Workgroup: Network Working Group
Internet-Draft: draft-li-ippm-otwamp-on-lag-03
Published: 7 March 2022
Intended Status: Standards Track
Expires: 8 September 2022
Authors: Z. Li (China Mobile)
T. Zhou (Huawei)
J. Guo (ZTE Corp.)
G. Mirsky (Ericsson)
R. Gandhi (Cisco)

One-way/Two-way Active Measurement Protocol Extensions for Performance Measurement on LAG

Abstract

This document defines extensions to One-way Active Measurement Protocol (OWAMP), and Two-way Active Measurement Protocol (TWAMP) 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 the 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 8 September 2022.

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 every 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.

OWAMP [RFC4656] and TWAMP [RFC5357] are two active measurement methods according to the classification given in [RFC7799], which can complement passive and hybrid methods. With both methods, running a single 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 OWAMP and TWAMP 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 SR-Policy [I-D.ietf-spring-segment-routing-policy].

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.

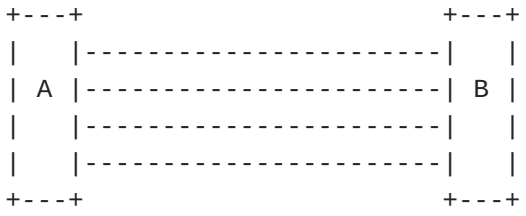


Figure 1: PM for LAG

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 layer, 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.

The micro sessions need to associate with the corresponding member links. For example, when the Server/Reflector/Receiver receives a Test packet, it needs to know from which member link the packet is received, and correlate it with a micro session.

This document defines new command types to indicate the set of micro sessions of a LAG. The details are described in Sections 3 and 4 of this document. Upon receiving a Test packet, the receiver uses the receiving link's identifier to correlate the packet to a particular micro session. In addition, Test packets MAY carry the member link information for validation check. For example, when a micro Session-Sender receives a reflected Test packet, it may need to check whether the Test packet is from the expected member link.

3. Mirco OWAMP Session

This document assumes that the OWAMP Server and the OWAMP Receiver of an OWAMP micro session are at the same end point.

3.1. Micro OWAMP-Control

To support the micro OWAMP session, a new command, Request-OW-Micro-Sessions (TBD1), is defined in this document. The Request-OW-Micro-Sessions command is based on the OWAMP Request-Session command, and uses the message format as described in Section 3.5 of OWAMP [RFC4656]. Test session creation of micro OWAMP session follows the same procedure as defined in Section 3.5 of OWAMP [RFC4656] with the following additions:

When an OWAMP Server receives a Request-OW-Micro-Sessions command, if the request is accepted, the OWAMP Server MUST build a set of micro sessions for all the member links of the LAG from which the Request-OW-Micro-Sessions message is received.

3.2. Micro OWAMP-Test

Micro OWAMP-Test reuses the OWAMP-Test packet format and procedures as defined in Section 4 of OWAMP [RFC4656] with the following additions:

The micro OWAMP Sender MUST send the micro OWAMP-Test packets over the member link with which the session is associated. When receives a Test packet, the micro OWAMP receiver MUST use the member link from which the Test packet is received to correlate the micro OWAMP

session. If there is no such a session, the Test packet MUST be discarded.

4. Mirco TWAMP Session

As above, this document assumes that the TWAMP Server and the TWAMP Session-Reflector of a micro OWAMP session are at the same end point.

4.1. Micro TWAMP-Control

To support the micro TWAMP session, a new command, Request-TW-Micro-Sessions (TBD2), is defined in this document. The Request-TW-Micro-Sessions command is based on the TWAMP Request-Session command, and uses the message format as described in Section 3.5 of TWAMP [RFC5357]. Test session creation of micro TWAMP session follows the same procedure as defined in Section 3.5 of TWAMP [RFC5357] with the following additions:

When a TWAMP Server receives a Request-TW-Micro-Sessions command, if the request is accepted, the TWAMP Server MUST build a set of micro sessions for all the member links of the LAG from which the Request-TW-Micro-Sessions message is received.

4.2. Micro TWAMP-Test

The micro TWAMP-Test protocol is based on the TWAMP-Test protocol [RFC5357] with the following extensions.

4.2.1. Sender Packet Format and Content

The micro TWAMP Session-Sender packet format is based on the TWAMP Session-Sender packet format as defined in Section 4.1.2 of [RFC5357]. Two new fields (Sender Micro-session ID and Reflector Micro-session ID) are added to carry the LAG member link identifiers.

For unauthenticated mode, the format is as below:

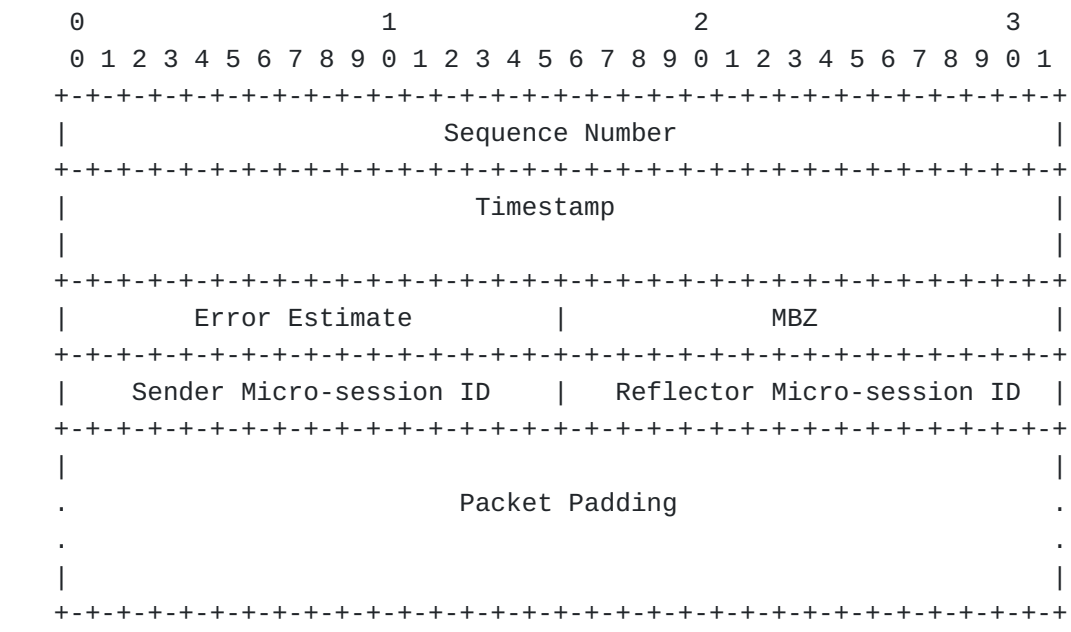


Figure 2: Micro Session-Sender Packet format in Unauthenticated Mode

For authenticated mode, the format is as below:

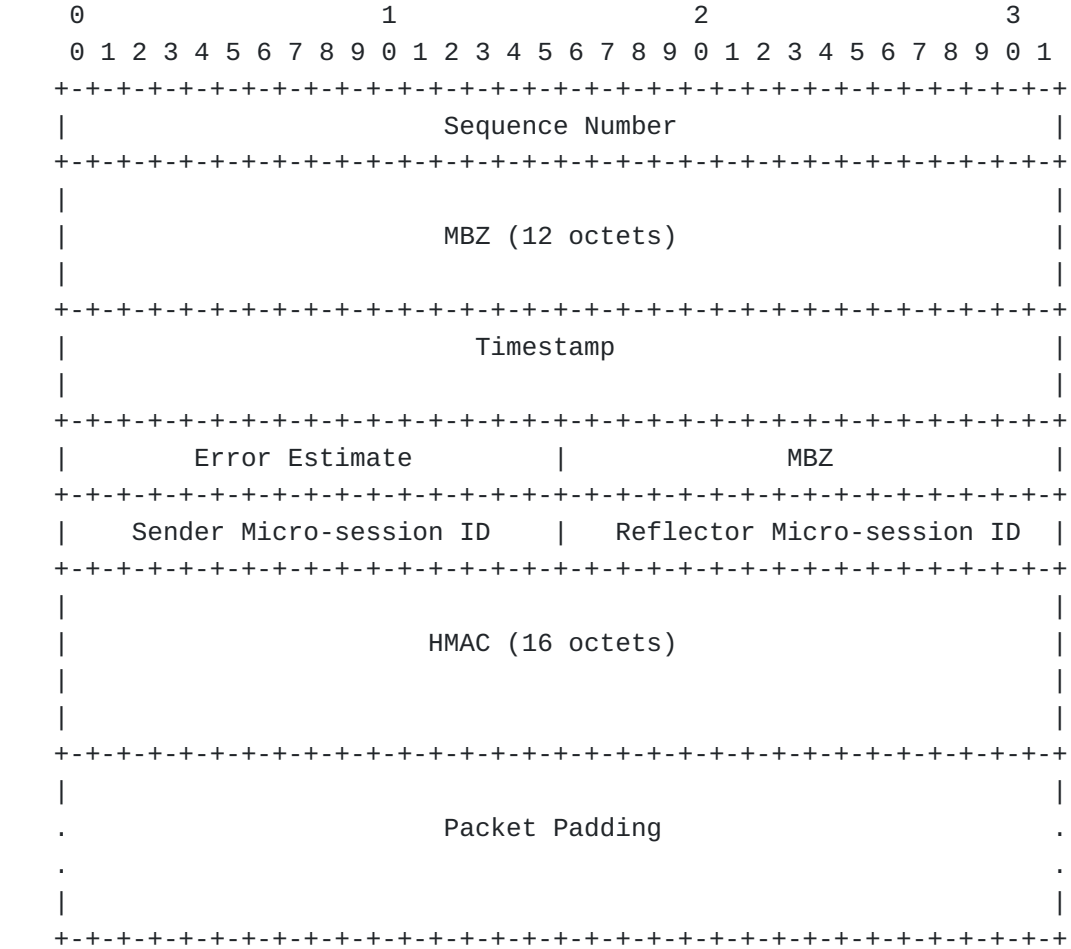


Figure 3: Micro Session-Sender Packet Format in Authenticated Mode

Except for the Sender/Reflector Micro-session ID field, all the other fields are the same as defined in Section 4.1.2 of TWAMP [RFC5357], which is defined in Section 4.1.2 of OWAMP [RFC4656]. Therefore, it follows the same procedure and guidelines as defined in Section 4.1.2 of TWAMP [RFC5357].

*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 Micro-session 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 Micro-session ID MUST be unique at the Session-Reflector.

4.2.2. Sender Behavior

The micro TWAMP Session-Sender inherits the behaviors of the TWAMP Session-Reflector as defined in Section 4.1 of [RFC5357]. In addition, the micro TWAMP Session-Sender MUST send the micro TWAMP-Test packets over the member link with which the session is associated.

When sending the Test packet, the micro TWAMP Session-Sender MUST put the Sender member link identifier that is associated with the micro TWAMP session in the Sender Micro-session ID. If the Session-Sender knows the Reflector member link identifier, the Reflector Micro-session ID field (see Figure 2 and Figure 3) MUST be set. Otherwise, the Reflector Micro-session ID field MUST be zero.

A Test packet with Sender member link identifier is sent to the Session-Reflector, and then is reflected with the same Sender member link identifier. So the Session-Sender can use the Sender member link identifier to check whether a reflected Test packet is received from the member link associated with the correct micro TWAMP session.

The Reflector member link identifier carried in the Reflector Micro-session ID field is used by the Session-Receiver to check whether a Test packet is received from the member link associated with the correct micro TWAMP session. It means that the Session-Sender has to learn the Reflector member link identifier. Once the Session-Sender knows the Reflector member link identifier, it MUST put the identifier in the Reflector Micro-session ID field (see Figure 2 or Figure 3) of the Test packets that will be sent to the Session-Reflector. The Reflector member link identifier can be obtained from pre-configuration or learned from the data plane (e.g., the

reflected Test packet). How to obtain/learn the Reflector member link identifier is out of the scope of this document.

When receiving a reflected Test packet, the micro TWAMP Session-Sender MUST use the receiving member link to correlate the reflected Test packet to a micro TWAMP session. If there is no such a session, the reflected Test packet MUST be discarded. If a matched session exists, 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.

4.2.3. Reflector Packet Format and Content

The micro TWAMP Session-Reflector packet format is based on the TWAMP Session-Reflector packet format as defined in Section 4.2.1 of [RFC5357]. Two new fields (Sender and Reflector Micro-session ID) are added to carry the LAG member link identifiers.

For unauthenticated mode, the format is as below:

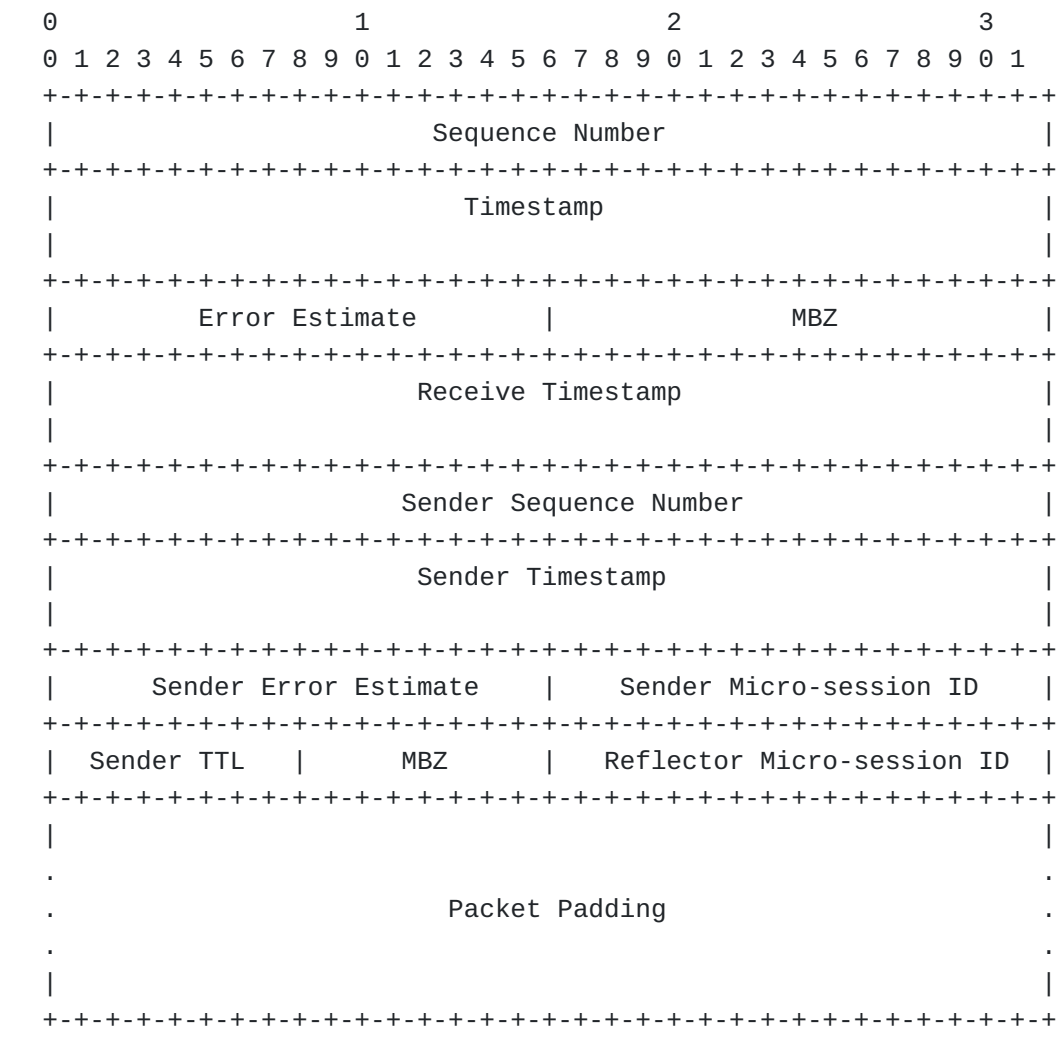
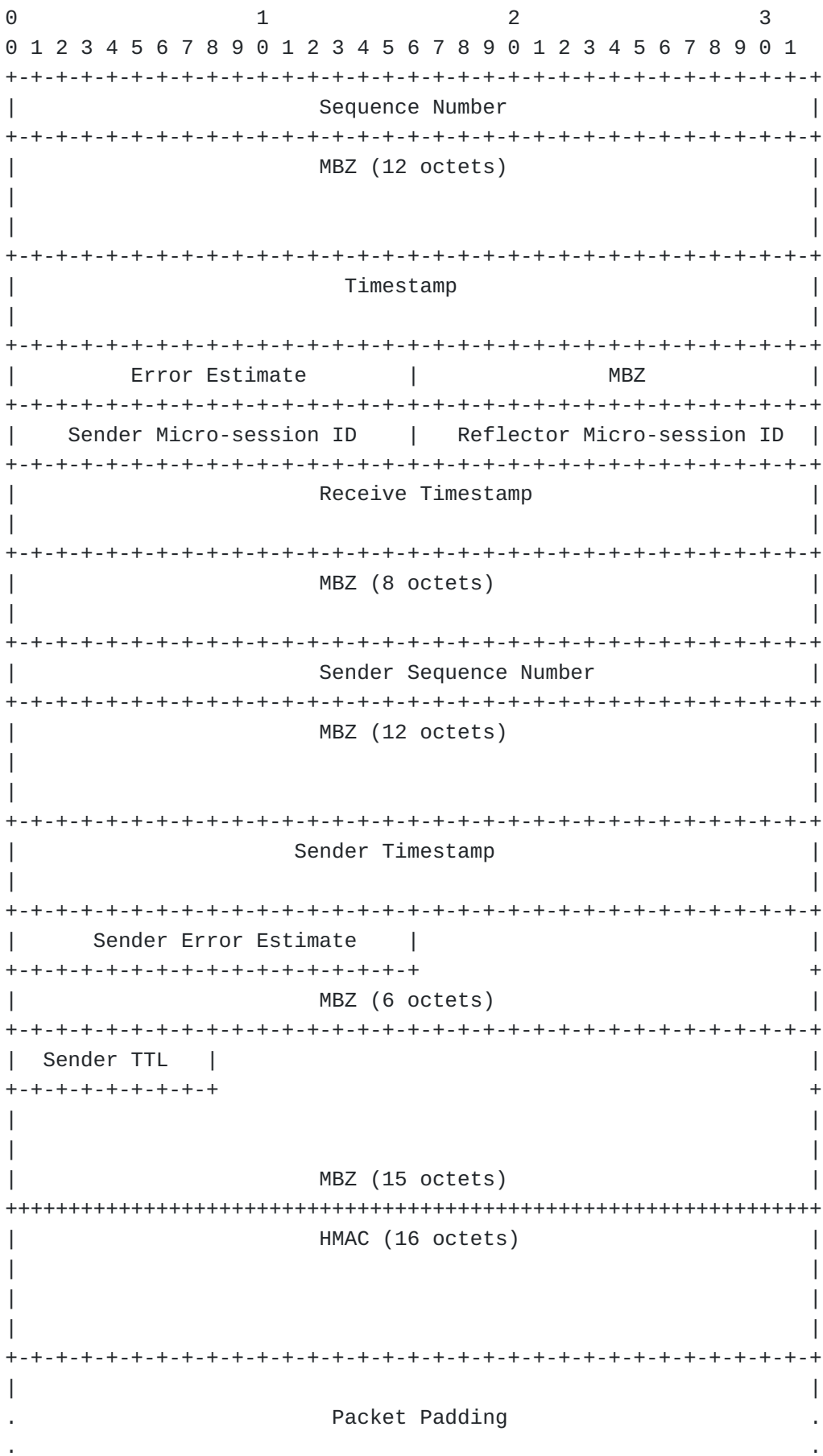


Figure 4: Micro Session-Reflector Packet Format in Unauthenticated Mode

For authenticated mode, the format is as below:


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
MBZ (12 octets)
Timestamp
Error Estimate
MBZ
Sender Micro-session ID
Reflector Micro-session ID
Receive Timestamp
MBZ (8 octets)
Sender Sequence Number
MBZ (12 octets)
Sender Timestamp
Sender Error Estimate
MBZ (6 octets)
Sender TTL
MBZ (15 octets)
HMAC (16 octets)
Packet Padding

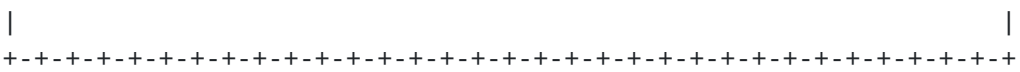


Figure 5: Micro Session-Reflector Packet Format in Authenticated Mode

Except for the Sender/Reflector Micro-session ID field, all the other fields are the same as defined in Section 4.2.1 of TWAMP [RFC5357]. Therefore, it follows the same procedure and guidelines as defined in Section 4.2.1 of TWAMP [RFC5357].

*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 Micro-session 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 Micro-session ID MUST be unique at the Session-Reflector.

4.2.4. Reflector Behavior

The micro TWAMP Session-Reflector inherits the behaviors of a TWAMP Session-Reflector as defined in Section 4.2 of [RFC5357].

In addition, when receiving a Test packet, the micro TWAMP Session-Reflector MUST use the receiving member link to correlate the Test packet to a micro TWAMP session. If there is no such a session, the Test packet MUST be discarded. If the Reflector Micro-session ID is not zero, the Reflector MUST use the Reflector Micro-session ID to validate whether it associates with the receiving member link. If the validation fails, the Test packet MUST be discarded.

When sending a response to the received Test packet, the micro TWAMP Session-Reflector MUST copy the Sender member link identifier from the received Test packet and put it in the Sender Micro-session ID field of the reflected Test packet (see Figure 4 and Figure 5). In addition, the micro TWAMP Session-Reflector MUST fill the Reflector Micro-session ID field (see Figure 2 and Figure 3) of the reflected Test packet with the member link identifier that is associated with the micro TWAMP session.

5. IANA Considerations

5.1. Mico OWAMP-Control Command

This document requires the IANA to allocate the following command type from OWAMP-Control Command Number Registry.

Value	Description	Semantics Definition
TBD1	Request-OW-Micro-Sessions	This document, Section 3.1

5.2. Mico TWAMP-Control Command

This document requires the IANA to allocate the following command type from TWAMP-Control Command Number Registry.

Value	Description	Semantics Definition
TBD2	Request-TW-Micro-Sessions	This document, Section 4.1

6. Security Considerations

This document does not introduce additional security requirements and mechanisms other than those described in [RFC4656], and [RFC5357].

7. Acknowledgements

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

8. References

8.1. 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>.

[RFC4656] Shalunov, S., Teitelbaum, B., Karp, A., Boote, J., and M. Zekauskas, "A One-way Active Measurement Protocol (OWAMP)", RFC 4656, DOI 10.17487/RFC4656, September 2006, <https://www.rfc-editor.org/info/rfc4656>.

[RFC5357] Hedayat, K., Krzanowski, R., Morton, A., Yum, K., and J. Babiarz, "A Two-Way Active Measurement Protocol (TWAMP)", RFC 5357, DOI 10.17487/RFC5357, October 2008, <https://www.rfc-editor.org/info/rfc5357>.

[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>.

[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>.

8.2. Informative References

[I-D.ietf-spring-segment-routing-policy]
Filsfils, C., Talaulikar, K., Voyer, D., Bogdanov, A., and P. Mattes, "Segment Routing Policy Architecture", Work in Progress, Internet-Draft, draft-ietf-spring-segment-routing-policy-18, 17 February 2022, <https://www.ietf.org/archive/id/draft-ietf-spring-segment-routing-policy-18.txt>.

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

Authors' Addresses

Zhenqiang Li
China Mobile
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