IPPM Working Group G. Mirsky
Internet-Draft ZTE Corp.
Intended status: Informational W. Lingqiang
Expires: August 31, 2018 G. Zhui
ZTE Corporation
February 27, 2018

Hybrid Two-Step Performance Measurement Method
draft-mirsky-ippm-hybrid-two-step-00

Abstract

Development of and advancements in automation of network operations brought new requirements toward measurement methodology. Among them is ability to collect the instant telemetry as the packet being processed by the networking elements along its path through the domain. This document introduces new hybrid measurement method, referred to as hybrid two-step, as it separates act of measuring and/or calculating performance metric from the act of collecting and transporting telemetry.

Status of This Memo

This Internet-Draft will expire on August 31, 2018.

Copyright Notice

Table of Contents

1. Introduction

Successful resolution of challenges of automated network operation, as part of overall life-cycle service orchestration, relies on collection of accurate and timely information that reflects the state of network elements on unprecedented massive, even grandiose scale. Because analysis and action upon the it requires considerable computing and storage resources, the network state information, also referred to as telemetry, is unlikely to be processed by network elements themselves but will be relayed into data lakes. The process of producing telemetry information, collecting and transporting it for post-processing should equally work with data flows and specially inserted in the network test packets. Per [RFC7799] classification such process classified as hybrid measurement method.

Several technical methods were proposed to enable collection of telemetry information instantaneous to the packet processing. Among them [P4.INT] and [I-D.ietf-ippm-ioam-data].

This document introduces new hybrid measurement method, referred to as Hybrid Two-step (HTS), that it separates measuring and/or calculating performance metric from the collecting and transporting telemetry. The hybrid two-step method extends two-step mode of Residence Time Measurement (RTM) defined in [RFC8169] to on-path telemetry collection and transport.

2. Conventions used in this document

2.1. Terminology

RTM	Residence Time Measurement
ECMP	Equal Cost Multipath
MTU	Maximum Transmission Unit
HTS	Hybrid Two-step

2.2. 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 BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

3. Problem Overview

Performance measurements are meant to provide data that characterize conditions experienced by data in the network and possibly trigger operations to re-route flows, allocate additional or free excess of resources. All changes to the network depend on the quality of collected data and calculated based on its performance metrics. The quality of measurements defined not only by resolution but by how consistent are performed measurements, how predictable is the moment of measurement making, of obtaining the data. Consider case of delay measurement that relies on collection of time of packet arrival at the ingress interface and time of packet transmission at egress interface. The ideal method may read wall clock value as the very first octet of the packet being received at ingress, and another value, as the first octet being transmitted. That way all nodal processing delays be accounted for as this method excludes packet queuing. But if the measurement method requires the original packet to carry either both time values of the calculated delay value, then the packet must be modified on-the-fly, while being transmitted. And that task may become even more challenging if the packet is encrypted. As result, at egress time may be obtained before the packet transmission begins, thus leaving variable delays unmeasured. Similar problem may cause lower quality of, for example, information that characterizes utilization of the egress interface. If unable to obtain the data consistently, without variable delays for additional processing, information may not accurately reflect the state at the egress interface. To mitigate this problem [RFC8169] defined RTM two-step mode.

Another challenge facing methods that collect telemetry into the actual data packet is risk of exceeding size of Maximum Transmission Unit (MTU), particularly if the packet traverses overlay domains or VPNs. Since the fragmentation is not available at the transport network, operators may have to reduce MTU size advertised to client layer or risk missing telemetry data for the part, most probably the latter part, of the path.

4. Theory of Operation

HTS method consists of the two phases:

- performing a measurement or obtaining telemetry information, one or more than one type, on a node;

- collecting and transporting the measurement.

HTS uses HTS Control message to define types of measurement or telemetry data collection requested from a node. HTS Control message may be inserted into the data packet, as meta-data or shim, or be transmitted in the specially constructed test packet.

To collect measurement and telemetry data from the nodes HTS method uses the follow-up packet. The node that creates the HTS Control message also originates the HTS follow-up packet. The follow-up packet contains characteristic information, copied from the data packet, sufficient for participating nodes to associate it with the original packet. Exact composition of the characteristic information is specific for each transport network and its definition is outside the scope of this document. The follow-up packet also uses the same encapsulation as the data packet. If not payload but only network information used to load-balance flows in equal cost multipath (ECMP), use of the network encapsulation identical to the data packet should guarantee that the follow-up packet remains in-band, i.e. traverses the same set of network elements, with the original data packet. Only one outstanding follow-up packet may be on the node for the given path. That means that if the node receives HTS Control message for the flow on which it still waits for the follow-up packet to the previous HTS Control message, the node will originate the follow-up packet to transport the former set of the telemetry data and transmit it before it transmits the follow-up packet with the latest set of telemetry information.

5. IANA Considerations

This document doesn't have any IANA requirements. The section may be deleted before the publication.

6. Security Considerations

Nodes that practice HTS method are presumed to share a trust model that depends on the existence of a trusted relationship among them. This is necessary as these nodes are expected to correctly modify specific content of the data in the follow-up packet, and degree to which HTS measurement is useful for network operation depends on this ability. In practice, this means that those portions of messages that contain the telemetry data cannot be covered by either confidentiality or integrity protection. Though there are methods that make it possible in theory to provide either or both such protections and still allow for intermediate nodes to make detectable but authenticated modifications, such methods do not seem practical at present, particularly for protocols that used to measure latency and/or jitter.

The ability to potentially authenticate and/or encrypt the telemetry data for scenarios both with and without participation of intermediate nodes that participate in HTS measurement is left for further study.

While it is possible for a supposed compromised node to intercept and modify the telemetry information in the follow-up packet, this is an issue that exists for nodes in general - for any and all data that may be carried over the particular networking technology - and is therefore the basis for an additional presumed trust model associated with existing network.

7. Acknowledgements

TBD

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>.

[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>.

8.2. Informative References

[I-D.ietf-ippm-ioam-data]
Brockners, F., Bhandari, S., Pignataro, C., Gredler, H., Leddy, J., Youell, S., Mizrahi, T., Mozes, D., Lapukhov, P., Chang, R., and d. daniel.bernier@bell.ca, "Data Fields for In-situ OAM", draft-ietf-ippm-ioam-data-01 (work in progress), October 2017.

[P4.INT] "In-band Network Telemetry (INT)", P4.org Specification, October 2017.

[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>.

[RFC8169] Mirsky, G., Ruffini, S., Gray, E., Drake, J., Bryant, S., and A. Vainshtein, "Residence Time Measurement in MPLS Networks", RFC 8169, DOI 10.17487/RFC8169, May 2017, <https://www.rfc-editor.org/info/rfc8169>.

Authors' Addresses

Greg Mirsky
ZTE Corp.

Email: gregimirsky@gmail.com

Wang Lingqiang
ZTE Corporation
No 19 ,East Huayuan Road
Beijing 100191
P.R.China

Phone: +86 10 82963945
Email: wang.lingqiang@zte.com.cn

Guo Zhui
ZTE Corporation
No 19 ,East Huayuan Road
Beijing 100191
P.R.China

Phone: +86 10 82963945
Email: guo.zhui@zte.com.cn