BIER Working Group
Internet-Draft
Intended status: Standards Track
Expires: 2 October 2022

G. Mirsky
Ericsson
L. Zheng
Individual Contributor
M. Chen
G. Fioccola
Huawei Technologies

31 March 2022

Performance Measurement (PM) with Marking Method in Bit Index Explicit Replication (BIER) Layer

draft-ietf-bier-pmmm-oam-12

Abstract

This document describes the applicability of a hybrid performance measurement method for packet loss and packet delay measurements of a multicast service through a Bit Index Explicit Replication domain.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

[RFC8279] introduces and explains the Bit Index Explicit Replication (BIER) architecture and how it supports the forwarding of multicast data packets. [RFC8296] specified that in the case of BIER encapsulation in an MPLS network, a BIER-MPLS label, the label that is at the bottom of the label stack, uniquely identifies the multicast flow. [I-D.fioccola-rfc8321bis] and [I-D.fioccola-rfc8889bis] describe a hybrid performance measurement method, according to the classification of measurement methods in [RFC7799]. The method, called Packet Network Performance Monitoring (PNPM), can be used to measure packet loss, latency, and jitter on live traffic complies with requirements R-5 and R-12 listed in [I-D.ietf-bier-oam-requirements]. Because this method is based on marking consecutive batches of packets, the method is often referred to as a marking method. Terms PNPM and "marking method" in this document are used interchangeably.

This document defines how the marking method can be used on the BIER layer to measure packet loss and delay metrics of a multicast flow in an MPLS network.

2. Conventions used in this document

2.1. Terminology

This document uses the terms related to the Alternate Marking Method as defined in [I-D.fioccola-rfc8321bis], [I-D.fioccola-rfc8889bis]. This document uses the terms related to the Bit Indexed Explicit Replication as defined in [RFC8296].

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. OAM Field in BIER Header

[RFC8296] defined the two-bits long field, referred to as OAM. The OAM field can be used for the marking performance measurement method. Because the setting of the field to any value does not affect forwarding and/or quality of service treatment of a packet, using the OAM field for PNPM in BIER layer can be viewed as the example of the hybrid performance measurement method.

Figure 1 displays the interpretation of the OAM field defined in this specification for the use of the PNPM method. The context of interpretation of the OAM field MAY be signaled via the control plane or configured using an extension to the BIER YANG data model [I-D.ietf-bier-bier-yang]. These extensions are outside the scope of this document.

```
  0
  0   1
+-+-+-+-+
| S | D |
+-+-+-+-+
```

Figure 1: OAM field of BIER Header format

where:

* S - Single-Marking flag;

* D - Double-Marking flag.

4. Theory of Operation

The marking method can be used in the multicast environment supported by BIER layer. Without limiting any generality consider multicast network presented in Figure 2. Any combination of markings can be applied to a multicast flow by the Bit Forwarding Ingress Router (BFIR) at either ingress or egress point to perform node, link, segment or end-to-end measurement to detect performance degradation defect and localize it efficiently.

```
                        -----
                      --| D |
              -----  /  -----
            --| B |--
           /  -----  \  -----
          /           --| E |
-----    /              -----
| A |---                -----
-----    \            --| F |
          \  -----   /  -----
           --| C |--
             -----   \  -----
                      --| G |
                        -----
```

Figure 2: Multicast network

Using the marking method, a BFIR creates distinct sub-flows in the particular multicast traffic over BIER layer. Each sub-flow consists of consecutive blocks of identically marked packets. For example, a block of N packets, with each packet being marked as X, is followed by the block of M packets with each packet being marked as Y. These blocks are unambiguously recognizable by a monitoring point at any Bit Forwarding Router (BFR) and can be measured to calculate packet loss and/or packet delay metrics. The marking method can be used on multiple flows concurently. Demultiplexing of monitored flows might be achived using n-tuple, for example, two-tuple as combination of the values in the Entropy and BFIR-id fields [RFC8296]. Also, that can be achieved by using an explicit Flow Identifiier. The definition of the Flow Identifier is outside the scope of this specification. It is expected that the marking values be set and cleared at the edge of BIER domain. Thus for the scenario presented in Figure 2 if the operator initially monitors the A-C-G and A-B-D segments he may enable measurements on segments C-F and B-E at any time.

4.1. Single-Marking Enabled Measurement

As explained in [I-D.fioccola-rfc8321bis], marking can be applied to delineate blocks of packets based either on the equal number of packets in a block or based on the equal time interval. The latter method offers better control as it allows a better account for capabilities of downstream nodes to report statistics related to batches of packets and, at the same time, time resolution that affects defect detection interval.

If the Single-Marking measurement is used to measure packet loss, then the D flag MUST be set to zero on transmit and ignored by the monitoring point.

The S flag is used to create sub-flows to measure the packet loss by switching the value of the S flag every N-th packet or at certain time intervals. Delay metrics MAY be calculated with the sub-flow using any of the following methods:

* First/Last Packet Delay calculation: whenever the marking, i.e., the value of S flag changes, a BFR can store the timestamp of the first/last packet of the block. The timestamp can be compared with the timestamp of the packet that arrived in the same order through a monitoring point at a downstream BFR to compute packet delay. Because timestamps collected based on the order of arrival this method is sensitive to packet loss and re-ordering of packets (see Section 4.3 for more details).

* Average Packet Delay calculation: an average delay is calculated by considering the average arrival time of the packets within a single block. A BFR may collect timestamps for each packet received within a single block. Average of the timestamp is the sum of all the timestamps divided by the total number of packets received. Then the difference between the average packet arrival time calculated for the downstream monitoring point and the same metric but calculated at the upstream monitoring point is the average packet delay on the segment between these two points. This method is robust to out of order packets and also to packet loss on the segment between the measurement points (packet loss may cause a minor loss of accuracy in the calculated metric because the number of packets used is different at each measurement point). This method only provides a single metric for the duration of the block, and it doesn't give the minimum and

maximum delay values. This limitation of producing only the single metric could be overcome by reducing the duration of the block. As a result, the calculated value of the average delay will better reflect the minimum and maximum delay values of the block's duration time.

4.2. Double-Marking Enabled Measurement

Double-Marking method allows measurement of minimum and maximum delays for the monitored flow, but it requires more nodal and network resources. If the Double-Marking method used, then the S flag is used to create the sub-flow, i.e., mark blocks of packets. The D flag is used to mark single packets within a block to measure delay and jitter.

The first marking (S flag alternation) is needed for packet loss and also for average delay measurement. The second marking (D flag is put to one) creates a new set of marked packets that are fully identified over the BIER network, so that a BFR can store the timestamps of these packets; these timestamps can be compared with the timestamps of the same packets on a second BFR to compute packet delay values for each packet. The number of measurements can be easily increased by changing the frequency of the second marking. On the other hand, the higher frequency of the second marking will cause a higher volume of the measurement data being transported through the BIER domain. An operator should consider and balance both effects. This method is useful to measure not only the average delay but also the minimum and maximum delay values and, in wider terms, to know more about the statistic distribution of delay values.

4.3. Operational Considerations

For the ease of operational procedures, the initial marking of a multicast flow is performed at BFIR. and cleared, by way of removing BIER encapsulation form a payload packet, at the edge of the BIER domain by BFERs.

Since at the time of writing this specification, there are no proposals to using auto-discovery or signaling mechanism to inform downstream nodes what methodology is used each monitoring point MUST be configured beforehand.

Section 5 [I-D.fioccola-rfc8321bis] provides a detailed analysis of how packet re-ordering and the duration of the block in the Single-Marking mode of the marking method impact the accuracy of the packet loss measurement. Re-ordering of packets in the Single-Marking mode will be noticeable only at the edge of a block of packets (re-ordering within the block cannot be detected in the Single-Marking mode). If the extra delay for some packets is much smaller than half of the duration of a block, then it should be easier to attribute re-ordered packets to the proper block and thus maintain the accuracy of the packet loss measurement.

Selection of a time interval to switch the marking of a batch of packets should be based on the service requirements. In the course of the regular operation, reports, including performance metrics like packet loss ratio, packet delay, and inter-packet delay variation, are logged every 15 minutes. Thus, it is reasonable to maintain the duration of the measurement interval at 5 minutes with 100 measurements per each interval. To support these measurements, marking of the packet batch is switched every 3 seconds. In case when performance metrics are required in near-real-time, the duration interval of a single batch of identically marked packets will be in the range of tens of milliseconds.

5. IANA Considerations

This document sets no requirements to IANA. This section can be removed before the publication.

6. Security Considerations

Regarding using the marking method, [I-D.fioccola-rfc8321bis] stressed two types of security concerns. First, the potential harm caused by the measurements, is a lesser threat as [RFC8296] defines OAM field used by the marking method so that the value of "two bits have no effect on the path taken by a BIER packet and have no effect on the quality of service applied to a BIER packet." Second security

concern, potential harm to the measurements can be mitigated by using policy, suggested in [RFC8296], to accept BIER packets only from trusted routers, not from customer-facing interfaces.

All the security considerations for BIER discussed in [RFC8296] are inherited by this document.

7. Acknowledgement

Comments from Alvaro Retana helped improve the document and are much appreciated.

Reviews and comments from Quan Xiong and Xiao Min are thankfully acknowledged.

8. References

8.1. Normative References

[I-D.fioccola-rfc8321bis]
Fioccola, G., Cociglio, M., Mirsky, G., Mizrahi, T., Zhou, T., and X. Min, "Alternate-Marking Method", Work in Progress, Internet-Draft, draft-fioccola-rfc8321bis-03, 23 February 2022, <https://datatracker.ietf.org/doc/html/draft-fioccola-rfc8321bis-03>.

[I-D.fioccola-rfc8889bis]
Fioccola, G., Cociglio, M., Sapio, A., Sisto, R., and T. Zhou, "Multipoint Alternate-Marking Method", Work in Progress, Internet-Draft, draft-fioccola-rfc8889bis-03, 23 February 2022, <https://datatracker.ietf.org/doc/html/draft-fioccola-rfc8889bis-03>.

[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>.

[RFC8296] Wijnands, IJ., Ed., Rosen, E., Ed., Dolganow, A., Tantsura, J., Aldrin, S., and I. Meilik, "Encapsulation for Bit Index Explicit Replication (BIER) in MPLS and Non-MPLS Networks", RFC 8296, DOI 10.17487/RFC8296, January 2018, <https://www.rfc-editor.org/info/rfc8296>.

8.2. Informative References

[I-D.ietf-bier-bier-yang]
Chen, R., Hu, F., Zhang, Z., Dai, X., and M. Sivakumar, "YANG Data Model for BIER Protocol", Work in Progress, Internet-Draft, draft-ietf-bier-bier-yang-07, 8 September 2020, <https://datatracker.ietf.org/doc/html/draft-ietf-bier-bier-yang-07>.

[I-D.ietf-bier-oam-requirements]
Mirsky, G., Kumar, N., Chen, M., and S. Pallagatti, "Operations, Administration and Maintenance (OAM) Requirements for Bit Index Explicit Replication (BIER) Layer", Work in Progress, Internet-Draft, draft-ietf-bier-oam-requirements-11, 15 November 2020, <https://datatracker.ietf.org/doc/html/draft-ietf-bier-oam-requirements-11>.

[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>.

[RFC8279] Wijnands, IJ., Ed., Rosen, E., Ed., Dolganow, A., Przygienda, T., and S. Aldrin, "Multicast Using Bit Index Explicit Replication (BIER)", RFC 8279, DOI 10.17487/RFC8279, November 2017, <https://www.rfc-editor.org/info/rfc8279>.

Authors' Addresses

Greg Mirsky
Ericsson
Email: gregimirsky@gmail.com

Lianshu Zheng
Individual Contributor
Email: veronique_zheng@hotmail.com

Mach Chen
Huawei Technologies
Email: mach.chen@huawei.com

Giuseppe Fioccola

Huawei Technologies
Email: giuseppe.fioccola@huawei.com