Workgroup: IPPM
Internet-Draft:
draft-ietf-ippm-stamp-on-lag-01
Published: 3 March 2023
Intended Status: Standards Track
Expires: 4 September 2023
Authors: Z. Li (China Mobile)
T. Zhou (Huawei)
J. Guo (ZTE Corp.)
G. Mirsky (Ericsson)
R. Gandhi (Cisco)

Simple Two-Way Active Measurement Protocol Extensions for Performance Measurement on LAG

Abstract

This document extends Simple Two-Way Active Measurement Protocol (STAMP) to implement performance measurement on every member link of a Link Aggregation Group (LAG). Knowing the measured metrics of each member link of a LAG enables operators to enforce a performance based traffic steering policy across the member links.

Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at https://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on 4 September 2023.

Table of Contents

1. Introduction

Link Aggregation Group (LAG), as defined in [IEEE802.1AX], provides mechanisms to combine multiple physical links into a single logical link. This logical link offers higher bandwidth and better resiliency, because if one of the physical member links fails, the aggregate logical link can continue to forward traffic over the remaining operational physical member links.

Usually, when forwarding traffic over LAG, the hash-based mechanism is used to load balance the traffic across the LAG member links. Link delay of each member link varies because of different transport paths. To provide low latency service for time sensitive traffic, we need to explicitly steer the traffic across the LAG member links based on the link delay, loss and so on. That requires a solution to measure the performance metrics of each member link of a LAG. Hence the measured performance metrics can work together with layer 2 bundle member link attributes advertisement [RFC8668] for traffic steering.

Simple Two-Way Active Measurement Protocol (STAMP) [RFC8762] is an active measurement method according to the classification given in [RFC7799], which can complement passive and hybrid methods. It provides a mechanism to measure both one-way and round-trip performance metrics, like delay, delay variation, and packet loss. Running a single STAMP test session over the aggregation without the knowledge of each member link would make it impossible to measure the performance of a given physical member link. The measured metrics can only reflect the performance of one member link or an average of some/all member links of the LAG.

This document extends STAMP to implement performance measurement on every member link of a LAG. The proposed method could also potentially apply to layer 3 ECMP (Equal Cost Multi-Path), e.g., with Segment Routing Policy [RFC9256].

2. Micro Session on LAG

This document intends to address the scenario (e.g., Figure 1) where a LAG (e.g., the LAG includes four member links) directly connects two nodes (A and B) . The goal is to measure the performance of each link of the LAG.

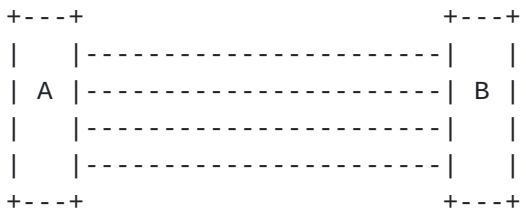

Figure 1: PM on LAG

To measure the performance metrics of every member link of a LAG, multiple sessions (one session for each member link) need to be established between the two end points that are connected by the LAG. These sessions are called micro sessions in the remainder of this document.

All micro sessions of a LAG share the same Sender IP Address and Receiver IP Address. As for the UDP Port, the micro sessions may share the same Sender Port and Receiver Port pair, or each micro session is configured with a different Sender Port and Receiver Port pair. But from the operational point of view, the former is simpler and is recommended.

At the Sender side, each micro STAMP session MUST be assgined with a unique SSID [RFC8972]. Both the micro STAMP Session Sender and Reflector MUST use SSID to correlate the Test packet to a micro

session. If there is no such a session, or the SSID is not correct, the Test packet MUST be discarded.

Test packets MAY carry the member link information for validation check. For example, when a micro STAMP Session-Sender receives a reflected Test packet, it may need to check whether the Test packet is from the expected member link. The detailed description about the member link validation is in section 3.

A micro STAMP Session-Sender MAY include the Follow-Up Telemetry TLV [RFC8972] to request information from the micro Session-Reflector. This timestamp might be important for the micro Session-Sender, as it improves the accuracy of network delay measurement by minimizing the impact of egress queuing delays on the measurement.

3. Member Link Validation

Test packets MAY carry the member link information for validation check. The micro Session Sender can verify whether the test packet is reveived from the expected member link. It can also verify whether the packet is sent from the expected member link at the Reflector side. The micro Session Reflector can verify whether the test packet is received from the expected member link.

3.1. Micro-session ID TLV

STAMP TLV [RFC8972] mechanism extends STAMP Test packets with one or more optional TLVs. This document defines the TLV Type (value TBA1) for the Micro-session ID TLV that carries the micro STAMP Session-Sender member link identifier and Session-Reflector member link identifier. The format of the Micro-session ID TLV is shown as follows:

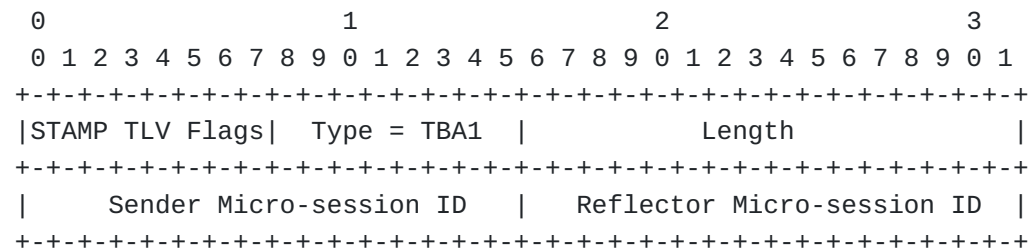

Figure 2: Micro-session ID TLV

*Type: A one-octet field. Value TBA1 is allocated by IANA (Section 5).

*Length: A two-octet field equal to the length of the Value field in octets. The Length field value MUST be 4 octets.

*Sender Micro-session ID (2-octets in length): it is defined to carry the Micro-session identifier of the Sender side. The value of the Sender Member Link ID MUST be unique at the Session-Sender.

*Reflector Micro-session ID (2-octets in length): it is defined to carry the Micro-session identifier of the Reflector side. The value of the Reflector Member ID MUST be unique at the Session-Reflector.

3.2. Micro STAMP-Test Procedures

The micro STAMP-Test reuses the procedures as defined in Section 4 of STAMP [RFC8762] with the following additions.

The micro STAMP Session-Sender MUST send the micro STAMP-Test packets over the member link with which the session is associated. The configuration and management of the mapping between a micro STAMP session and the Sender/Reflector member link identifiers are outside the scope of this document.

When sending a Test packet, the micro STAMP Session-Sender MUST set the Sender Micro-session ID field with the member link identifier associated with the micro STAMP session. If the Session-Sender knows the Reflector member link identifier, the Reflector Micro-session ID field MUST be set. Otherwise, the Reflector Micro-session ID field MUST be zero. The Reflector member link identifier can be obtained from pre-configuration or learned from data plane (e.g., the reflected Test packet). How to obtain/learn the Reflector member link identifier is outside of this document's scope.

When the micro STAMP Session-Reflector receives a Test packet, if the Reflector Micro-session ID is not zero, the micro STAMP Session-Reflector MUST use the Reflector member link identifier to check whether it is associated with the micro STAMP session. If the validation fails, the Test packet MUST be discarded. If all validations passed, the Session-Reflector sends a reflected Test packet to the Session-Sender. The micro STAMP Session-Reflector MUST put the Sender and Reflector member link identifiers that are associated with the micro STAMP session in the Sender Micro-session ID and Reflector Micro-session ID fields respectively. The Sender member link identifier is copied from the received Test packet.

When receiving a reflected Test packet, the micro Session-Sender MUST use the Sender Micro-session ID to validate whether the reflected Test packet is correctly transmitted over the expected member link. If the validation fails, the Test packet MUST be discarded. The micro Session-Sender MUST use the Reflector Micro-

session ID to validate the Reflector's behavior. If the validation fails, the Test packet MUST be discarded.

Two modes of the STAMP Session-Reflector, stateless and stateful, characterize the expected behavior. The micro STAMP-Test supports both stateless and stateful modes. However, the micro STAMP-Test does not introduce any additional state to STAMP, i.e, any procedure with regard to the Micro-session ID is stateless.

4. IANA Considerations

In the "STAMP TLV Types" registry created for [RFC8972], a new STAMP TLV Type for Micro-session ID TLV is requested from IANA as follows:

STAMP TLV Type Value	Description	Semantics Definition	Reference
TBA1	Micro-session ID TLV	Section 3	This Document

Figure 3: New STAMP TLV Type

5. Security Considerations

The STAMP extension defined in this document is intended for deployment in LAG scenario where Session-Sender and Session-Reflector are directly connnected. As such, it's assumed that a node involved in STAMP protocol operation has previously verified the integrity of the LAG connection and the identity of its one-hop-away peer node.

This document does not introduce any additional security issues and the security mechanisms defined in [RFC8762] and [RFC8972] apply in this document.

6. Acknowledgements

The authors would like to thank Mach Chen, Min Xiao, Fang Xin for the valuable comments to this work.

7. References

7.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/

RFC2119, March 1997, <https://www.rfc-editor.org/info/rfc2119>.

[RFC7799] Morton, A., "Active and Passive Metrics and Methods (with Hybrid Types In-Between)", RFC 7799, DOI 10.17487/RFC7799, May 2016, <https://www.rfc-editor.org/info/rfc7799>.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <https://www.rfc-editor.org/info/rfc8174>.

[RFC8762] Mirsky, G., Jun, G., Nydell, H., and R. Foote, "Simple Two-Way Active Measurement Protocol", RFC 8762, DOI 10.17487/RFC8762, March 2020, <https://www.rfc-editor.org/info/rfc8762>.

[RFC8972] Mirsky, G., Min, X., Nydell, H., Foote, R., Masputra, A., and E. Ruffini, "Simple Two-Way Active Measurement Protocol Optional Extensions", RFC 8972, DOI 10.17487/RFC8972, January 2021, <https://www.rfc-editor.org/info/rfc8972>.

7.2. Informative References

[IEEE802.1AX] IEEE Std. 802.1AX, "IEEE Standard for Local and metropolitan area networks - Link Aggregation", November 2008.

[RFC8668] Ginsberg, L., Ed., Bashandy, A., Filsfils, C., Nanduri, M., and E. Aries, "Advertising Layer 2 Bundle Member Link Attributes in IS-IS", RFC 8668, DOI 10.17487/RFC8668, December 2019, <https://www.rfc-editor.org/info/rfc8668>.

[RFC9256] Filsfils, C., Talaulikar, K., Ed., Voyer, D., Bogdanov, A., and P. Mattes, "Segment Routing Policy Architecture", RFC 9256, DOI 10.17487/RFC9256, July 2022, <https://www.rfc-editor.org/info/rfc9256>.

Authors' Addresses

Zhenqiang Li
China Mobile
No. 29 Finance Avenue, Xicheng District
Beijing
China

Email: li_zhenqiang@hotmail.com

Tianran Zhou

Huawei
China

Email: zhoutianran@huawei.com

Jun Guo
ZTE Corp.
China

Email: guo.jun2@zte.com.cn

Greg Mirsky
Ericsson
United States of America

Email: gregimirsky@gmail.com

Rakesh Gandhi
Cisco
Canada

Email: rgandhi@cisco.com