Network Working Group
Internet-Draft
Intended status: Informational
Expires: May 2, 2017

T. Mizrahi
Marvell
G. Fioccola
Telecom Italia
M. Chen
L. Zheng
Huawei Technologies
G. Mirsky

October 29, 2016

Passive Performance Monitoring using a Multiplexed Marking Field
draft-mizrahi-ippm-multiplexed-alternate-marking-00

Abstract

This memo introduces a marking method that uses a single marking bit and allows accurate loss and delay measurement.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

Alternate marking, defined in [I-D.ietf-ippm-alt-mark], 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 text packets. In contrast, the alternate marking method does not require control packets to be exchanged. Instead, every data packet carries a color indicator, which divides the traffic into consecutive blocks of packets. The color value is toggled periodically, as illustrated in Figure 1.

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

Packets      AAAAAAAAAA BBBBBBBBBB AAAAAAAAAA BBBBBBBBBB AAAAAAAAAA
   Time   ---------------------------------------------------------->
            |          |          |          |          |
            | Block 1  | Block 2  | Block 3  | Block 4  | Block 5 ...
            |          |          |          |          |
Color        0000000000 1111111111 0000000000 1111111111 0000000000
```

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 [I-D.ietf-ippm-alt-mark]):

o Single marking: 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-of-order, as the two measurement may have a different view of which packet was the first in the block.

o Average delay: 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 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 [I-D.ietf-ippm-alt-mark] is 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.

This memo extends the marking method of [I-D.ietf-ippm-alt-mark]. The method introduced in this document uses a single marking bit in the packet header, while providing the advantages of the double marking method. In a nutshell, the color indicator and the timestamp indicator are multiplexed into a single bit. There is an underlying assumption that the two MPs that take part in the measurement are time-synchronized.

2. Terminology

2.1. 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 RFC 2119 [RFC2119].

2.2. Abbreviations

MP Measurement Point

DSCP Differentiated Services Code Point

3. Alternate Marking using a Multiplexed Marking Bit

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 2. 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      AAAAAAAAAA BBBBBBBBBB AAAAAAAAAA BBBBBBBBBB AAAAAAAAAA
      Time   ----------------------------------------------------------->
               |          |          |          |          |
               | Block 1  | Block 2  | Block 3  | Block 4  | Block 5 ...
               |          |          |          |          |
   Color        0000000000 1111111111 0000000000 1111111111 0000000000
                    ^          ^          ^           ^        ^
     Packets        |          |          |           |        |
     marked for     |          |          |           |        |
     timestamping   |          |          |           |        |
                    v          v          v           v        v
   Muxed bit    0000100000 1111011111 0000100000 1111101111 0001000000
```

Figure 2: Alternate marking with multiplexed bit.

3.2. Timing and Synchronization Aspects

It is assumed that all MPs are synchronized to a common reference time with an accuracy of +/- A/2. Thus, the difference between the clock values of any two MPs is bounded by A. Clocks can be synchronized for example using NTP [RFC5905], PTP [IEEE1588], or by other means. 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.

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 3. 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.

```
                 +--- Beginning of measurement period
                 |
                 v

    ...BBBBBBBBB | AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA | BBBBBBBB...
                 |<======================================>|
                 |                   L                    |
       <========>|<========><==================><========>|<========>
           L/4       L/4            L/2             L/4       L/4

       <===================><==================><===================>
           Detect color     Detect timestamping      Detect color
             change              indication           change
```

Figure 3: 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 4. 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 d time units at the beginning and end of the L/2-period.

```
                 +--- Beginning of measurement period
                 |
                 v

    ...BBBBBBBBB | AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA | BBBBBBBB...
                 |<======================================>|
                 |                   L                    |
       <========>|<========>|<================>|<========>|
           L/4       L/4    |      L/2       |    L/4
                         <=>|<=>            <=>|<=>
                          d   d              d   d
                                <===========>
                                permissible
                                timestamping
                                indication
                                interval
```

Figure 4: 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 4 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 [I-D.ietf-ippm-alt-mark] 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).

4. IANA Considerations

This memo includes no requests from IANA.

5. Security Considerations

The security considerations of the alternate marking method are discussed in [I-D.ietf-ippm-alt-mark]. Specifically, the method 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].

6. References

6.1. Normative References

[I-D.ietf-ippm-alt-mark]
Fioccola, G., Capello, A., Cociglio, M., Castaldelli, L., Chen, M., Zheng, L., Mirsky, G., and T. Mizrahi, "Alternate Marking method for passive performance monitoring", draft-ietf-ippm-alt-mark-02 (work in progress), October 2016.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <http://www.rfc-editor.org/info/rfc2119>.

6.2. Informative References

[IEEE1588]
IEEE, "IEEE 1588 Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems Version 2", 2008.

[RFC5905] Mills, D., Martin, J., Ed., Burbank, J., and W. Kasch, "Network Time Protocol Version 4: Protocol and Algorithms Specification", RFC 5905, DOI 10.17487/RFC5905, June 2010, <http://www.rfc-editor.org/info/rfc5905>.

[RFC7384] Mizrahi, T., "Security Requirements of Time Protocols in Packet Switched Networks", RFC 7384, DOI 10.17487/RFC7384, October 2014, <http://www.rfc-editor.org/info/rfc7384>.

Authors' Addresses

Tal Mizrahi
Marvell
6 Hamada st.
Yokneam
Israel

Email: talmi@marvell.com

Giuseppe Fioccola
Telecom Italia
Via Reiss Romoli, 274
Torino 10148
Italy

Email: giuseppe.fioccola@telecomitalia.it

Mach(Guoyi) Chen
Huawei Technologies

Email: mach.chen@huawei.com

Lianshu Zheng
Huawei Technologies

Email: vero.zheng@huawei.com

Greg Mirsky
USA

Email: gregimirsky@gmail.com