process for early allocation

draft-ietf-ippm-ioam-ipv6-options-01: Next steps

- WGLC?

IOAM Flags
draft-ietf-ippm-ioam-flags-01

Changes Since Version -00

Clarifications about the loopback flag.

Text has been added about the purpose of the active flag.

Security considerations updated:

- Amplification attacks.
- Measures to limit the impact of amplification attacks:
  - Rate limiting.
  - Data minimization: up to one data field per exported packet.

- Seeking feedback from the WG.

Open Issue

- Loopback on the reverse path:
  - Pushing IOAM data on the reverse path is not necessary.
  - Problem: how do transit nodes know that a looped back packet is in transit on the reverse path?
    - New flag?
    - New IOAM type?
    - Clearing the RemainingLen field when the packet is looped back?

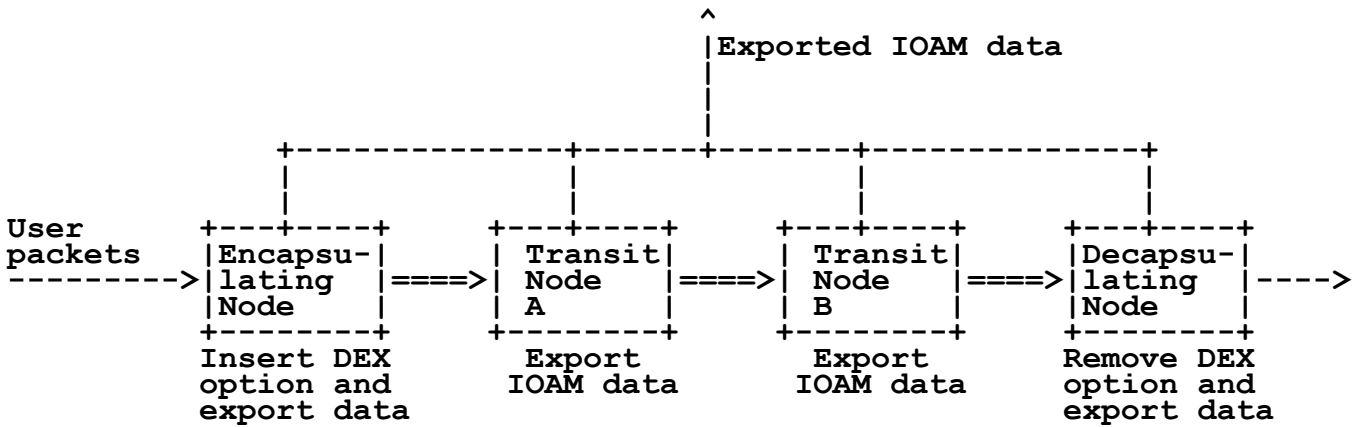
IOAM Direct Export
draft-ietf-ippm-ioam-direct-export-00

Direct Exporting (DEX) – Overview

IOAM data is exported without modifying data packets.

Simplifies transit node processing.

Reduces the data plane on-the-wire overhead of IOAM.



The History / State of this Draft

This draft combines two somewhat similar approaches:

- The PBT-I concept from draft-song-ippm-postcard-based-telemetry
- The Immediate Export flag from draft-mizrahi-ippm-ioam-flags

This draft is the product of a design team that worked on combining the two concepts.

December 2019 – adopted by the IPPM WG.

February 2019 – draft-ietf-00.

Open Issue – Hop Count

Question: Should the DEX option include an explicit Hop Count field, or is the Hop_Lim/Node_ID data field sufficient?

No Hop Count:

- Using existing functionality: Hop_Lim/Node_ID data field can be used, copied from the TTL/Hop Limit from the encap protocol, and included in the exported packet.
- The DEX option does not need to be modified by transit switches.

Explicit Hop Count:

- The lower layer TTL may not be accurate, e.g., L2 or hierarchical VPN.
- Allows to detect IOAM-capable node that fails to export packets.

IETF 107, IPPM WG draft-geib-ippm-connectivity-monitoring

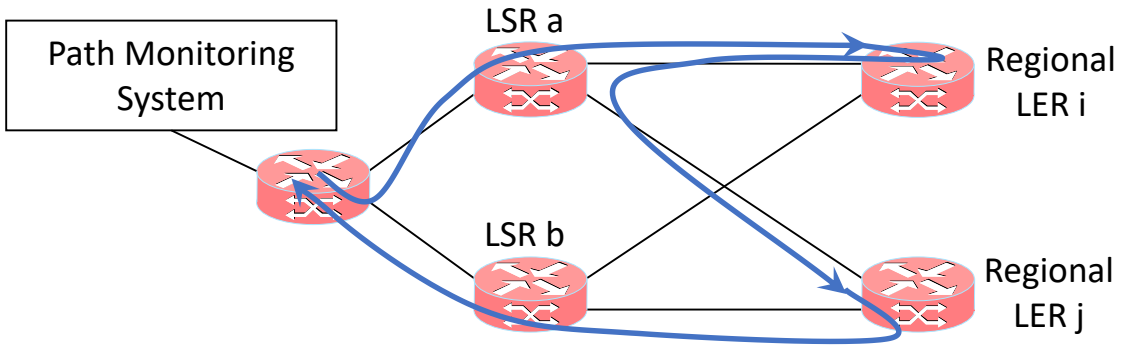
Scope:
Monitor Segment Routed subpaths or links to detect and locate loss of connectivity and congestion.

Properties:
- Designed to monitor network sections with known topology.
- One measurement path per monitored link (or 0,5 per monitored interface).
- Change point detection doesn‘t require well synchronised timestamps.
- Applies segment routing to allow for an a priori designed network tomography evaluation and a limited number of monitoring systems.
- Round-Trip Delay and One-Way Delay estimate of the monitored link or path.

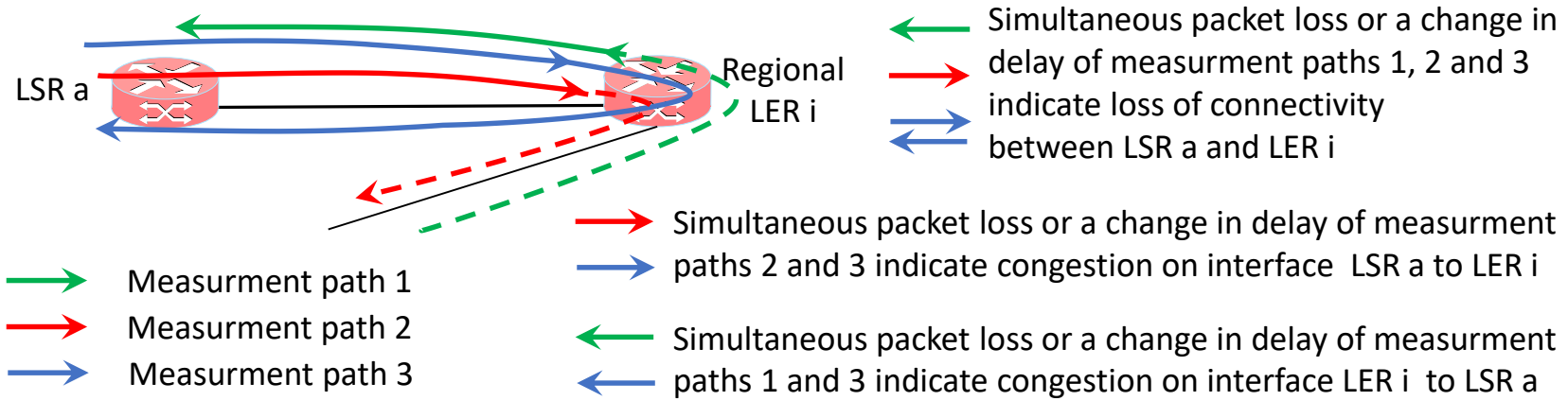
Aim:
- Standardise suitable metrics.
- Interested? Read and comment.

IETF 107, IPPM WG draft-geib-ippm-connectivity-monitoring

Set-up of each individual measurement path (one only shown):
1 round trip, 1 downstream & 1 upstream pass of different monitored LSR – LER IFs.



Detection of events (different measurement paths combine as shown below to create an individual measurement path combination per monitored interface):



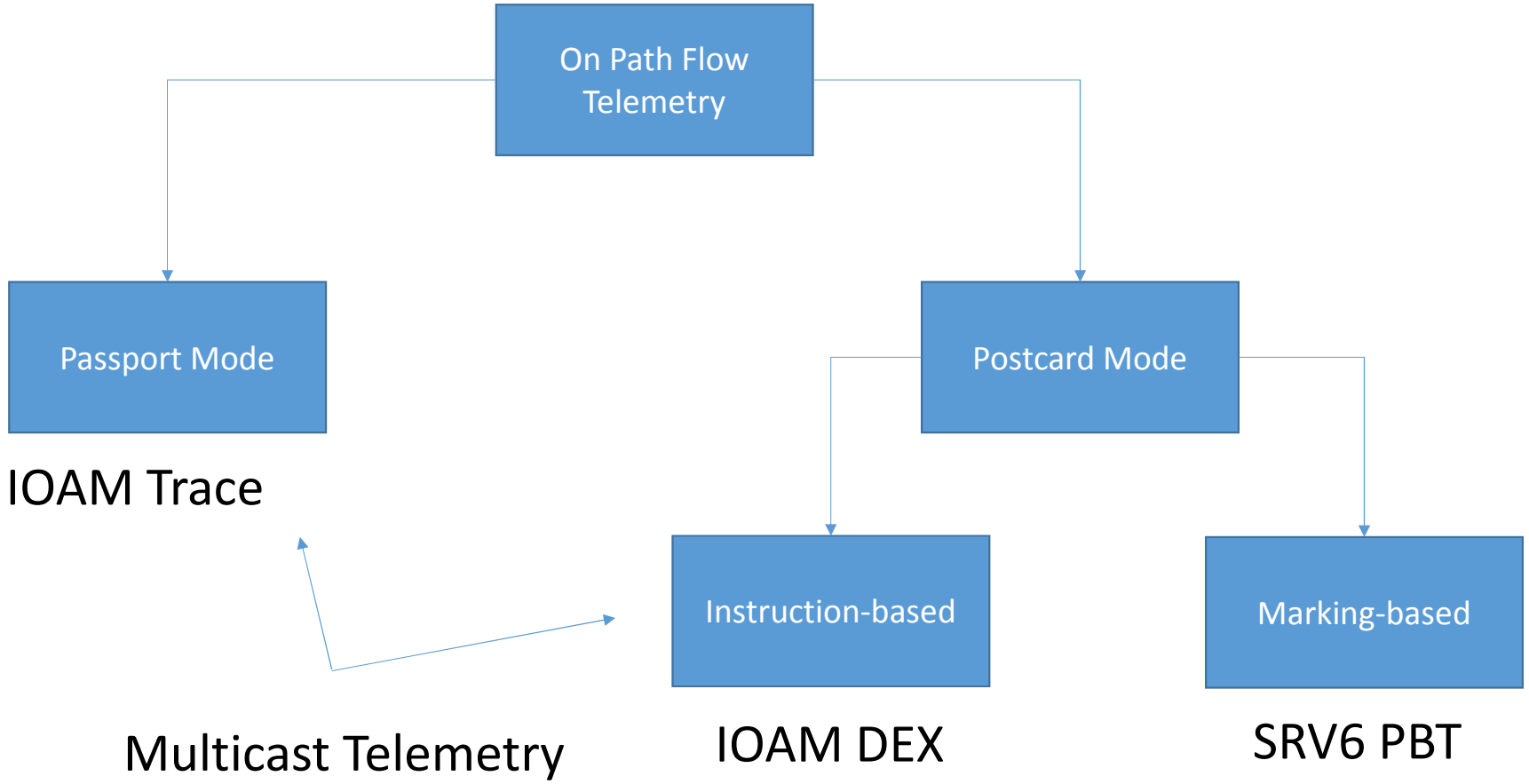
Postcard-based On-Path Flow Data Telemetry

draft-song-ippm-postcard-based-telemetry-06

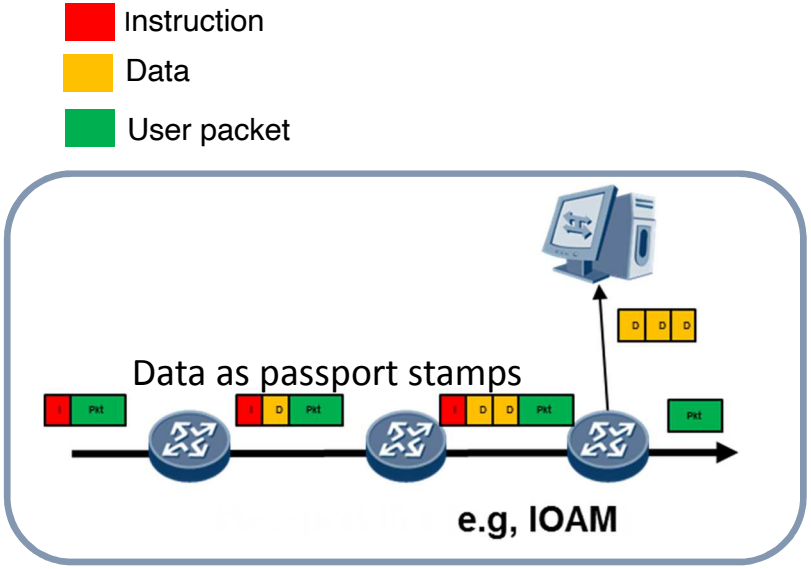
Haoyu Song (Futurewei)
Tianran Zhou (Huawei)
Zhenbin Li (Huawei)
Jongyoon Shin (SK Telecom)
Kyungtae Lee (LG U+)

What's New

- Change the status of the draft to "Informational"
- Position PBT-M and PBT-I as two high level approaches of PBT
  - An implementation of PBT-I as an IOAM option: Direct Export
  - An implementation of PBT-M in SRv6


On Path Flow Telemetry
Passport Mode
Postcard Mode
IOAM Trace
Instruction-based
Marking-based
Multicast Telemetry
IOAM DEX
SRV6 PBT

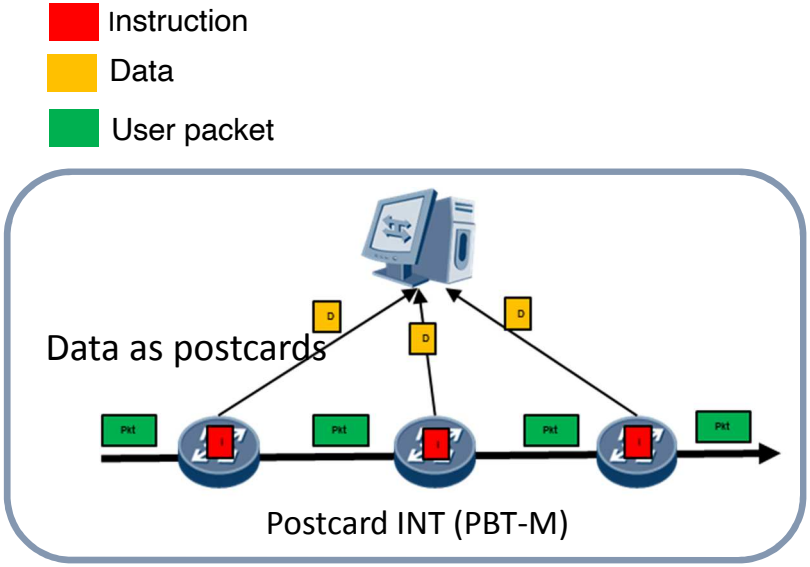
Passport-based On-path Telemetry: IOAM Trace



- Forwarding performance impact
- Packet size inflation
- Encapsulation
- Security
- Drop awareness & localization

- Self-describing
- Export overhead
- Configuration overhead

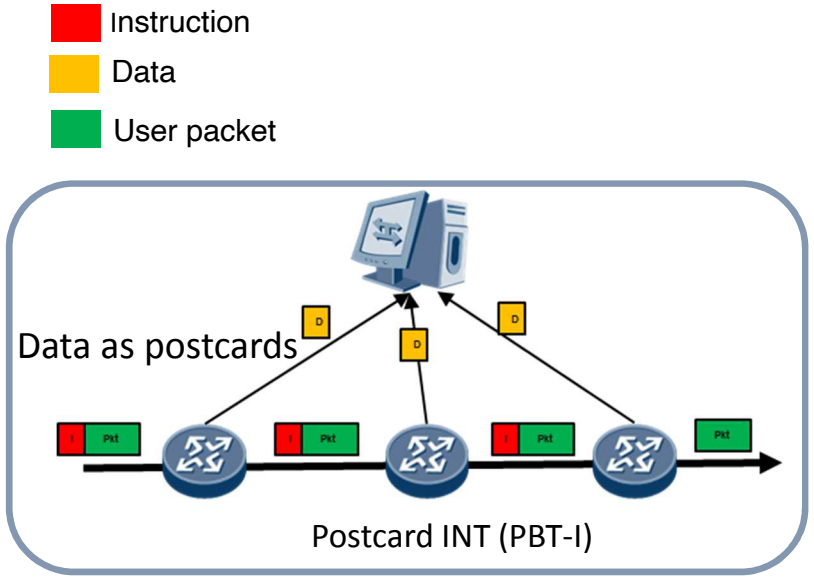
Postcard-based On-path Telemetry: PBT-M



- Forwarding performance impact
- Packet size inflation
- Encapsulation
- Security
- Drop localization

- Data correlation
- Export overhead
- Configuration overhead

Postcard-based On-path Telemetry: IOAM DEX



- Forwarding performance impact
- Packet size inflation
- Encapsulation
- Security
- Drop localization
- Data correlation
- Export overhead
- Configuration overhead

Why this document

- Describes the high level approach and classification
- Summarizes the pros and cons of each approach
- Details the marking-base PBT which is not covered anywhere else

- This work has motivated several other works
  - Embodiments or standardizations of each approach are covered by separate documents

Next Steps

- Request for WG adoption

Client-Server Explicit Performance Measurements: 2bit Packet Loss

draft-cfb-ippm-spinbit-measurements-01

1st of April 2020, IPPM WG interim meeting

Fabio BULGARELLA (Telecom Italia - TIM)

Mauro COCIGLIO (Telecom Italia - TIM)

Giuseppe FIOCCOLA (Huawei)

Massimo NILO (Telecom Italia - TIM)

Riccardo SISTO (Politecnico di Torino)

IETF

2bit Packet Loss Explicit Measurement

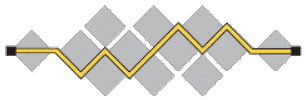
Our 2bit Packet Loss measurement it's an enhancement of the methodology described in: draft-ferrieuxhamchaoui-quic-lossbits

In the above draft two bits of a protocol header (e.g QUIC, TCP,...) are used to mark the production traffic between Client and Server.

The 2 bits are the sQuare bit (Q-bit) and Loss bit (L-bit):

- The Q-bit creates square waves of a known length (e.g. 64 packets): RFC 8321 Alternate Marking
- The L-bit is set in a packet by the end-point when the protocol signals a retransmission.

In our draft the R-bit substitutes the L-bit.



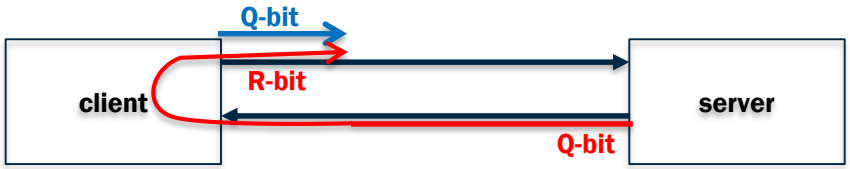
The Reflection square bit (R-bit).

Our idea is to reflect the Q-bit in the opposite direction using the R-bit.

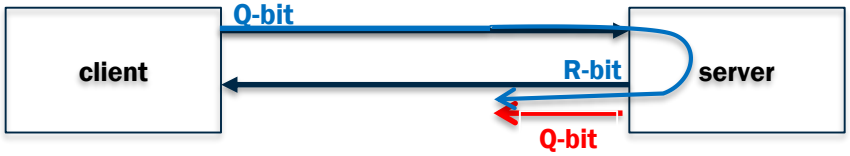
The sizes of the transmitted R-bit blocks are the **"average sizes"** of the received Q-bit blocks.

This idea allows to have continuous alternate marked packet blocks in both directions.

The Client generates the Q-bit signal and reflects the received Q-bit signal using the R-bit signal:



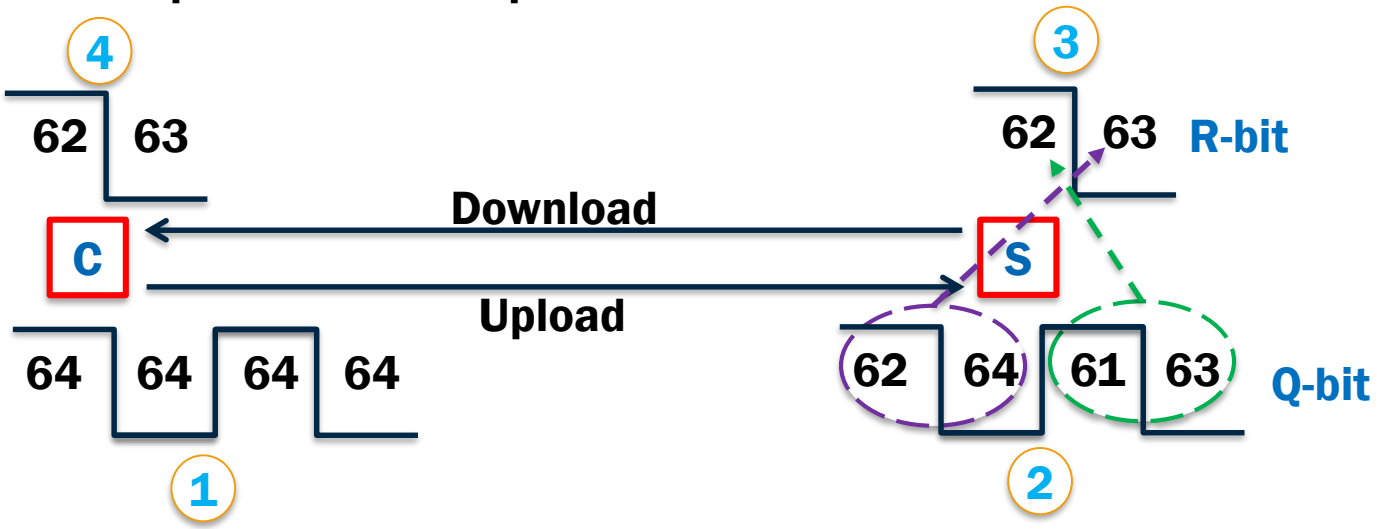
The Server does the same in the opposite direction:



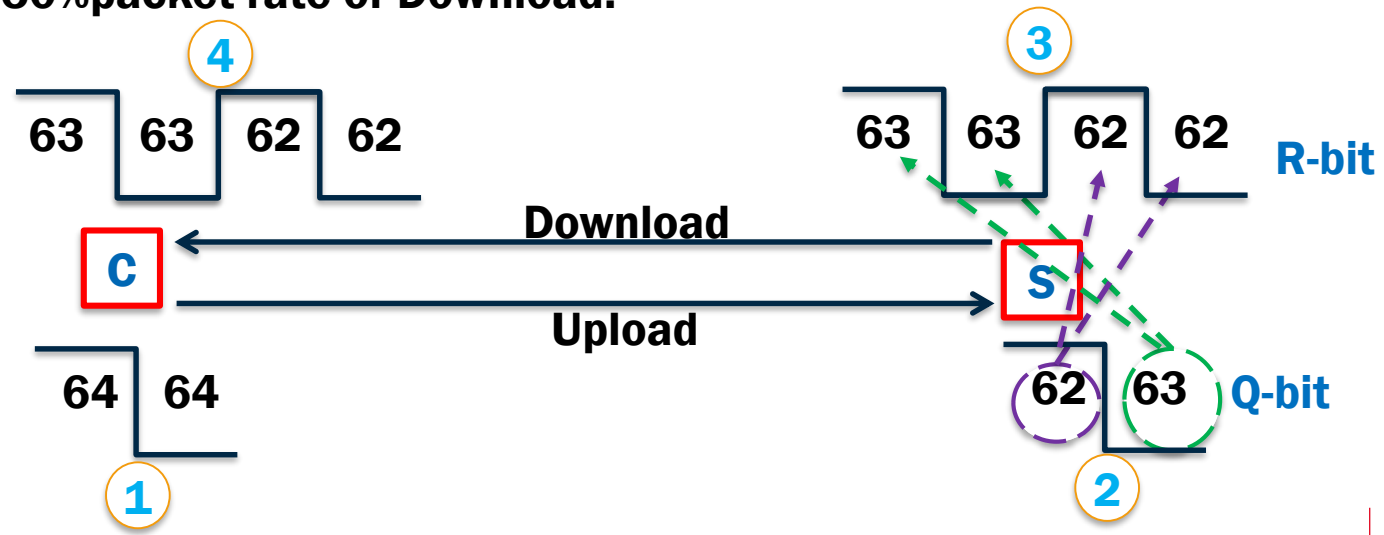
IETF®

Example: Upload Packet Loss with different packet rates

Download 50% packet rate of Upload:



Upload 50%packet rate of Download:



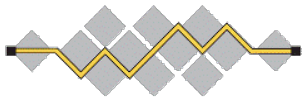
R-bit Algorithm

«When the transmission of a new R-block starts, its size M is set equal to the size of the last Q-marked period whose reception has been completed;

if, before transmission of the R-block is terminated, the reception of at least one further Q-marked period is completed, the size of the R-block is updated to the average size of the further received Q-marked periods»

Algorithm properties:

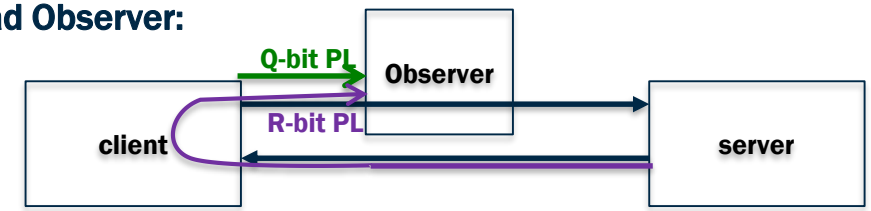
- It works in both cases when the reflected packets number is greater than those received and when the reflected packets number is lower.
- All traffic is measured (all the production traffic has both the Q-bit and the R-bit marked)



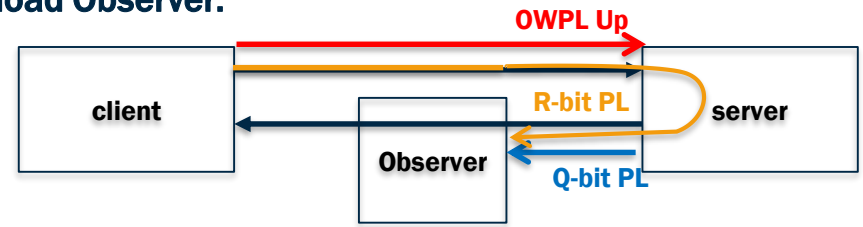
One direction Observer:

- **Download Observer:**

- **Upload Observer:**

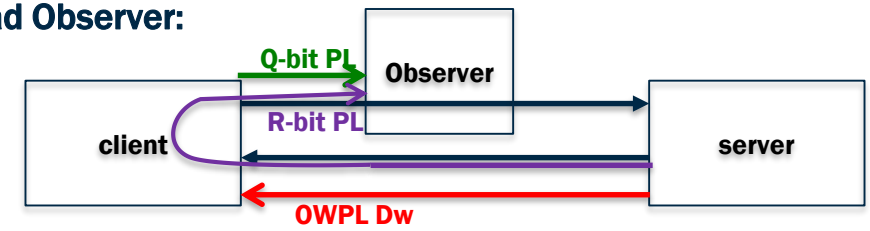


- **Download Observer:**



OWPL Up = R-bit PL Dw – Q-bit PL Dw

- **Upload Observer:**

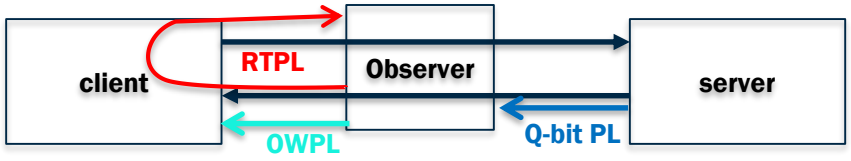


OWPL Dw = R-bit PL Up – Q-bit PL Up

OWPL: One Way Packet Loss

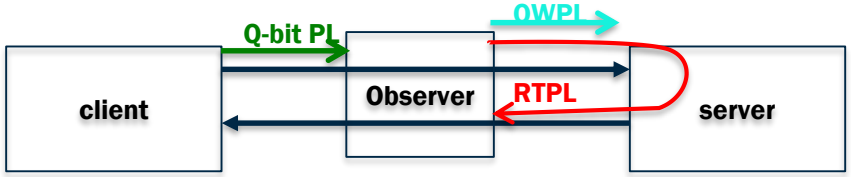
Two direction Observer:

- **Observer-Client RTPL and OWPL:**



RTPL =R-bit PL Up – Q-bit PL Dw ⇨ RTPL – Q-bit PL Up= OWPL

- **Observer-Server RTPL and OWPL:**



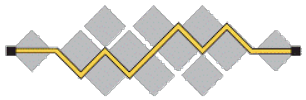
RTPL =R-bit PL Dw – Q-bit PL Up ⇨ RTPL – Q-bit PL Dw= OWPL

RTPL: Round Trip Packet Loss
OWPL: One Way Packet Loss

L-bit versus R-bit

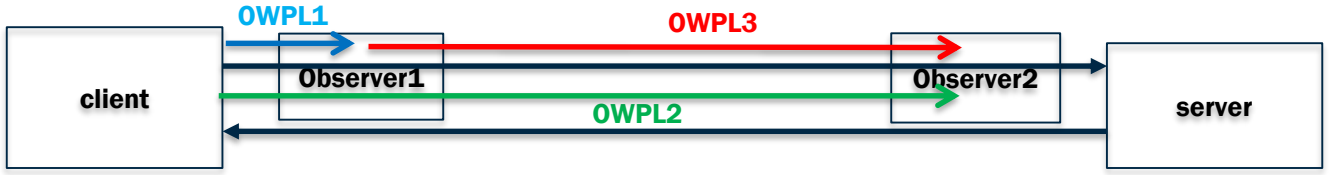
L-bit weaknesses (& R-bit strengths):

1. The dependence from an internal protocol variable not directly connected to losses but to retransmissions.
2. Ack packet losses are not correctly detected.
3. The loss measurement signal is inaccurate in case of losses.



2Point One-Way Packet Loss (Q-bit only)

- Observer2-Observer1 OWPL:



Observer2-Observer1 One-Way: OWPL2 – OWPL1 = OWPL3

- Observer1-Observer2 OWPL:

Observer1-Observer2 One-Way: OWPL1 – OWPL2 = OWPL3

Performance Measurement Using STAMP for Segment Routing Networks

draft-gandhi-spring-stamp-srpm-00

*(*previously *draft-gandhi-spring-twamp-srpm-07)*

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
- History of the Draft
- Updates Since IETF-106
- Summary
- Next Steps

Requirements and Scope

Requirements:

- Delay and Loss Performance Measurement (PM)
  - ✓ Links and End-to-end P2P/P2MP SR Policies
  - ✓ Applicable to SR-MPLS/SRv6 data planes
- No need to bootstrap PM session (e.g. to negotiate UDP port) - spirit of SR
  - ✓ Stateless on egress node - spirit of SR
- Handle ECMP for SR Policies
- Support stand-alone direct-mode loss measurement

Scope:

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

History of the Draft

- Feb 2019
  - Draft was published - *draft-gandhi-spring-twamp-srpm-00*
- May 2019
  - Added STAMP TLV for Return Path
- Mar 2019
  - Presented version-00 at IETF 104 Prague in SPRING WG
- July 2019
  - Presented version-01 at IETF 105 Montreal in IPPM WG
    - Slide 9 Titled - Applicability of STAMP – STAMP is supported
- Aug 2019
  - Version-02 updates included a section on stand-alone LM messages
- Nov 2019
  - **SPRING Chairs announced in the meeting the agreement with IPPM chairs to progress the draft in SPRING WG**
  - Presented version-04 at IETF 106 Singapore in SPRING WG
- Mar 2020
  - Moved SRPM STAMP support to *draft-gandhi--spring-**stamp**-srpm-00*
  - Keep SRPM TWAMP Light support in *draft-gandhi-spring-**twamp**-srpm-08*

Updates Since IETF-106 (Version-04)

Updates:

1. Defined Control Code for “In-band Response Requested” for STAMP
   - ✓ Updated Two-way mode procedure using the Control Code
2. Defined Destination Address in STAMP Node Address TLV to identify the intended Destination node
3. Added Return Address Sub-TLV in the STAMP Return Path TLV to send response to a specific node
4. Various editorial changes

Open Items:

- Identify TLV as Mandatory or Optional
- Update IANA registry action

STAMP Control Code Field

For a Query: Sender Control Code

0x0: Out-of-band Response Requested.
This is also the default 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. The bidirectional path does not have to be an SR path.

For a Response: Reflector Control Code

0x1: Error - Invalid Message.
Indicates that the operation failed because the received query message could not be processed.
0xN: Additional Error will be defined in future.

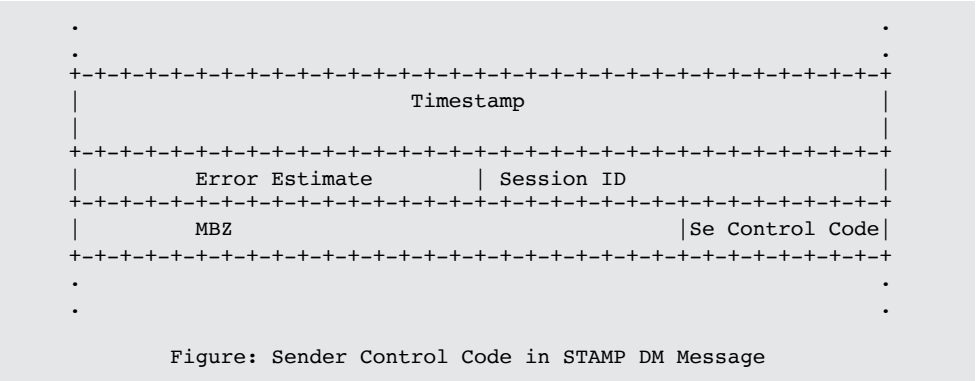


Figure: Sender Control Code in STAMP DM Message

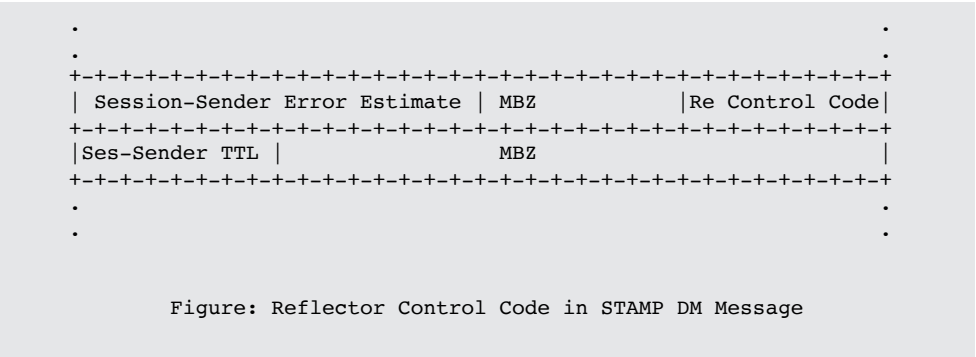


Figure: Reflector Control Code in STAMP DM Message

Performance Measurement Modes

- One-way Measurement Mode
  - Reply sent “out of band” on IP/UDP path by default
- Two-way Measurement Mode
  - Reply sent “in-band” on reverse SR path
    - Based on Control Code from the probe query message
  - Use Return Path TLV for STAMP from the probe query message
- Loopback Measurement Mode
  - Probe message carries the return path in the header of the packet

Destination Address in STAMP Node Address TLV

Destination Node Address (value TBA1):

- Indicates the address of the intended recipient node of the query message.
- The reflector node SHOULD NOT send response if it is not the intended destination node of the query.
- Useful when query is sent with 127/8 destination address.

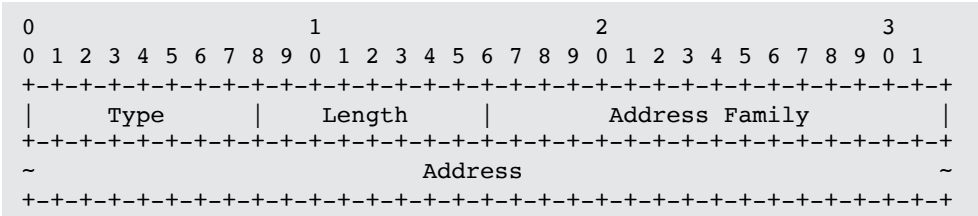


Figure: Node Address TLV Format

Return Address in STAMP Return Path TLV

Return Path (value TBA2):

Sub-TLVs:

- Type (value 0): Return Address. Target node address of the response different than the Source Address in the query
- Type (value 1): SR-MPLS Label Stack of the Reverse SR Path
- Type (value 2): SR-MPLS Binding SID [draft-ietf-pce-binding-label-sid] of the Reverse SR Policy
- Type (value 3): SRv6 Segment List of the Reverse SR Path
- Type (value 4): SRv6 Binding SID [draft-ietf-pce-binding-label-sid] of the Reverse SR Policy

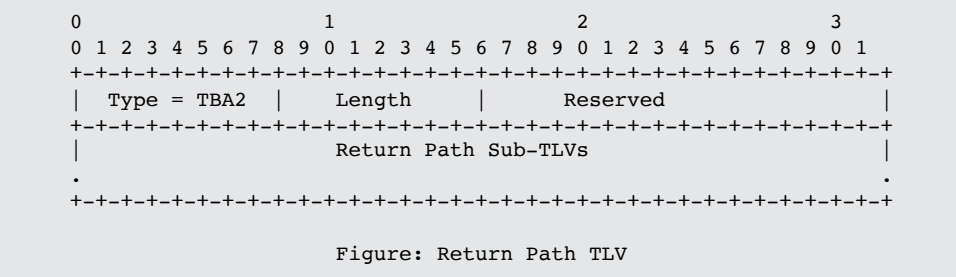


Figure: Return Path TLV

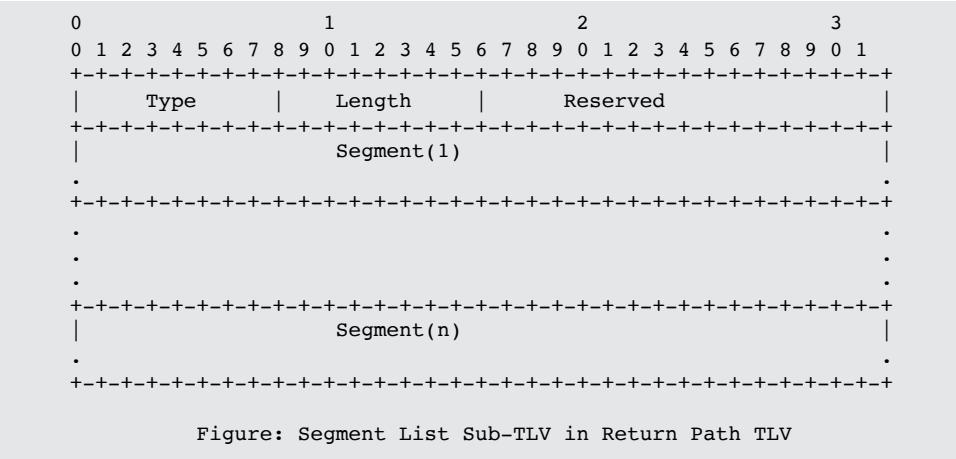
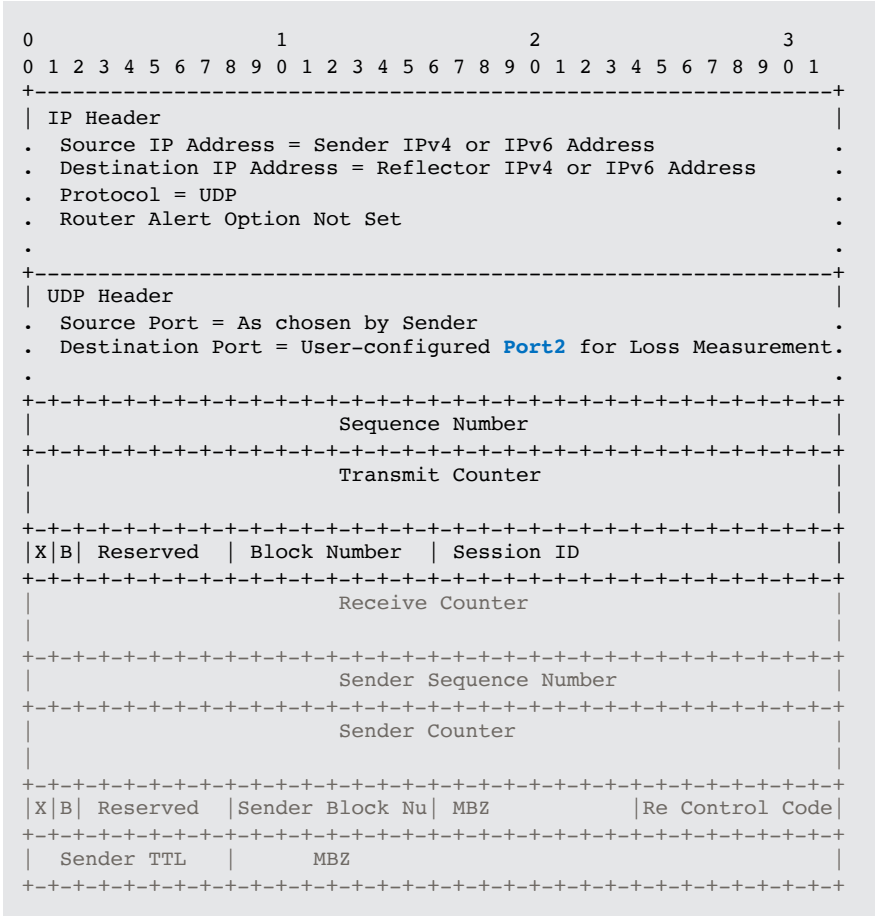


Figure: Segment List Sub-TLV in Return Path TLV

Stand-alone LM Message Format for STAMP

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



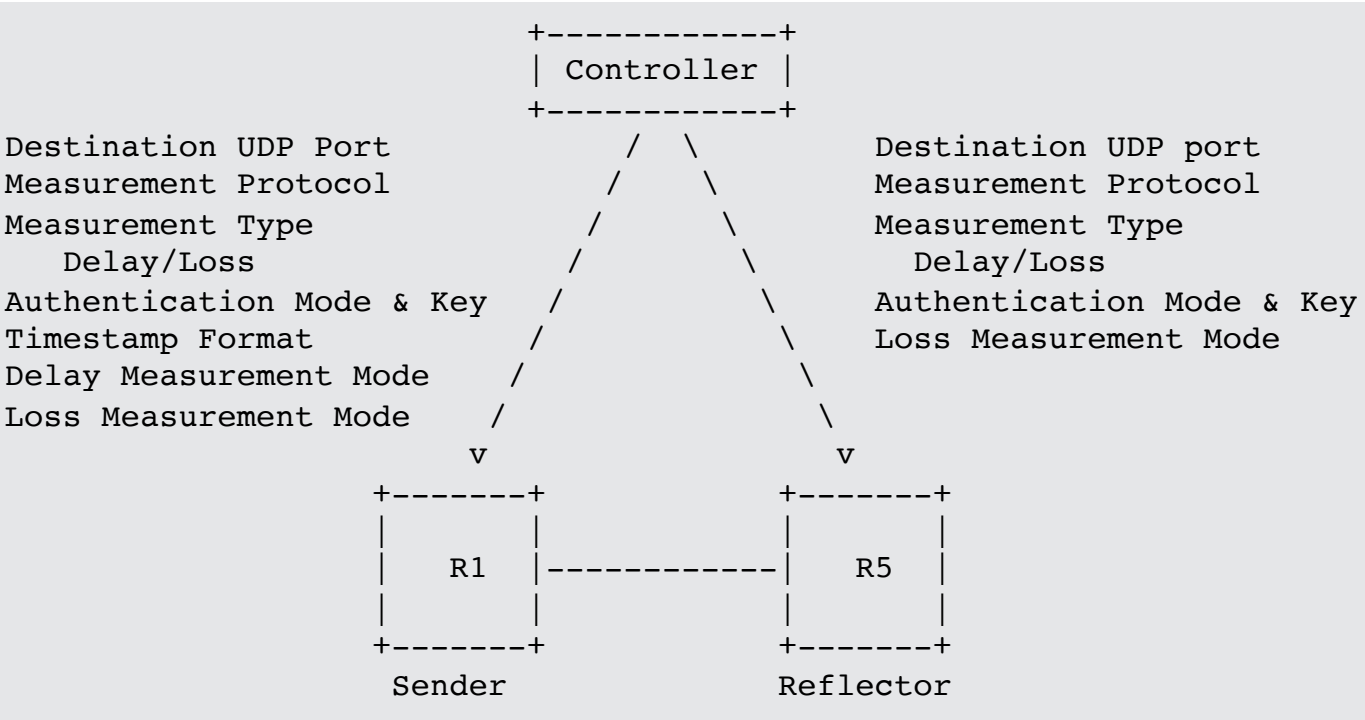
Next Steps

- Welcome your comments and suggestions
- Implementation exists
- In WG adoption (SPRING WG) queue
- Keep IPPM WG in the loop about the milestones

Thank you

Backup

Example Provisioning Model



Probe Query for Links

- User-configured destination UDP **port1** is used for DM probe messages in unauthenticated mode and **port2** is used for LM probe messages in unauthenticated mode.

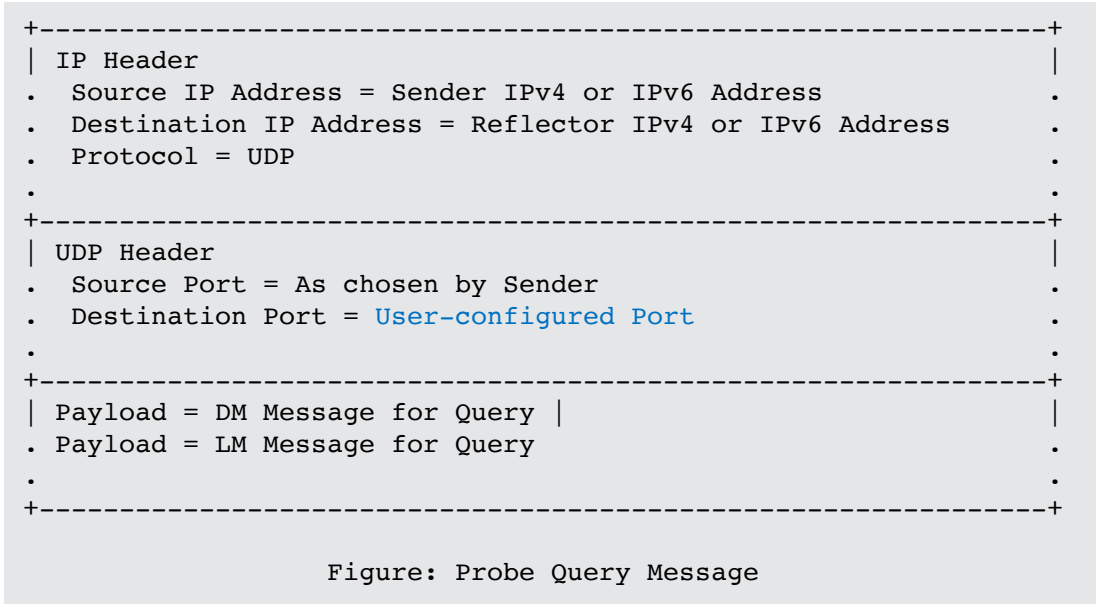


Figure: Probe Query Message

Probe Query for SR-MPLS and SRv6 Policy

For **end-to-end** performance delay/loss measurement of SR Policy, the probe query messages are sent on the SR Policy path with:

1. MPLS label stack for SR-MPLS Policies
2. SRv6 SRH [RFC 8754] with SID list for SRv6 Policies

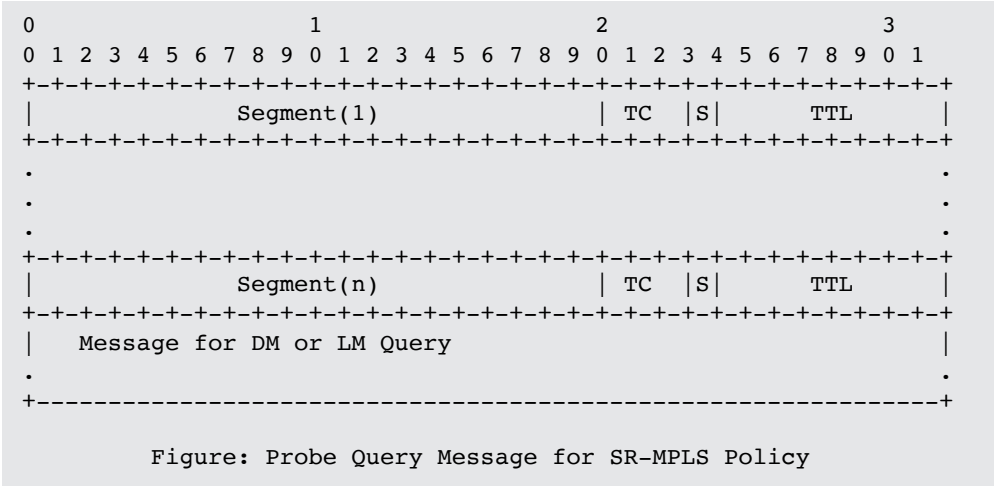


Figure: Probe Query Message for SR-MPLS Policy

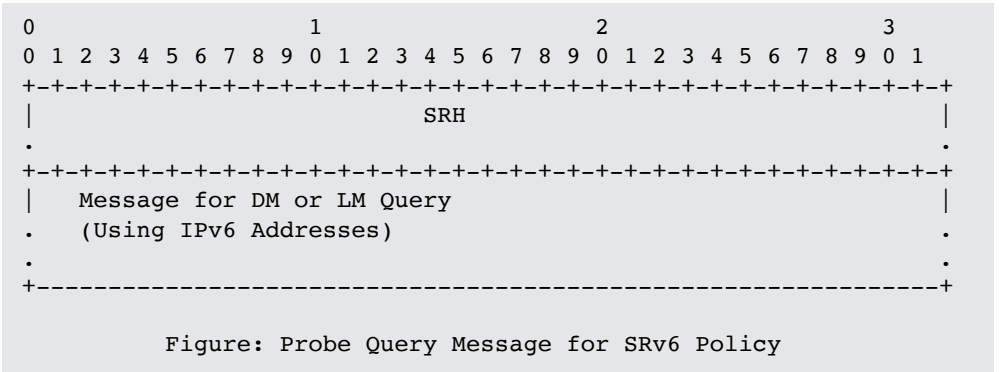


Figure: Probe Query Message for SRv6 Policy

Probe Response Message

- The probe response message is sent using the IP/UDP information from the probe query message.

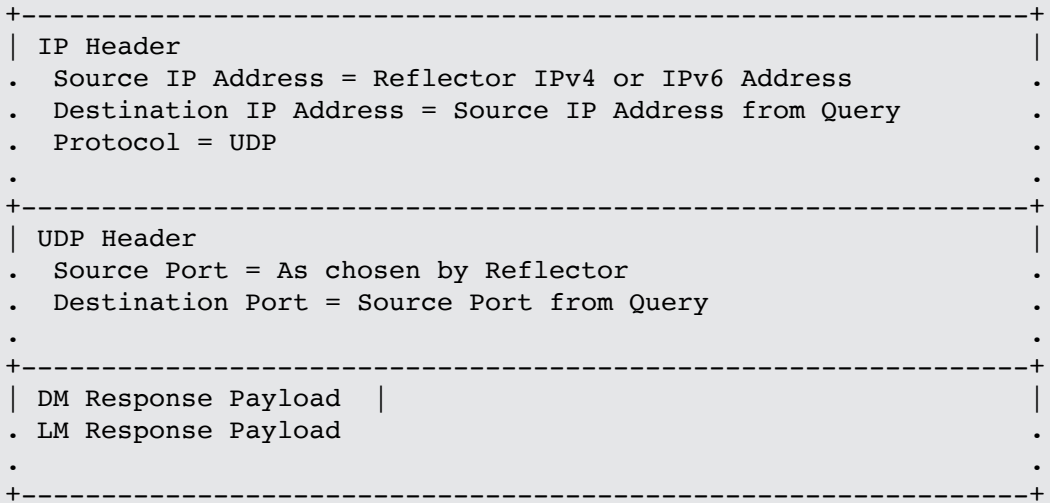


Figure: Probe Response Message

ECMP Support for SR Policy

- SR Policy can have ECMP between the ingress and transit nodes, between transit nodes and between transit and egress nodes.
- Sending PM probe queries that can take advantage of the hashing function in forwarding plane.
- Existing forwarding mechanisms are applicable to PM probe messages:
  - For IPv4
    - Destination addresses in IPv4 header (e.g. 127/8)
  - For IPv6
    - Destination addresses in IPv6 header (e.g. FFFF:7F00/104)
    - Flow label in IPv6 header

Backup

STAMP DM Message with Direct Measurement TLV (DM+LM Combined Probe Message)

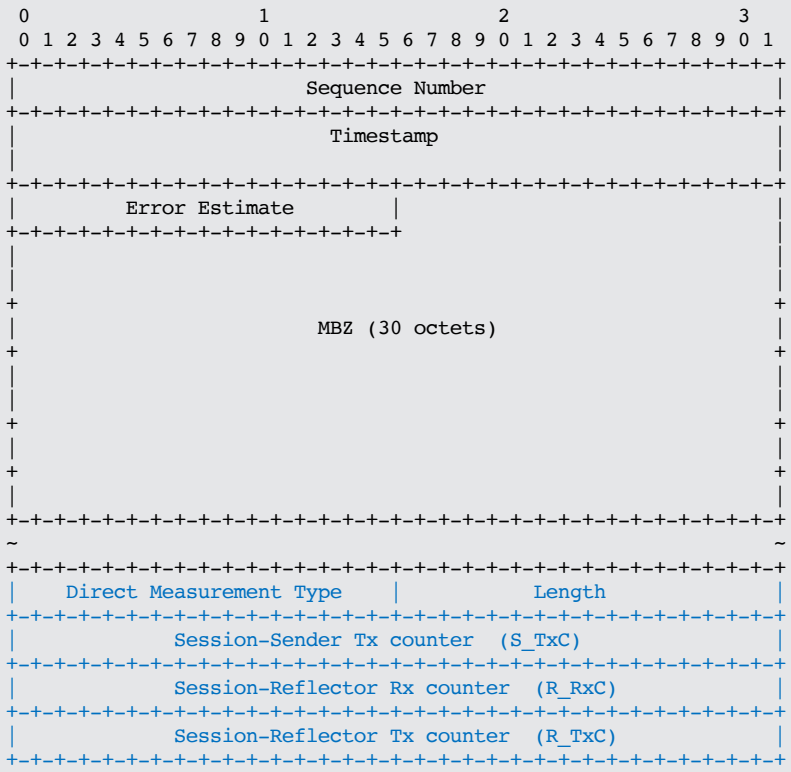


Figure: Sender Message Format

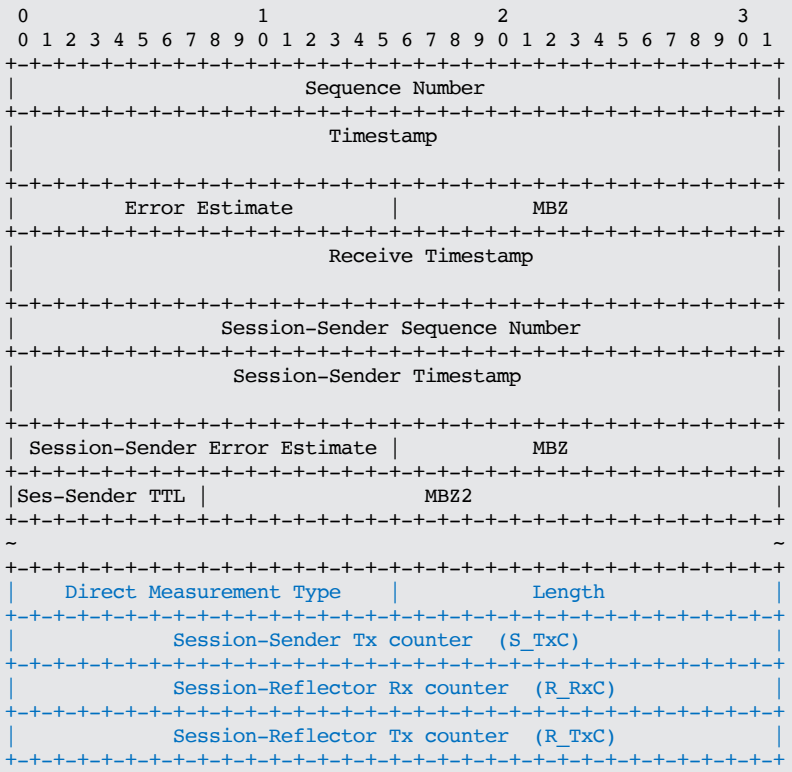


Figure: Reflector Message Format

Thank you