Network Working Group Internet-Draft Intended status: Informational Expires: January 7, 2020 T. Mizrahi Huawei Network.IO Innovation Lab C. Arad

> G. Fioccola Huawei Technologies M. Cociglio Telecom Italia M. Chen L. Zheng Huawei Technologies G. Mirsky ZTE Corp. July 6, 2019

Compact Alternate Marking Methods for Passive and Hybrid Performance Monitoring draft-mizrahi-ippm-compact-alternate-marking-05

Abstract

This memo introduces new alternate marking methods that require a compact overhead of either a single bit per packet, or zero bits per packet. This memo also presents a summary of alternate marking methods, and discusses the tradeoffs among them. The target audience of this document is network protocol designers; this document is intended to help protocol designers choose the best alternate marking method(s) based on the protocol's constraints and requirements.

Status of This Memo

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

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 January 7, 2020.

Mizrahi, et al.

Expires January 7, 2020

Copyright Notice

Copyright (c) 2019 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to <u>BCP 78</u> and the IETF Trust's Legal Provisions Relating to IETF Documents (<u>https://trustee.ietf.org/license-info</u>) 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

$\underline{1}$. Introduction	<u>3</u>
<u>1.1</u> . Background	<u>3</u>
<u>1.2</u> . The Scope of This Document	<u>4</u>
<u>2</u> . Terminology	<u>5</u>
<u>2.1</u> . Requirements Language	<u>5</u>
<u>2.2</u> . Abbreviations	<u>5</u>
<u>3</u> . Marking Abstractions	<u>5</u>
$\underline{4}$. Double Marking	7
<u>5</u> . Single-bit Marking	<u>8</u>
<u>5.1</u> . Single Marking Using the First Packet	<u>8</u>
5.2. Single Marking using the Mean Delay	<u>8</u>
<u>5.3</u> . Single Marking using a Multiplexed Marking Bit	<u>8</u>
<u>5.3.1</u> . Overview	<u>8</u>
<u>5.4</u> . Pulse Marking	<u>9</u>
<u>6</u> . Zero Marking Hashed	<u>10</u>
<u>6.1</u> . Hash-based Sampling	<u>10</u>
<u>6.1.1</u> . Hashed Pulse Marking	<u>11</u>
<u>6.1.2</u> . Hashed Step Marking	<u>11</u>
<u>7</u> . Single Marking Hashed	<u>11</u>
8. Timing and Synchronization Aspects	<u>12</u>
<u>8.1</u> . Synchronization Aspects in Multiplexed Marking	<u>13</u>
9. Multipoint Marking Methods	<u>14</u>
<u>10</u> . Summary of Marking Methods	<u>15</u>
<u>11</u> . Alternate Marking using Reserved Values	<u>19</u>
<u>12</u> . IANA Considerations	<u>20</u>
<u>13</u> . Security Considerations	<u>20</u>
<u>14</u> . References	<u>20</u>
<u>14.1</u> . Normative References	<u>20</u>
<u>14.2</u> . Informative References	<u>20</u>
Authors' Addresses	21

Internet-Draft

<u>1</u>. Introduction

<u>1.1</u>. Background

Alternate marking, defined in [RFC8321], is a method for measuring packet loss, packet delay, and packet delay variation. Typical delay measurement protocols require the two measurement points (MPs) to exchange timestamped test packets. In contrast, the alternate marking method does not require control packets to be exchanged. Instead, every data packet carries a marking bit, which is used for triggering measurement events. Note that the frequency of these measurement events is dependent on the users' application(s) and the node characteristics.

The marking bit can be used as a color indication, as defined in [<u>RFC8321</u>], which is toggled periodically. This approach is illustrated in Figure 1.

```
A: packet with color 0
B: packet with color 1
```

Packets	AAAAAAAAA BBBBBBBBB AAAAAAAA BBBBBBBBB AAAAAA
Time	>
	Block 1 Block 2 Block 3 Block 4 Block 5
Color	000000000 111111111 0000000000 11111111

Figure 1: Alternate marking: packets are monitored on a per-color basis.

Alternate marking is used between two MPs, the initiating MP, and the monitoring MP. The initiating MP incorporates the marking field into en-route packets, allowing the monitoring MP to use the marking field in order to bind each packet to the corresponding block.

Each of the MPs maintains two counters, one per color. At the end of each block the counter values can be collected by a central management system, and analyzed; the packet loss can be computed by comparing the counter values of the two MPs.

When using alternate marking delay measurement can be performed in one of three ways (as per [<u>RFC8321</u>]):

o Single marking using the first packet: in this method each packet uses a single marking bit, used as a color indicator. The first packet of each block is used by both MPs as a reference for delay

measurement. The timestamp of this packet is measured by the two measurement points, and can be collected by the mangement system from each of the measurement points, which can compute the path delay by comparing the two timestamps. The drawback of this approach is that it is not accurate when packets arrive out-oforder, as the two MPs may have a different view of which packet was the first in the block.

- o Single marking using the mean delay: as in the previous method, each packet uses a single marking method, indicating the color. Each of the MPs computes the average packet timestamp of each block. The management system can then compute the delay by comparing the average times of the two MPs. The drawback of this approach is that it may be computationally heavy, or difficult to implement at the data plane.
- o Double marking: each packet uses two marking bits. One bit is used as a color indicator, and one is used as a timestamping indicator. This method resolves the drawbacks raised for the two previous methods, at the expense of an extra bit in the packet header.

The double marking method is the most straightforward approach. It allows for accurate measurement without incurring expensive computational load. However, in some cases allocating two bits for passive measurement is not possible. For example, if alternate marking is implemented over IPv4, allocating 2 marking bits in the IPv4 header is challenging, as every bit in the 20-octet header is costly; one of the possible approaches discussed in [<u>RFC8321</u>] is to reserve one or two bits from the DSCP field for remarking. In this case every marking bit comes at the expense of reducing the DSCP range by a factor of two.

<u>1.2</u>. The Scope of This Document

This memo extends the marking methods of [<u>RFC8321</u>], and introduces methods that require a single marking bit, or zero marking bits.

Two single-bit marking methods are proposed, multiplexed marking and pulse marking. In multiplexed marking the color indicator and the timestamp indicator are multiplexed into a single bit, providing the advantages of the double marking method while using a single bit in the packet header. In pulse marking both delay and loss measurement are triggered by a 'pulse' value in a single marking field.

This document also discusses zero-bit marking methods that leverage well-known hash-based selection approaches ([<u>RFC5474</u>], [<u>RFC5475</u>]).

Alternate marking is discussed in this memo as a single-bit or a twobit marking method. However, these methods can similarly be applied to larger fields, such as an IPv6 Flow Label or an MPLS Label; single-bit marking can be applied using two reserved values, and twobit marking can be applied using four reserved values. Marking based on reserved values is further discussed in this document, including its application to MPLS and IPv6.

Finally, this memo summarizes the alternate marking methods, and discusses the tradeoffs among them. It is expected that different network protocols will have different constraints, and therefore may choose to use different alternate marking methods. In some cases it may be preferable to support more than one marking method; in this case the particular marking method may be signaled through the control plane.

2. Terminology

<u>2.1</u>. Requirements Language

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 <u>RFC 2119</u> [<u>RFC2119</u>].

2.2. Abbreviations

The following abbreviations are used in this document:

DSCP	Differentiated Services Code Point
DM	Delay Measurement
LM	Loss Measurement
LSP	Label Switched Path
MP	Measurement Point
MPLS	Multiprotocol Label Switching
SFL	Synonymous Flow Label [<u>I-D.ietf-mpls-sfl-framework</u>]

3. Marking Abstractions

The marking methods that were discussed in <u>Section 1</u>, as well as the methods introduced in this document, use two basic abstractions, pulse detection, and step detection.

The common thread along the various marking methods is that one or two marking bits are used by the MPs to signal a measurement event. The value of the marking bit indicates when the event takes place, in one of two ways:

Pulse An event is detected when the value of the marking bit is toggled in a single packet.

Step An event is detected when the value of the marking bit is toggled, and remains at the new value.

The double marking method (<u>Section 1</u>) uses pulse-based detection for DM, and step-based detection for LM.

Pulse-based detection affects the processing of a single packet; the packet that indicates the pulse is processed differently than the packets around it. For example, in the double marking method, the marked packet is timestamped for DM, without affecting the packets before or after it. Note that if the marked packet is lost, no pulse is detected, yielding a missing measurement (see Figure 2).

P: indicates a packet

Packets	РРРРРРРР	PPPPPPPPP	PPPPPPPPP	PPPPPPPPP	PPPPPPPPP
Time					>
Marking bit	0000010000	0000010000	0000010000	0000010000	00000 0000
	\wedge	\wedge	\wedge	Λ	Λ
Pulse-based	k	I	I	I	
detection		I			
					opped packet: detection

Figure 2: Pulse-based Detection.

In step-based detection the event is detected by observing a value change in stream of packets. Specifically, when the step approach is used for LM (as in the double marking method), two counters are used per flow; each MP decides which counter to use based on the value of the marking bit. Thus, the step-based approach allows accurate counting even when packets arrive out-of-order (see Figure 3). When the step approach is used for DM (e.g., single marking using the first packet), out-of-order causes the delay measurement to be false, without any indication to the management system.

Internet-Draft

P: indicates a packet

Packets Time Λ \wedge Λ Λ Step-based detection out-of-order

Figure 3: Step-based Detection.

4. Double Marking

The two-bit marking method of [RFC8321] uses two marking bits: a color indicator, and a delay measurement indicator. The color bit is used for step-based LM, while the delay bit is used as a pulse-based DM trigger. This double marking approach is the most straightforward of the approaches discussed in this memo, as it allows accurate measurement, it is resilient to out-of-order delivery, and is relatively simple to implement. The main drawback is that it requires two bits, which are not always available.

Figure 4 illustrates the double marking method: each block of packets includes a packet that is marked for timestamping, and therefore has its delay bit set.

```
A: packet with color 0
B: packet with color 1
```

Packets Time -----> | Block 1 | Block 2 | Block 3 | Block 4 | Block 5 ... Λ Λ Λ Λ Λ Packets marked for timestamping 1

Figure 4: The double marking method.

Internet-Draft

5. Single-bit Marking

5.1. Single Marking Using the First Packet

This method uses a single marking bit that indicates the color, as described in [RFC8321]. Both LM and DM are implemented using a stepbased approach; LM is implemented using two color-based counters per flow. The first packet of every period is used by the two MPs as the reference for measuring the delay. As denoted above, the delay computed in this method may be erroneous when packets are delivered out-of-order.

```
A: packet with color 0
B: packet with color 1
```

AAAAAAAAAA				
Block 1	Block 2	Block 3	Block 4	Block 5
	1111111111	0000000000	 	0000000000
^	^	^	^	^
1	1	I	I	1
	Block 1 00000000000	 Block 1 Block 2 0000000000 111111111	 Block 1 Block 2 Block 3 0000000000 111111111 0000000000	AAAAAAAAA BBBBBBBBB AAAAAAAAAA BBBBBBBBB Block 1 Block 2 Block 3 Block 4 000000000 111111111 000000000 111111111 ^ ^ ^ ^

Figure 5: Single marking using the first packet of the block.

<u>5.2</u>. Single Marking using the Mean Delay

As in the first-packet approach, in the mean delay approach ([RFC8321]) a single marking bit is used to indicate the color, enabling step-based loss measurement. Delay is measured in each period by averaging the measured delay over all the packets in the period. As discussed above, this approach is not sensitive to out-of-order delivery, but may be heavy from a computational perspective.

5.3. Single Marking using a Multiplexed Marking Bit

5.3.1. Overview

This section introduces a method that uses a single marking bit that serves two purposes: a color indicator, and a timestamp indicator. The double marking method that was discussed in the previous section uses two 1-bit values: a color indicator C, and a timestamp indicator T. The multiplexed marking bit, denoted by M, is an exclusive or between these two values: M = C XOR T.

An example of the use of the multiplexed marking bit is depicted in Figure 6. The example considers two routers, R1 and R2, that use the multiplexed bit method to measure traffic from R1 to R2. In each block R1 designates one of the packets for delay measurement. In each of these designated packets the value of the multiplexed bit is reversed compared to the other packets in the same block, allowing R2 to distinguish the designated packets from the other packets.

```
A: packet with color 0
B: packet with color 1
```

Packets	АААААААА	BBBBBBBBBB	АААААААА	BBBBBBBBBB	ΑΑΑΑΑΑΑΑ
Time					>
	Block 1	Block 2	Block 3	Block 4	Block 5
Color	0000000000	1111111111	0000000000	1111111111	0000000000
	Λ	Λ	Λ	\wedge	Λ
Packets				I	
marked for				I	
timestampir	ng			I	
	V	V	V	V	V
Muxed bit	0000100000	1111011111	0000100000	1111101111	0001000000

Figure 6: Alternate marking with multiplexed bit.

<u>5.4</u>. Pulse Marking

Pulse marking uses a single marking bit that is used as a trigger for both LM and DM. In this method the two MPs maintain a single perflow counter for LM, in contrast to the color-based methods which require two counters per flow. In each block one of the packets is marked. The marked packet triggers two actions in each of MPs:

o The timestamp is captured for DM.

o The value of the counter is captured for LM.

In each period, each of the MPs exports the timestamp and counterstamp to the management system, which can then compute the loss and delay in that period. It should be noted that as in [RFC8321], if the length of the measurement period is L time units, then all network devices must be synchronized to the same clock reference with an accuracy of +/- L/2 time units.

The pulse marking approach is illustrated in Figure 7. Since both LM and DM use a pulse-based trigger, if the marked packet is lost then no measurement is available in this period. Moreover, the LM accuracy may be affected by out-of-order delivery.

P: packet - all packets have the same color

Packets	РРРРРРРРР	РРРРРРРР	РРРРРРРР	РРРРРРРР	РРРРРРРР
Time					>
			1		
	Block 1	Block 2	Block 3	Block 4	Block 5
			1	1	
	Λ	Λ	\wedge	\wedge	\wedge
Packets	I		I		
marked for	I		I	I	
DM and LM	I		I	I	
	V	V	V	V	V

Marking bit 0000100000 0000100000 0000100000 0000010000 000100000

Figure 7: Pulse marking method.

<u>6</u>. Zero Marking Hashed

6.1. Hash-based Sampling

Hash based selection [<u>RFC5475</u>] is a well-known method for sampling a subset of packets. As defined in [<u>RFC5475</u>]:

A Hash Function h maps the Packet Content c, or some portion of it, onto a Hash Range R. The packet is selected if h(c) is an element of S, which is a subset of R called the Hash Selection Range.

Hash-based selection can be leveraged as a marking method, allowing a zero-bit marking approach. Specifically, the pulse and step abstractions can be implemented using hashed selection:

- o Hashed pulse-based trigger: in this approach, a packet is selected if h(c) is an element of S, which is a strict subset of the hash range R. When |S|<<|R|, the average sampling period is long, reducing the probability of ambiguity between consecutive packets. |S| and |R| denote the number of elements in S and R, respectively.
- o Hashed step-based trigger: the hash values of a given traffic flow are said to be monotonically increasing if for two packets p1 and

p2, if p1 is sent before p2 then h(p1)<=h(p2). If it is guaranteed that the hash values of a flow are monotonically increasing, then a step-based approach can be used on the range R. For example, in an IPv4 flow the Identification field can be used as the hash value of each packet. Since the Identification field is monotonically increasing, the step-based trigger can be implemented using consecutive ranges of the Identification value. For example, the fourth bit of the Identification field is toggled every 8 packets. Thus, a possible hash function simply takes the fourth bit of the Identification field as the hash value. This hash value is toggled every 8 packets, simulating the alternate marking behavior of Section 4.

Note that as opposed to the double marking and single marking methods, hashed sampling is not based on fixed time intervals, as the duration between sampled packets depends only on the hash value.

It is also important to note that all methods that use hash-based marking require the hash function and the set S to be configured consistently across the MPs.

6.1.1. Hashed Pulse Marking

In this approach a hash is computed over the packet content, and both LM and DM are triggered based on the pulse-based trigger (Section 6.1). A pulse is detected when the hash value h(c) is equal to one of the values in S. The hash function h and the set S determine the probability (or frequency) of the pulse event.

6.1.2. Hashed Step Marking

As in the previous approach, hashed step marking also uses a hash that is computed over the packet content. In this approach DM is performed using a pulse-based trigger, whereas the LM trigger is step-based (<u>Section 6.1</u>). The main drawback of this method is that the step-based trigger is possible only under the assumption that the hash function is monotonically increasing, which is not necessarily possible in all cases. Specifically, a measured flow is not necessarily an IPv4 5-tuple. For example, a measured flow may include multiple IPv4 5-tuple flows, and in this case the Identification field is not monotonically increasing.

7. Single Marking Hashed

Mixed hashed marking combines the single marking approach with hashbased sampling. A single marking bit is used in the packet header as a color indicator, while a hash-based pulse is used to trigger DM. Although this method requires a single bit, it is described in this

section as it is closely related to the other hash-based methods that require zero marking bits.

The hash-based selection for DM can be applied in one of two possible approaches: the basic approach, and the dynamic approach. In the basic approach, packets forwarded between two MPs, MP1 and MP2, are selected using a hash function, as described above. One of the challenges is that the frequency of the sampled packets may vary considerably, making it difficult for the management system to correlate samples from the two MPs. Thus, the dynamic approach can be used.

In the dynamic hash-based sampling, alternate marking is used to create divide time into periods, so that hash-based samples are divided into batches, allowing to anchor the selected samples to their period. Moreover, by dynamically adapting the length of the hash value, the number of samples is bounded in each marking period. This can be realized by choosing first the maximum number of samples (NMAX) to be used with the initial hash length. The algorithm starts with only few hash bits, that permit to select a greater percentage of packets (e.g. with 1 bit of hash half of the packets are sampled). When the number of selected packets reaches NMAX, a hashing bit is added. As a consequence, the sampling proceeds at half of the original rate and the packets already selected that do not match the new hash are discarded. This step can be repeated iteratively. It is assumed that each sample includes the timestamp (used for DM) and the hash value, allowing the management system to match the samples received from the two MPs.

The dynamic process statistically converges at the end of a marking period and the number of selected samples beyond the initial NMAX samples mentioned above is between NMAX/2 and NMAX. Therefore, the dynamic approach paces the sampling rate, allowing to bound the number of sampled packets per sampling period.

8. Timing and Synchronization Aspects

As pointed out in [RFC8321], it is assumed that all MPs are synchronized to a common reference time with an accuracy of +/- L/2, where L is the periodic measurement interval. Thus, the difference between the clock values of any two MPs is bounded by L. Note that this is a relatively relaxed synchronization requirement that does not require complex means of synchronization. Clocks can be synchronized for example using NTP [RFC5905], PTP [IEEE1588], or by other means.

In the step-based approaches the common reference time is used for dividing the time domain into equal-sized measurement periods, such

that all packets forwarded during a measurement period have the same color, and consecutive periods have alternating colors. In the pulse-based approaches the synchronization helps the management system to correlate measurements from multiple measurement points without ambiguity.

8.1. Synchronization Aspects in Multiplexed Marking

The single marking bit incorporates two multiplexed values. From the monitoring MP's perspective, the two values are Time-Division Multiplexed (TDM), as depicted in Figure 8. It is assumed that the start time of every measurement period is known to both the initiating MP and the monitoring MP. If the measurement period is L, then during the first and the last L/4 time units of each block the marking bit is interpreted by the monitoring MP as a color indicator. During the middle part of the block, the marking bit is interpreted as a timestamp indicator; if the value of this bit is different than the color value, the corresponding packet is used as a reference for delay measurement.

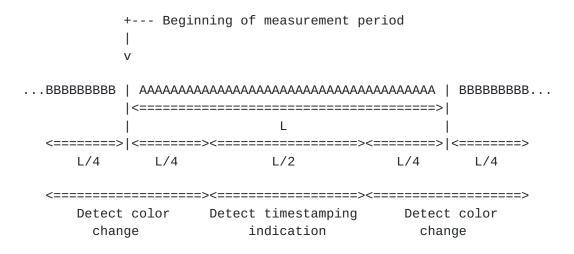


Figure 8: Multiplexed marking field interpretation at the receiving measurement point.

In order to prevent ambiguity in the receiver's interpretation of the marking field, the initiating MP is permitted to set the timestamp indication only during a specific interval, as depicted in Figure 9. Since the receiver is willing to receive the timestamp indication during the middle L/2 time units of the block, the sender refrains from sending the timestamp indication during a guardband interval of time units at the beginning and end of the L/2-period.

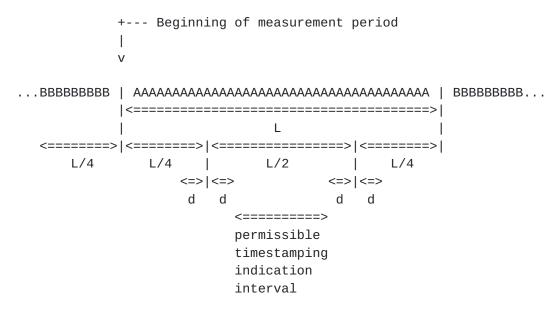


Figure 9: A time domain view.

The guardband d is given by d = A + D_max - D_min, where A is the clock accuracy, D_max is an upper bound on the network delay between the MPs, and D_min is a lower bound on the delay. It is straightforward from Figure 9 that d < L/4 must be satisfied. The latter implies a minimal requirement on the synchronization accuracy.

All MPs must be synchronized to the same reference time with an accuracy of +/- L/8. Depending on the system topology, in some systems the accuracy requirement will be even more stringent, subject to d < L/4. Note that the accuracy requirement of the conventional alternate marking method [RFC8321] is +/- L/2, while the multiplexed marking method requires an accuracy of +/- L/8.

Note that we assume that the middle L/2-period is designated as the timestamp indication period, allowing a sufficiently long guardband between the transitions. However, a system may be configured to use a longer timestamp indication period or a shorter one, if it is guaranteed that the synchronization accuracy meets the guardband requirements (i.e., the constraints on d).

9. Multipoint Marking Methods

It should be noted that most of the marking methods that were presented in this memo are intended for point-to-point measurements, e.g., from MP1 to MP2 in Figure 10. In point-to-multipoint measurements, the mean delay method can be used to measure the loss and delay of the entire point-to-multipoint flow (which includes all the traffic from MP3 to either MP4 or MP5), while other methods such as double marking can be used to measure the point-to-point

performance, for example from MP3 to MP5. Alternate marking in multipoint scenarios is discussed in detail in [I-D.ietf-ippm-multipoint-alt-mark].

MP4 MP1 MP3 MP2 +--+ +--+ +--+ +--+ + - - + | |---->| | | |---->| |---->| | +--+ +--+ +--+ +--+ +--+ MP5 +--+ +---->| | +--+

Point-to-point measurement Point-to-multipoint measurement

Figure 10: Point-to-point and point-to-multipoint measurements.

10. Summary of Marking Methods

This section summarizes the marking methods described in this memo. Each row in the table of Figure 11 represents a marking method. For each method the table specifies the number of bits required in the header, the number of counters per flow for LM, the methods used for LM and DM (pulse or step), and also the resilience to disturbances.

Mizrahi, et al. Expires January 7, 2020 [Page 15]

+----+ | Method |# of|# of|LM |DM |Resilience to|Resilience to| |bits|coun|Method|Method|Reordering |Packet drops | | | | | LM | DM | LM | DM | |Single marking| 1 | 2 |Step | Step | + | -- | + | -- | |- 1st packet | | | | | | | | +----+ |Single marking| 1 | 2 |Step |Mean | + | + | + | - | |- mean delay | | | | | |Double marking| 2 | 2 |Step |Pulse | + | + | = |Single marking| 1 | 2 |Step |Pulse | + | + | + | = | 1 |multiplexed | | | | | | | | |Pulse marking | 1 | 1 |Pulse |Pulse | -- | + | - | = | |Zero marking | 0 | 1 |Hashed|Hashed| -- | + | - | + |hashed | |(2) |pulse |pulse | (-) | | | |(step)| | | 1 |Single marking| 1 | 2 |Step |Hashed| + | + | + | + | |hashed | | | |pulse | | |

- + Accurate measurement.
- = Invalidate only if a measured packet is lost (detectable)
- No measurement in case of disturbance (detectable).
- -- False measurement in case of disturbance (not detectable).

Figure 11: Detailed Summary of Marking Methods

In the context of this comparison two possible disturbances are considered: out-of-order delivery, and packet drops. Generally speaking, pulse based methods are sensitive to packet drops, since if the marked packet is dropped no measurement is recorded in the current period. Notably, a missing measurement is detectable by the management system, and is not as severe as a false measurement. Step-based triggers are generally resilient to out-of-order delivery for LM, but are not resilient to out-of-order delivery for DM. Notably, a step-based trigger may yield a false delay measurement when packets are delivered out-of-order, and this inaccuracy is not detectable.

As mentioned above, the double marking method is the most straightforward approach, and is resilient to most of the

disturbances that were analyzed. Its obvious drawback is that it requires two marking bits.

Several single marking methods are discussed in this memo. In this case there is no clear verdict which method is the optimal one. The first packet method may be simple to implement, but may present erroneous delay measurements in case of dropped or reordered packets. Arguably, the mean delay approach and the multiplexed approach may be more difficult to implement (depending on the underlying platform), but are more resilient to the disturbances that were considered here. Note that the computational complexity of the mean delay approach can be reduced by combining it with a hashed approach, i.e., by computing the mean delay over a hash-based subset of the packets. The pulse marking method requires only a single counter per flow, while the other methods require two counters per flow.

The hash-based sampling approaches reduce the overhead to zero bits, which is a significant advantage. However, the sampling period in these approaches is not associated with a fixed time interval. Therefore, in some cases adjacent packets may be selected for the sampling, potentially causing measurement errors. Furthermore, when the traffic rate is low, measurements may become signifcantly infrequent.

It is clear from the previous table that packet loss measurement can be considered resilient to both reordering and packet drops if at least one bit is used with a step-based approach. Thus, since the packet loss can be considered obvious, the previous table can be simplified into Figure 12, where only the characteristics of delay measurements are highlighted. This more compact table allows room for an additional column referring to multipoint-to-multipoint (Section 9) delay measurement compatibility.

Mizrahi, et al. Expires January 7, 2020 [Page 17]

Internet-Draft

+	+	+	+	+	++
		on All	Resilience to	Resilience	compatible
Single marking - 1st packet		Yes 		- - 	NO
Single marking - mean delay	1 +	Yes 	+ +	- - 	Yes
Double marking	2 +	Yes	+ +	=	No ++
Single marking multiplexed	1	Yes	+	=	No
Pulse marking	1	NO	+	=	NO
Zero marking hashed 	0 	No 	+ 	+ 	Yes
Single marking hashed +	1 +	Yes 	 + 	 + +	Yes

- + Accurate measurement.
- = Invalidate only if a measured packet is lost (detectable)
- No measurement in case of disturbance (detectable).
- -- False measurement in case of disturbance (not detectable).

Figure 12: Summary of Marking Methods: focus on Delay Measurement

In the context of delay measurement, both zero marking hashed and single marking hashed are resilient to packet drops. Using double marking it could also be possible to perform an accurate measurement in case of packet drops, as long as the packet that is marked for DM is not dropped.

The single marking hashed method seems the most complete approach, especially because it is also compatible with multipoint-tomultipoint measurements.

<u>11</u>. Alternate Marking using Reserved Values

As mentioned in <u>Section 1</u>, a marking bit is not necessarily a single bit, but may be implemented by using two well-known values in one of the header fields. Similarly, two-bit marking can be implemented using four reserved values.

A notable example is MPLS Synonymous Flow Labels (SFL), as defined in [I-D.ietf-mpls-rfc6374-sfl]. Two MPLS Label values can be used to indicate the two colors of a given LSP: the original Label value, and an SFL value. A similar approach can be applied to IPv6 using the Flow Label field.

The following example illustrates how alternate marking can be implemented using reserved values. The bit multiplexing approach of Section 5.3 is applicable not only to single-bit color indicators, but also to two-value indicators; instead of using a single bit that is toggled between '0' and '1', two values of the indicator field, U and W, can be used in the same manner, allowing both loss and delay measurement to be performed using only two reserved values. Thus, the multiplexing approach of Figure 6 can be illustrated more generally with two values, U and W, as depicted in Figure 13.

```
B: packet with color 1
Packets
    ----->
Time
    1
           1
        1
               | Block 1 | Block 2 | Block 3 | Block 4 | Block 5 ...
```

Color	00000000000000	11111111111	0000000000	11111111111	00000000000
	Λ	Λ	Λ	Λ	Λ
Packets	I	I	I	I	
marked for					
timestampi	ng			I	I
	V	V	V	V	V
Muxed marking	UUUUWUUUUU	WWWWUWWWWW	UUUUWUUUUU	WWWWWUWWWW	UUUWUUUUUU

values

A: packet with color 0

Figure 13: Alternate marking with two multiplexed marking values, U and W.

<u>12</u>. IANA Considerations

This memo includes no requests from IANA.

13. Security Considerations

The security considerations of the alternate marking method are discussed in [RFC8321]. The analysis of Section 10 emphasizes the sensitivity of some of the alternate marking methods to packet drops and to packet reordering. Thus, a malicious attacker may attempt to tamper with the measurements by either selectively dropping packets, or by selectively reordering specific packets. The multiplexed marking method Section 5.3 that is defined in this document requires slightly more stringent synchronization than the conventional marking method, potentially making the method more vulnerable to attacks on the time synchronization protocol. A detailed discussion about the threats against time protocols and how to mitigate them is presented in [RFC7384].

<u>14</u>. References

<u>14.1</u>. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", <u>BCP 14</u>, <u>RFC 2119</u>, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/info/rfc2119>.
- [RFC8321] Fioccola, G., Ed., Capello, A., Cociglio, M., Castaldelli, L., Chen, M., Zheng, L., Mirsky, G., and T. Mizrahi, "Alternate-Marking Method for Passive and Hybrid Performance Monitoring", <u>RFC 8321</u>, DOI 10.17487/RFC8321, January 2018, <<u>https://www.rfc-editor.org/info/rfc8321</u>>.

<u>14.2</u>. Informative References

```
[I-D.ietf-ippm-multipoint-alt-mark]
```

Fioccola, G., Cociglio, M., Sapio, A., and R. Sisto, "Multipoint Alternate Marking method for passive and hybrid performance monitoring", <u>draft-ietf-ippm-</u> <u>multipoint-alt-mark-02</u> (work in progress), July 2019.

[I-D.ietf-mpls-rfc6374-sfl]

Bryant, S., Chen, M., Li, Z., Swallow, G., Sivabalan, S., Mirsky, G., and G. Fioccola, "<u>RFC6374</u> Synonymous Flow Labels", <u>draft-ietf-mpls-rfc6374-sfl-03</u> (work in progress), December 2018.

[I-D.ietf-mpls-sfl-framework]

Bryant, S., Chen, M., Li, Z., Swallow, G., Sivabalan, S., and G. Mirsky, "Synonymous Flow Label Framework", <u>draft-</u> <u>ietf-mpls-sfl-framework-04</u> (work in progress), December 2018.

[IEEE1588]

IEEE, "IEEE 1588 Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems Version 2", 2008.

- [RFC5474] Duffield, N., Ed., Chiou, D., Claise, B., Greenberg, A., Grossglauser, M., and J. Rexford, "A Framework for Packet Selection and Reporting", <u>RFC 5474</u>, DOI 10.17487/RFC5474, March 2009, <<u>https://www.rfc-editor.org/info/rfc5474</u>>.
- [RFC5475] Zseby, T., Molina, M., Duffield, N., Niccolini, S., and F. Raspall, "Sampling and Filtering Techniques for IP Packet Selection", <u>RFC 5475</u>, DOI 10.17487/RFC5475, March 2009, <<u>https://www.rfc-editor.org/info/rfc5475</u>>.
- [RFC5905] Mills, D., Martin, J., Ed., Burbank, J., and W. Kasch, "Network Time Protocol Version 4: Protocol and Algorithms Specification", <u>RFC 5905</u>, DOI 10.17487/RFC5905, June 2010, <https://www.rfc-editor.org/info/rfc5905>.
- [RFC7384] Mizrahi, T., "Security Requirements of Time Protocols in Packet Switched Networks", <u>RFC 7384</u>, DOI 10.17487/RFC7384, October 2014, <<u>https://www.rfc-editor.org/info/rfc7384</u>>.

Authors' Addresses

Tal Mizrahi Huawei Network.IO Innovation Lab Israel

Email: tal.mizrahi.phd@gmail.com

Carmi Arad

Email: carmi.arad@gmail.com

Giuseppe Fioccola Huawei Technologies

Email: giuseppe.fioccola@huawei.com

Mauro Cociglio Telecom Italia Via Reiss Romoli, 274 Torino 10148 Italy

Email: mauro.cociglio@telecomitalia.it

Mach(Guoyi) Chen Huawei Technologies

Email: mach.chen@huawei.com

Lianshu Zheng Huawei Technologies

Email: vero.zheng@huawei.com

Greg Mirsky ZTE Corp.

Email: gregimirsky@gmail.com

Mizrahi, et al. Expires January 7, 2020 [Page 22]