bmwg L. Contreras
Internet-Draft Telefonica
Intended status: Informational 7 July 2025
Expires: 8 January 2026

Calibration of Measured Time Values between Network Elements
draft-contreras-bmwg-calibration-03

Abstract

Network devices are incorporating capabilities for time stamping certain packets so that such time stamps can reference the time at which a packet passes through a given device (at ingress or egress). Those stamps or marks are relevant to calculating the measured delay and can be used for traffic engineering purposes.

To ensure consistency and accuracy across different network element implementations, a benchmarking procedure is necessary to calibrate the timestamps. Such a procedure can permit the identification and correction of any deviations or biases in the time stamps produced by diverse devices. This document proposes a methodology for Calibrating the measurements from different network element implementations in a common measurement scenario.

Table of Contents

1. Introduction

Latency is becoming one of the major drivers for network traffic engineering and planning due to the introduction of novel services sensitive to both latency and its variability.

Network devices incorporate capabilities for time-stamping certain packets so that such time stamps can reference the time at which a packet passes through a given device (at ingress or egress). Those stamps or marks are relevant to calculating the measured delay and can be used for traffic engineering purposes. That is the case of the Flex Algo algorithms [RFC9502] [RFC9351] aiming to represent low latency topologies being based on the time reference measured among devices. Furthermore, services sensitive to latency, such as the deterministic ones [RFC8557] or those following the network slicing paradigm [RFC9543], require from accurate information in order to place flows in the network.

The measurements here considered are those performed between neighbor nodes in a network, so that an end-to-end delay metric can be composed by aggregating partial measurements. To ensure consistency and accuracy across different network element implementations, a benchmarking procedure is necessary to calibrate the timestamps. This allows correction of observed measurements and improves the reliability of latency-based traffic engineering decisions.

Different protocols are implemented nowadays on network devices to perform such measurements, such as TWAMP-light [RFC5357] [RFC8545] or STAMP [RFC8972]. It is then important to ensure that the time stamp produced by the network device is accurate so that the traffic engineering decisions based on such measurements are correct. Differences on accuracy of the measurements among devices can be due to different causes, such as different solutions for time stamping (e.g., hardware-based vs software-based), different module involved in the time stamping process (e.g., central processor vs line card), different type of component taking care of the timestamping (e.g., Network Processor, ASIC, hardware accelerator, etc), etc.

The traffic engineering decisions will be typically implemented by centralized controllers, which can process the measurements and decide on the specific traffic engineering path. This can be the case of the Path Computing Element (PCE) [RFC5440]. It is important to note that any deviations or bias on the time stamps of the packets can lead to inaccurate calculations or decisions.

An example is the following: Network Device A can introduce some delay (e.g., due to less powerful processing capabilities) in the timestamp versus Network Device B. In a dual-plane backbone, with similar characteristics in terms of physical path length, where one of the planes is built with network devices of type A while the other one is built with network devices of type B, the difference between the time stamps in A and B can lead to all the traffic being delivered by one of the planes if low latency flex algo topology is used.

The purpose of this document is to present a benchmarking procedure to calibrate the time stamps of different network device implementations so that the observed measured can be corrected if needed during the processing of the measured time-stamped data.

2. Baseline measurement scenario

In order to create a reference or baseline to later on compare the measurements obtained from distinct network elements, a reference scenario is defined. This reference scenario will serve to calibrate the time stamps of the different network devices.

The reference scenario assume an instrumentation device generating and receiving TWAMP-light or STAMP packets. Those packets are delivered on a fiber spool of a given known length, L. Typical lengths for testing purposes are in the order of 20 - 100 kms. It can be assumed a delay of 5 us/km on that fiber spool [ITU-T-G.114]. The measurement setup is as follows:

```
                          +------------+
                          |            |
             +------------|   tester   |<-------------+
             |            |            |              |
             |            +------------+              |
             |                                        |
             |                                        |
             |        Fiber spool of length L         |
             +----------------------------------------+
```

Figure 1: Baseline measurement scenario

The tester will generate the baseline measurement for the time spent on delivering packets through the link represented by the fiber spool. Such a reference value can be defined as T_baseline and will serve for comparison of the measurements performed by the distinct implementations of network elements for calibration purposes.

3. Network element calibration tests

Similar to the case before, it is possible to perform the same kind of test with actual network elements. Two scenarios are possible.

* Performing the test with just one single network element, similar to the baseline scenario. This can be useful for characterizing the behavior of one specific network device implementation, for both hardware and software.

* Performing the test between two network elements. This can be useful for determining interworking scenarios, considering either network elements from the same vendor but with different characteristics (e.g., network device model, hardware or software characteristics, etc), or network elements from different vendors.

3.1. Single network element test

The scenario is similar to the baseline scenario, as follows.

```
                               +------------+
                               |  network   |
                  +------------|  element   |<-------------+
                  |            |     A      |              |
                  |            +------------+              |
                  |                                        |
                  |                                        |
                  |        Fiber spool of length L         |
                  +----------------------------------------+
```

Figure 2: Single network element test

As before, the network element generates measurement packets through the link represented by the fiber spool. The obtained value can be referred as T_ne.

The calibration exercise results from the comparison between T_baseline and T_ne.

Note for discussion: consider the implications of connecting the fiber spool to ports / line cards of different characteristics in terms of hardware implementation. Maybe a restriction should be imposed in this kind of test to be performed considering the same kind of termination for the fiber spool.

3.2. Paired network elements test

This scenario considers the measurement to be performed between a couple of network elements, as follows.

```
               +------------+                         +------------+
               |  network   | Fiber spool of length L |  network   |
               |  element   |<----------------------->|  element   |
               |     A      |                         |     B      |
               +------------+                         +------------+
```

Figure 3: Paired network elements test

The network elements can be from the same or different vendor. The purpose of this test is to calibrate measurements when the devices involved represent different implementations of the same network device functionality, including the protocol used for time stamping.

In this case, one of the network elements generates measurement packets through the link represented by the fiber spool of known length (and previously measured in the reference scenario). The obtained value can be referred as T_ne_ab, when the measurement is triggered from A to B. Similarly, it is possible to obtain T_ne_ba on the other direction.

Assuming that both T_ne_ab and T_ne_ba produce the same result, the calibration exercise results from the comparison between T_baseline and T_ne_ab (or T_ne_ba).

3.3. Measurement conditions

The measurements proposed should specify:

* Network device model, hardware and software description.

* Length of the fiber spool used as reference for the measurement.

* Description of the ports / line card where the fiber spool is connected, since different line cards could produce measurements with different accuracy levels.

* Protocol used in the measurement (e.g., TWAMP-light, STAMP, ...).

* Duration of the test for statistical consistency (e.g., period for average, min, max calculations)

* Network tester used for generating the baseline values.

(Note: to consider scalability aspects on the collected measurements that could impact the calibration exercise. For instance, if multiple sessions are running in parallel from one node to several neighbor nodes, with one session per neighbor).

(Note: to be discussed the possibility of considering asymmetric scenarios in the connection of network elements).

(Note: to assess the extendibility of this approach to any network element, e.g. switches).

3.4. Resolution of the calibration process

The calibration process of the time-stamps observed during the benchmarking methodology needs to take into account the time units of relevance for the purpose of the usage of such measurements.

The timestamp format used in TWAMP, TWAMP Light, and STAMP protocols is based on the NTP format, as defined in [RFC8877]. These protocols employ a 64-bit timestamp, where:

* The first 32 bits represent the number of seconds since the NTP epoch (January 1, 1900)

* The remaining 32 bits represent fractional seconds.

This structure provides a theoretical resolution of approximately 233 picoseconds.

For instance, with this in mind, if the purpose is the generation of FlexAlgo topologies based on delay, it should be taken into account that the delay field advertised when using IS-IS [RFC8570] or OSPF [RFC7471] is expressed in microseconds. Thus, differences below a microsecond are not relevant. OSPF or IS-IS with traffic engineering extensions encode the "delay" value in microseconds using 24-bit fields. This naturally results in lower precision compared to the timestamp format used in active measurement protocols.

When calibrating measurements for use in network operations or control plane mechanisms, it is crucial to consider the specific use case of the timestamp to ensure appropriate granularity and accuracy. For example, in the case of constructing low-latency FlexAlgo topologies, it is necessary to align the measurement resolution to microseconds, as this is the resolution that the IGP protocol can propagate towards the control plane generating the low-latency topology. To achieve this, the measured delay collected by means of TWAMP, TWAMP Light, and STAMP protocols SHOULD be adjusted using a ceiling function, i.e., rounding up to the nearest integer value in microseconds.

Other use cases, however, may require different resolution levels, rounding techniques, or adjustments depending on operational goals.

4. Further discussions

This is a placeholder for capturing topics that will be considered for next versions of the document.

4.1. Consideration of hybrid methods (e.g., IOAM) as measurement mechanisms for calibration assessment

On-path hybrid methods, such as In situ Operations, Administration, and Maintenance (IOAM) [RFC9197] [RFC9326] allow the obtention of telemetry information that can be used for measuring delays in network links. The IOAM data can be embedded in different encapsulations.

While these methods coud be used for collecting the metric of interest, in the context of calibrating hop-by-hop measurements, it would impose practical limitations. For instance, the need of generating individual flows per hop for the solely purpose of measuring the link delay. Furthermore, it would require the specification of the measurement conditions in terms of the flows used for embedding the telemetry data.

4.2. Usage of instrumentation instead of physical fiber spools for determining the baseline value

Instrumentation equipment can introduce impairments in a link in a controlled manner, either as a deterministic or statistical manner. For the purpose of the calibration pursued in this document, the usage of instrumentation equipment can be considered instead of physical fiber spools by configuring deterministic delay values.

While this can be a proper approach, it presents some practical shortcomings. For instance, the need of supporting all the variety of interface bit rates of interest in the platforms to be calibrated.

5. Acknowledgements

The author would like to thank Miguel Cros (Telefonica) for the discusions on the topic. Moreover the author would like also thank Giuseppe Fioccola, Tal Mizrahi, Carsten Rossenhoevel for the valuable comments received during the BMWG meetings.

This work has been partially funded by the European Commission through the Horizon Europe SNS JU PREDICT-6G project (GA 101095890).

6. References

6.1. Normative References

[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>.

[RFC5440] Vasseur, JP., Ed. and JL. Le Roux, Ed., "Path Computation Element (PCE) Communication Protocol (PCEP)", RFC 5440, DOI 10.17487/RFC5440, March 2009, <https://www.rfc-editor.org/info/rfc5440>.

[RFC7471] Giacalone, S., Ward, D., Drake, J., Atlas, A., and S. Previdi, "OSPF Traffic Engineering (TE) Metric Extensions", RFC 7471, DOI 10.17487/RFC7471, March 2015, <https://www.rfc-editor.org/info/rfc7471>.

[RFC8545] Morton, A., Ed. and G. Mirsky, Ed., "Well-Known Port Assignments for the One-Way Active Measurement Protocol (OWAMP) and the Two-Way Active Measurement Protocol (TWAMP)", RFC 8545, DOI 10.17487/RFC8545, March 2019, <https://www.rfc-editor.org/info/rfc8545>.

[RFC8570] Ginsberg, L., Ed., Previdi, S., Ed., Giacalone, S., Ward, D., Drake, J., and Q. Wu, "IS-IS Traffic Engineering (TE) Metric Extensions", RFC 8570, DOI 10.17487/RFC8570, March 2019, <https://www.rfc-editor.org/info/rfc8570>.

[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>.

[RFC9197] Brockners, F., Ed., Bhandari, S., Ed., and T. Mizrahi, Ed., "Data Fields for In Situ Operations, Administration, and Maintenance (IOAM)", RFC 9197, DOI 10.17487/RFC9197, May 2022, <https://www.rfc-editor.org/info/rfc9197>.

[RFC9326] Song, H., Gafni, B., Brockners, F., Bhandari, S., and T. Mizrahi, "In Situ Operations, Administration, and Maintenance (IOAM) Direct Exporting", RFC 9326, DOI 10.17487/RFC9326, November 2022, <https://www.rfc-editor.org/info/rfc9326>.

[RFC9351] Talaulikar, K., Ed., Psenak, P., Zandi, S., and G. Dawra, "Border Gateway Protocol - Link State (BGP-LS) Extensions for Flexible Algorithm Advertisement", RFC 9351, DOI 10.17487/RFC9351, February 2023, <https://www.rfc-editor.org/info/rfc9351>.

[RFC9502] Britto, W., Hegde, S., Kaneriya, P., Shetty, R., Bonica, R., and P. Psenak, "IGP Flexible Algorithm in IP Networks", RFC 9502, DOI 10.17487/RFC9502, November 2023, <https://www.rfc-editor.org/info/rfc9502>.

6.2. Informative References

[ITU-T-G.114]
"One-way transmission time", May 2003.

[RFC8557] Finn, N. and P. Thubert, "Deterministic Networking Problem Statement", RFC 8557, DOI 10.17487/RFC8557, May 2019, <https://www.rfc-editor.org/info/rfc8557>.

[RFC8877] Mizrahi, T., Fabini, J., and A. Morton, "Guidelines for Defining Packet Timestamps", RFC 8877, DOI 10.17487/RFC8877, September 2020, <https://www.rfc-editor.org/info/rfc8877>.

[RFC9543] Farrel, A., Ed., Drake, J., Ed., Rokui, R., Homma, S., Makhijani, K., Contreras, L., and J. Tantsura, "A Framework for Network Slices in Networks Built from IETF Technologies", RFC 9543, DOI 10.17487/RFC9543, March 2024, <https://www.rfc-editor.org/info/rfc9543>.

Author’s Address

Luis M. Contreras
Telefonica
Ronda de la Comunicacion, s/n
28050 Madrid
Spain
Email: luismiguel.contrerasmurillo@telefonica.com