

SFC Working Group
INTERNET-DRAFT
Intended Status: Standards Track

Anil Kumar S N
Gaurav Agrawal
Vinod Kumar S
Huawei Technologies India
June 28, 2016

Expires: December 30, 2016

Packet Delay Measurement for SFC
draft-agv-sfc-packet-delay-measurement-01

Abstract

This document describes performance measurement(PM) packet delay measurement for Service Function Chains (SFCs) to assess the operational status and behavior of a SF, a subset of SFs, a whole SFC as a function of the actual deterministic SF/SFC. Packet loss mechanism described in this document is based on [[SFC-PM-arch](#)].

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 <http://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 December 30, 2016

Copyright and License Notice

Copyright (c) 2016 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to [BCP 78](#) and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1	Introduction	4
1.1	Terminology	5
1.2	Terms & Definition	5
1.3	Applicability and Scope	5
2	Message Format	6
2.1	Performance Measurement Context Header	6
2.2	Packet Delay PM Context Header	6
2.3	Performance Measurement Flow ID	8
3	Operations	9
3.1	Delay Measurement Options	9
3.2	Packet Delay Parameters for Option 1	9
3.3	Packet Delay Parameters for Option 2	10
3.4	Initiating a Packet Delay Measurement	11
3.5	Encapsulation of Classified Packets	11
3.6	Incoming Packets Processing at MA	12
3.7	Outgoing Packets Processing at MA	13
3.8	MA Reporting the Statistics to Collector	13
3.9	Packet Delay Calculation at Collector	14
3.9.1	Packet Delay Calculation at Collector (option 1)	14
3.9.1	Packet Delay Calculation at Collector (option 2)	14
4	Security Considerations	15
5	IANA Considerations	15
5.1	TLV Class	15
6	Acknowledgments	15
7	References	16
7.1	Normative References	16

7.2 Informative References	16
Authors' Addresses	17

1 Introduction

The delivery of end-to-end services often requires various service functions. These include traditional network service functions such as firewalls and traditional IP Network Address Translators (NATs), as well as application-specific functions. The definition and instantiation of an ordered set of service functions and subsequent "steering" of traffic through them is termed Service Function Chaining (SFC).

The common reasons for packet delay are nodal processing (Algorithmic, Packetization etc...), queuing, transmission delay, propagation delay and service function processing delay.

Proper operation of a SFC depends on the ability to monitor and quickly identify faults and focus attention on the root cause of the problem. Also service provider service level agreements (SLAs) depends on the capability to measure and monitor performance metrics for packet delay. SFC delay measurement is one of the important tool to achieve above requirements.

Packet delay measurement capability provides operators with greater visibility into the performance characteristics of their networks, thereby facilitating planning, troubleshooting, and network performance evaluation.

This document specifies best possible efficient and accurate mechanism for packet delay measurement for Service Function Chains (SFCs) for a SFC network domain.

The packet delay measurement described in this document is realized by adding a new context header in the network service header.

1.1 Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC 2119](#) [RFC2119].

1.2 Terms & Definition

Refer to [RFC 7665](#) for definitions of SFC terms.

Measurement Collector: An operational function that collects measurement data from a measurement agent. Measurement collector is responsible for collecting the performance measurement data from measurement agent. Measurement collector functionality could be integrated with one of the MA or with the controller itself.

Measurement Agent: An operational function that contains one or more measurement functions. Measurement agents is responsible for understand and analyze performance measurement control information encoded in NSH metadata and perform the performance data collecting and report the same to measurement collector with key information to identify performance measurement instance along with data collected.

Measurement Controller: An operational function that controls running, scheduling, and general coordination of measurement functions by instructing a measurement agent using NSH metadata. Measurement controller is responsible for configuring the performance measurement instance. Optionally performance measurement instance can be configured manually at the ingress in which case controller is not required

1.3 Applicability and Scope

This document defines the implementation mechanism for the packet delay measurement as per performance measurement architecture [SFC-PM-arch]. This document defines a new NSH message format for carrying packet delay measurement related control information. It also defines operations to be carried out for packet loss measurement.

Communication mechanism between measurement controller, measurement collector and MA is out of scope of this document.

2 Message Format

2.1 Performance Measurement Context Header

The format of the performance measurement context headers, is as described below.

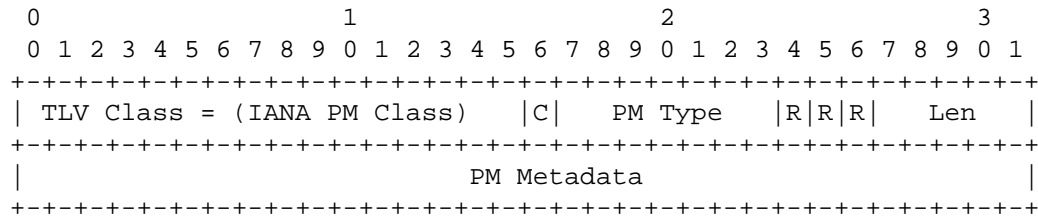


Figure 1: Performance measurement context header

TLV Class: As per [I-D.ietf-sfc-nsh] the TLV class value for performance measurement needs to be applied from IANA.

PM Type: indicates the type of performance measurement that needs to be carried out by the MA.

Length: Length of the variable PM metadata, in 4-byte words.

2.2 Packet Delay PM Context Header

The format of the packet loss PM context headers, is as described below.

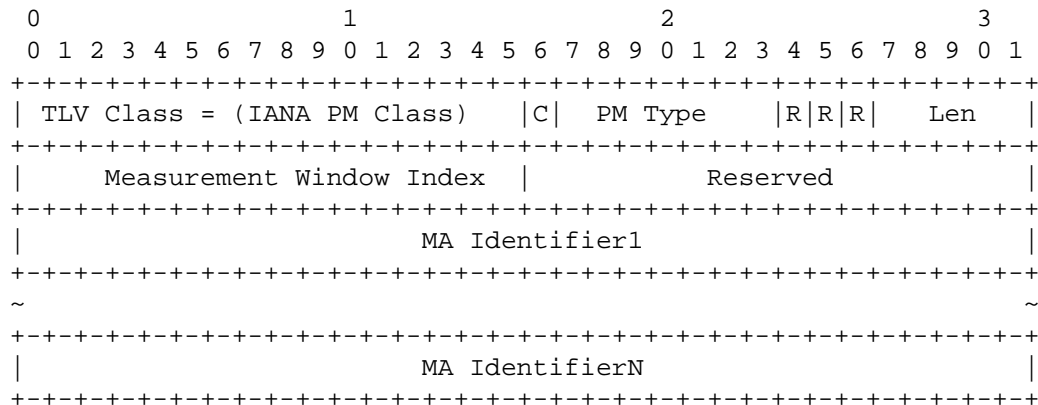


Figure 2: Packet loss PM context header

TLV Class: As per [I-D.ietf-sfc-nsh] the TLV class value for performance measurement needs to be applied from IANA.

PM Type: Indicates the type of performance measurement that needs to be carried out by the MA. (IANA assigned PM type PM delay value)

This memo defines the following values.

0x6 - Packet delay : indicates PM meta data contains info for packet delay. Specified SF(s) SHOULD participate in packet delay PM.

0x7 - SF hop by hop packet delay : indicates PM meta data contains info for packet delay. All the SF(s) in the SFP SHOULD participate in the Packet Delay PM

0x8 - SFF Hop by Hop Packet Delay : indicates PM meta data contains info for packet delay. All the SFF(s) in the SFP SHOULD participate in the packet delay PM

0x9 - All hop by hop packet delay : indicates PM meta data contains info for packet delays. All the SF(s) and SFF(s) in the SFP SHOULD participate in the packet delay PM

0xA - SFP as a whole : indicates PM meta data contains information for packet delay. Classifier (source MA) and SF domain boundary SFF (sink MA) SHOULD participate in the packet delay computation.

Additional values can be added for future PM types and scenarios.

Length: Length of the variable PM metadata, in 4-byte words. It will have a minimum length of 1 indicating the presence of measurement window index.

Measurement Window Index: 16 bit number to denote the current measurement window to which the packet belongs.

Reserved: Reserved 16 bits for future purpose.

MA identifier list: List of participating MA in the SFP.

MA identifier has two parts

- 1) Node identifier - 24 bit
 - a) For SFF: MUST be unique number assigned by controller
 - b) For SF: All zero. Context identifier itself identifies SF node.
- 2) Context identifier - 8 bit
 - a) For SFF: Service index of next SF.
 - b) For SF: Service index

MA identifier SHOULD be in decreasing order of the SI for optimized traversal of the SI participation.

The Length of the packet delay PM context header is fixed to 1

when the PM Type is 2 to 5. Since all SF's in the SFP has to perform the Delay measurement specified in PM Type field.

```

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
+-----+-----+-----+-----+-----+-----+-----+-----+
| TLV Class = (IANA PM Class) |C| Type 2to5 |R|R|R| Len=0x1 |
+-----+-----+-----+-----+-----+-----+-----+-----+
| Measurement Window Index | Reserved |
+-----+-----+-----+-----+-----+-----+-----+

```

2.3 Performance Measurement Flow ID

The method of encoding the PMF id is done using the flow id defined in [I-D.ietf-sfc-nsh].

```

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
+-----+-----+-----+-----+-----+-----+-----+-----+
| TLV Class = 0x0 |C| Type=0x7 |R|R|R| L=0x1 |
+-----+-----+-----+-----+-----+-----+-----+-----+
| Performance Measurement Flow ID |
+-----+-----+-----+-----+-----+-----+-----+

```

Sample encoding of PMF using flow ID

```

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
+-----+-----+-----+-----+-----+-----+-----+-----+
|Ver|O|C|R|R|R|R|R|R| Length | MD-type=0x2 | Next Protocol |
+-----+-----+-----+-----+-----+-----+-----+-----+
| Service Path ID | Service Index |
+-----+-----+-----+-----+-----+-----+-----+-----+
| TLV Class=IANA |C| Type=0x2 |R|R|R| Len |
+-----+-----+-----+-----+-----+-----+-----+-----+
| Measurement Window Index | Reserved |
+-----+-----+-----+-----+-----+-----+-----+-----+
| TLV Class = 0x0 |C| Type=0x7 |R|R|R| L=0x1 |
+-----+-----+-----+-----+-----+-----+-----+-----+
| Performance Measurement Flow ID |
+-----+-----+-----+-----+-----+-----+-----+

```


3 Operations

For packet delay measurement all the MA's time should be synchronized from a centralized stable clock. Among many available time synchronization protocols all MAs's must use same protocol and synchronize time to the same centralized stable clock. Selection of type of time synchronization protocol is for MAs is out of scope of this document.

Every MA needs to maintain the packet delay statistics which they immediately report it to collector or may accumulate and report to collector at a reporting interval which depends on local policy.

3.1 Delay Measurement Options

There are 2 options for packet delay measurement

Option 1: Delay Measurement is performed using single packet in the measurement window.

Option 2: Delay Measurement is performed using multiple packet in a measurement window.

In case of Option 2 is used for delay measurement, MA sums the measured time for the received window and reports the summed time together with the count of samples to the collector. Collector will co-relate the time info received across multiple report and compute the average time for a window.

3.2 Packet Delay Parameters for Option 1

Parameters for Option 1:

1) Rx-Time[P][M][W]

2) Tx-Time[P][M][W]

These parameters will be created and maintained at every MA

Rx-Time[P][M][W]: This 3 dimensional Parameter is maintained at every MA. In this P stands for PMF ID, M stands for MA identifier and W stands for window ID. It will contain the system time (NTP UTC time format) when packets is received for a given PMF ID + MA identifier + window.

PMF ID is unique within a SFC domain, for a given PMF ID there could be multiple windows. This parameter is created when a packet delay PM packet is received at a first time within a reporting interval. The

parameter is deleted upon expiry of reporting interval at which it will be sent to collector and deleted from MA.

Tx-Time[P][M][W]: This 2 dimensional Parameter is maintained at every MA. In this P stands for PMF ID, M stands for MA identifier and W stands for window ID. It will contain the system time (NTP UTC time format) when packets is sent out for a given PMF ID + MA identifier + window.

PMF ID is unique within a SFC domain, for a given PMF ID there could be multiple windows. This Parameter is created when a packet delay PM packet is sent at a first time within a reporting interval. The parameter is deleted upon expiry of reporting interval at which it will be sent to collector and deleted from MA.

3.3 Packet Delay Parameters for Option 2

Parameter for Option 2:

- 1) Rx-Sum-Time[P][M][W]
- 2) Rx-Sample-Count[P][M][W]
- 3) Tx-Sum-Time[P][M][W]
- 4) Tx-Sample-Count[P][M][W]

These parameters will be created and maintained at every MA

Rx-Time[P][M][W]: This 3 dimensional parameter is maintained at every MA. In this P stands for PMF ID, M stands for MA identifier and W stands for window ID. It will contain the sum of system time when packets are received for a given PMF ID + MA identifier + window.

PMF ID is unique within a SFC Domain, for a given PMF ID there could be multiple MA identifier and there could be multiple windows. This parameter is created when a packet delay PM packet is received at a first time within a reporting interval. The parameter is deleted upon expiry of reporting interval at which it will be sent to collector and deleted from MA.

Rx-Sample-Count[P][M][W]: This 3 dimensional parameter is maintained at every MA. In this P stands for PMF ID, M stands for MA identifier and W stands for window ID. It will contain the count of the PM delay sample packets received for a given PMF ID + MA identifier + window.

Tx-Time[P][M][W]: This 3 dimensional parameter is maintained at very MA. In this P stands for PMF ID, M stands for MA identifier and W

stands for window ID. It will contain the sum of system time when packets are sent for a given PMF ID + MA identifier + window.

PMF ID is unique within a SFC Domain, for a given PMF ID there could be multiple MA identifier and there could be multiple windows. This parameter is created when a packet delay PM packet is sent at a first time within a reporting interval. The Parameter is deleted upon expiry of Reporting interval at which it will be sent to Collector and deleted from MA.

Tx-Sample-Count[P][M][W]: This 3 dimensional Parameter is maintained at every MA. In this P stands for PMF ID, M stands for MA identifier and W stands for window ID. It will contain the count of the PM delay sample packets sent for a given PMF ID + MA identifier + window.

3.4 Initiating a Packet Delay Measurement

Measurement controller programs the PM instance at the classifier. Within the PM schedule as programmed in PM instance, classifier starts the packet classification (based on the classification rule programmed in PM instance). The packet classification policy may be customer/network/service specific.

3.5 Encapsulation of Classified Packets

For the classified packet, PM context header is prepared with following information

- Set the type class value to the IANA allocated PM class.
- Set the type value to the required packet delay measurement type.
- If Type value = 0x6
 - o Encode the list of MA identifier that need to participate in packet delay measurement.
- If Value = 0x7, 0x8, 0x9, 0xA
 - o No need to encode SI, since the type dictates the involved participating MA
- Set the window identifier as the current running window number
- Encode the 32 bit globally unique PMF ID using the flow ID context header as defined in [I-D quinn-sfc-nsh-tlv].

3.6 Incoming Packets Processing at MA

On receiving the packet with NSH header following operations are carried out:

Step 1: Detection of PM context header in a packet, by verifying the PM TLV Class as allocated by IANA.
(If not detected, move to step 6)

Step 2: Check if PM type field value is 6 to 10.
(If not move to step 6).

Step 3: If PM Type value = 7 to 10 move directly to step 5

Step 4: Check presence of self MA identifier in MA identifier list
(If not present, move to step 6)

Step 5: If option 1 is used then go to step 5a, if option 2 is used then go to step 5b

Step 5a: Record the value for PMF ID and window ID; record the time when the packet is received Rx-Time[P][M][W].

Step 5b: Record the value for PMF ID and window ID; record the time when the packet is received and sum it up in Rx-Sum-Time[P][M][W]. If Rx-Sum-Time[P][M][W] doesn't exist then create it and initialize it with recorded time value.

Also increment the value of statistics counter Rx-Sample-Count[P][M][W] by 1. If statistics counter doesn't exist, create it and initialize it with 1.

Step 6: Packet is sent for normal processing

3.7 Outgoing Packets Processing at MA

At outgoing port of MA following operations are carried out.

Step 1: If option 1 is used then go to step 2a, if option 2 is used then go to step 2b

Step 2: If option 1 is used then go to step 2a, if option 2 is used then go to step 2b

Step 2a: Record the value for PMF ID and window ID; Record the time when the packet is sent Tx-Time[P][M][W].

Step 2b: Record the value for PMF ID and window ID; Record the time when the packet is sent and sum it up in Tx-Sum-Time[P][M][W]. If Tx-Sum-Time[P][M][W] doesn't exist then create it and initialize it with recorded time value.

Also increment the value of statistics counter Tx-Sample-Count[P][M][W] by 1. If statistics counter doesn't exist, create it and initialize it with 1.

Step 3: Packet is sent for normal processing

3.8 MA Reporting the Statistics to Collector

Reporting timer will run on each MA. Consistency of this timer should be ensured across the entire MA in the SFP, ensuring the same is out of scope of this document.

On expiry of this timer following information needs to be sent to the collector.

- MA identifier
- PM type value
- Value of all the collected Rx-Time[P][M][W], Rx-Sum-Time[P][M][W], & Rx-Sample-Count[P][M][W] parameters along with the corresponding PMF ID and window Id
- Value of all the collected Tx-Time[P][M][W], Tx-Sum-Time[P][M][W], & Tx-Sample-Count[P][M][W] parameters along with the corresponding PMF ID and window Id

MA may delete the parameters after sending the same to collector.

3.9 Packet Delay Calculation at Collector

Collector accumulates the parameters per PMF per MA per window.
 collector computes the packet delay using the accumulated parameters
 per PMF per MA per window.

3.9.1 Packet Delay Calculation at Collector (option 1)

Packet delay between MA (for example from MA1 to MA2)

$$= (MA2)Rx-Time[P][M][W] - (MA1)Tx-Time[P][M][W]$$

Packet delay at MA = $(MA)Tx-Time[P][M][W] - (MA)Rx-Time[P][M][W]$

3.9.1 Packet Delay Calculation at Collector (option 2)

On receipt of report from a MA collector performs following operation

- 1) For a PMF if it has $Rx-Sum-Time[P][M][W]/Tx-Sum-Time[P][M][W]$ related for an already collected window, received value will be summed up with an existing one, otherwise a new entry is created.
- 2) For a PMF if it has $Rx-Sample-Count[P][M][W]/Tx-Sample-Count[P][M][W]$ related for an already collected window, received value will be summed up with an existing one, otherwise a new entry is created.
- 3) Average time computation by Collector for every MA
 $Average-Rx-Time[P][M][W] = Rx-Sum-Time[P][M][W]/Rx-Sample-Count[P][M][W]$
 $Average-Tx-Time[P][M][W] = Tx-Sum-Time[P][M][W]/Tx-Sample-Count[P][M][W]$
- 4) Packet delay between MA (for example from MA1 to MA2)

$$= (MA2)Average-Rx-Time[P][M][W] - (MA1)Average-Tx-Time[P][M][W]$$

Packet delay at MA

$$= (MA)Average-Tx-Time[P][M][W] - (MA)Average-Rx-Time[P][M][W]$$

4 Security Considerations

No specific security considerations for this document

5 IANA Considerations

5.1 TLV Class

IANA is requested to allocate a new class value for performance measurement from "Network Service Header (NSH) Parameters" registry.

Value	Description	Reference
-----	-----	-----
New	Performance Measurement	[SFC-PM-arch]

6. Acknowledgments

We would like to thank Nobin Mathew for his review and comments.

7 References

7.1 Normative References

- [I-D.ietf-sfc-nsh]
Quinn, P. and U. Elzur, "Network Service Header", [draft-ietf-sfc-nsh-01](#) (work in progress), July 2015.
- [RFC1305] Mills, D., "Network Time Protocol (Version 3) Specification, Implementation and Analysis", [RFC 1305](#), March 1992.
- [RFC5905] Mills, D., Martin, J., Ed., Burbank, J., and W. Kasch, "Network Time Protocol Version 4: Protocol and Algorithms Specification", [RFC 5905](#), June 2010.

7.2 Informative References

- [RFC2784] Farinacci, D., Li, T., Hanks, S., Meyer, D., and P. Traina, "Generic Routing Encapsulation (GRE)", [RFC 2784](#) DOI 10.17487/RFC2784, March 2000, <http://www.rfc-editor.org/info/rfc2784>.
- [RFC5226] Narten, T. and H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", [BCP 26](#), [RFC 5226](#), DOI 10.17487/RFC5226, May 2008, <http://www.rfc-editor.org/info/rfc5226>.
- [RFC7498] Quinn, P., Ed. and T. Nadeau, Ed., "Problem Statement for Service Function Chaining", [RFC 7498](#), DOI 10.17487/RFC7498, April 2015, <http://www.rfc-editor.org/info/rfc7498>.
- [SFC-arch]
Quinn, P., Ed. and J. Halpern, Ed., "Service Function Chaining (SFC) Architecture", 2014, <http://datatracker.ietf.org/doc/draft-quinn-sfc-arch>.
- [SFC-PM-arch]
Anil, Gaurav and Vinod., "Performance Measurement Architecture for SFC", 2015,
<TODO>
- [SFC-PM-Loss]
Anil, Gaurav and Vinod., "Packet Loss Measurement for SFC", 2015,
<TODO>

Authors' Addresses

Anil Kumar S N
Huawei Technologies India Pvt. Ltd,
Near EPIP Industrial Area,
Kundalahalli Village,
Whitefield,
Bangalore - 560037

EMail: anil.sn@huawei.com

Gaurav Agrawal
Huawei Technologies India Pvt. Ltd,
Near EPIP Industrial Area,
Kundalahalli Village,
Whitefield,
Bangalore - 560037

EMail: gaurav.agrawal@huawei.com

Vinod Kumar S
Huawei Technologies India Pvt. Ltd,
Near EPIP Industrial Area,
Kundalahalli Village,
Whitefield,
Bangalore - 560037

EMail: vinods.kumar@huawei.com