IPPM Working Group
Internet-Draft
Intended status: Standards Track
Expires: July 9, 2024

H. Wang
Y. Liu
China Mobile
C. Lin
New H3C Technologies
X. Min
ZTE Corporation
G. Mirsky
Ericsson

January 9, 2024

Flow Measurement in IPv6 Network
draft-wang-ippm-ipv6-flow-measurement-06

Status of this Memo

Copyright Notice

Abstract

This document describes how to deploy in-situ flow performance measurement based on Alternate-Marking method in IPv6 domain.

Table of Contents

1. Introduction

The Alternate-Marking method, as presented in [I-D.draft-ietf-ippm-rfc8321bis], can be applied to perform packet loss, delay, and jitter measurements on live traffic. Likewise, [I-D.draft-ietf-ippm-rfc8889bis] generalizes and expands this methodology to measure any kind of unicast flow whose packets can follow several different paths in the multipoint-to-multipoint network.

The Alternate-Marking method, as described in [I-D.draft-ietf-ippm-rfc8321bis] and [I-D.draft-ietf-ippm-rfc8889bis], allows the synchronization of the measurements in different points by dividing the packet flow into batches. So it is possible to get coherent counters and show what is happening in every marking period for each monitored flow. Based on this ability, the method could be used to perform packet loss, delay and jitter measurements on live traffic.

Based on the Alternate-Marking method, this document discusses how to deploy in-situ flow performance measurement in IPv6 domain. The Flow Measurement Operation is described and the applications are proposed in Section 5.

In combination with the scalability of the IPv6 packet header and other in-situ flow measurement functions that may be supported in the future, a specific data structure is defined to carry the marking bits and other information required for flow measurement. The structure is called Flow Monitor Option, and details are in Section 3.

How to encapsulate the Flow Monitor Option in IPv6 traffic flow is discussed in Section 2. A new type of IPv6 Extension Header Option is proposed, Flow Monitor Option is encapsulated in Hop-by-Hop options Header or Destination Options Header depending on the measurement type.

1.1. 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.

1.2. Terminology

The definitions of the basic terms are identical to those found in Alternate Marking [I-D.draft-ietf-ippm-rfc8321bis] and Multipoint Alternate-Marking [I-D.draft-ietf-ippm-rfc8889bis].

The important new terms that need to be explained are listed below:

ACL: access-control list

2. Flow Measurement in IPv6 Network

2.1. Carrying Flow Measurement Indicators

The flow measurement method described in this document needs to add monitoring information for performing measurement to the flow. In IPv6, the general way to carry packet's additional information is IPv6 Extension Header. Several IPv6 Extension Headers have been defined in [RFC8200]. It is necessary to determine suitable IPv6 Extension Header to carry measuring data for deploying of performance measure in IPv6. In the domain where flow measurement is enabled, only the traffic to be measured carries the Flow Measurement Indicators structure.

There are two measurement types: End-to-End and Hop-by-Hop. The participating nodes in two types are different.

The source node allocates Flow Measurement Indicators structure defined in Section 2.2 and encodes it in packet. For End-to-End measurement, just destination node processes the Flow Measurement Indicators structure. According to Section 4.1 of [RFC8200], IPv6 Destination Options Header before the upper-layer header is appropriate for End-to-End measurement.

For Hop-by-Hop measurement, all nodes on the delivery path are expected to examine and process the Flow Measurement Indicators. According to [RFC8200], the Flow Measurement Indicators can be carried as an option of Hop-by-Hop Options Header.

2.2. Flow Measurement Indicators Definition

As description in Section 2.1, Flow Measurement Indicators is encoded in IPv6 Destination Options Header or IPv6 Hop-by-Hop Options Header. The Flow Measurement Indicators structure must be defined following IPv6 Option's principle.

This document defines Flow Monitor Option for flow measurement. Using Flow Monitor Option to marking packets required by Alternate-Marking, and to carry flow identity and measure parameters.

3. Definition of Flow Monitor Option

Flow Monitor Option is defined to carry Flow Measurement Indicators, below is detailed description.

3.1. Data Fields Format

The following figure shows the data field's format for Flow Monitor Option. This Flow Measurement Indicators structure can be encapsulated in the Hop-by-Hop Options Header and Destination Options Header.

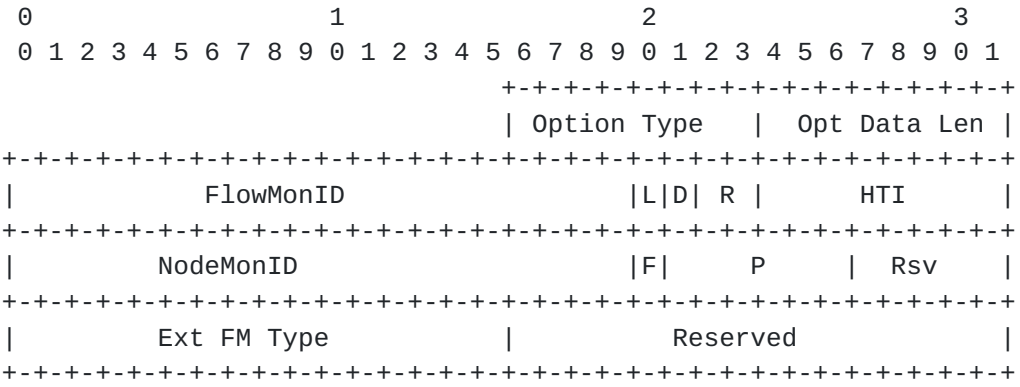


Figure 1: Flow Monitor Option

where:

- Option Type: 8-bit identifier of the type of Flow Monitor Option. The encoding format references Section 4.2 of [RFC8200]. The value is to be assigned by IANA.

- Opt Data Len: The length of the Option Data Fields of this option in bytes.

- FlowMonID: 20 bits unsigned integer. The FlowMon identifier is used to identify one flow in the node. See Section 4.1 for details.

- L: Loss Flag, a marking bit of packet loss measurement.

- D: Delay Flag, a marking bit of packet delay measurement.

- R: Reserved for future use, now initialized to zero for transmission and ignored on reception.

- HTI: Header Type Indication. It indicates the type of the option header, has the following value:

  0: Reserved, indicate that the format of Flow Monitor Option is the same as [I-D.ietf-6man-ipv6-alt-mark].

  1~15: Private type.

16~255: Extensible type value. When the value is 16, the format of the option header is as shown in Figure 2.

- NodeMonID: 20 bits unsigned integer. It is used to identify a node in the measurement domain, combined with the FlowMonID field to uniquely identify a monitored flow. Detail description sees Section 4.1.

- F: The marking bit of two-way flow measurement. If the field is set to 1, the end node generates reverse flow measurement configuration dynamically according to the current flow.

- P: 6 bits, measurement period. It has the following values:

000000: 1 second

000001: 10 seconds

000010: 30 seconds

000011: 60 seconds

000100: 300 seconds

Others: Reserved

- Ext FM Type: A 16 bits Bitmap for Extendable Flow Measurement type. The Bitmap can present 15 different measurement types. From bit 0 to 14, each bit presents a specific measurement type. The bit15 is reserved for extension Bitmap, 1 indicates carrying the extension Bitmap. The use case about Ext FM Type is described in Section 5.6.

4. Encapsulating Flow Monitor Option

When flow measurement is enabled, source node allocates Flow Monitor Option for monitored flows, fills measurement parameters, sets marking bits, and adds an extension header for packet encapsulating the Flow Monitor Option.

For Hop-by-Hop measurement, the Flow Monitor Option is encapsulated in the Hop-by-Hop Options Header.

For End-to-End measurement, the Flow Monitor Option is encapsulated in the Destination Options Header before the upper-layer header.

4.1. Flow Monitoring Identification

The Flow Monitoring Identification is required for some general reasons:

First, it helps to reduce the per node configuration. Otherwise, each node needs to configure an access-control list (ACL) for each of the monitored flows. Moreover, using a Flow Monitoring Identification allows a flexible granularity for the flow definition.

Second, it simplifies the counters handling. Hardware processing of flow tuples (and ACL matching) is challenging and often incurs into performance issues, especially in tunnel interfaces.

Third, it eases the data export encapsulation and correlation for the collectors.

The NodeMon identifier (NodeMonID) field is filled with the source node's identifier. The NodeMonID as configuration is set on the source node by the central controller. The controller ensures NodeMonID is unique within the measurement domain.

The FlowMon identifier (FlowMonID) field is used to uniquely identify a monitored flow within a specified source node. The FlowMonID can be uniformly assigned by the central controller, also can be algorithmically generated by the source node based on the flow information.

Using the combination of FlowMonID and NodeMonID, a monitored flow can be uniquely identified within the measurement domain. The FlowMonID field and NodeMonID field are set at the source node.

5. Flow Measurement Operation

[I-D.draft-ietf-ippm-rfc8321bis] describes a method to perform packet loss, delay and jitter measurements on live traffic. This section describes how the method can be applied in IPv6 network.

5.1. Packet Loss Measurement

The L marking bit in the Flow Monitor Option is used to color the flows that need packet loss measurement. By setting the L marking bit to 1 or 0 according to the measurement period filled in P field in the source node, the monitored flows can be split into consecutive blocks. The intermediate and end nodes read the L marking bit and identify the packet blocks. By counting the number of packets in each block and comparing the values measured by

different nodes along the path, it is possible to measure packet loss occurred in any single block between any two points.

5.2. Packet Delay Measurement

The same principle used to measure packet loss also can be applied to one-way delay measurement. Packet delay measurement references Double-Marking Method described in [I-D.draft-ietf-ippm-rfc8321bis] using the L marking bit and D marking bit in Flow Monitor Option.

The L marking bit is used to mark the alternate flow. By marking the L marking bit to 1 or 0, the monitored flows can be split into consecutive blocks. And, within this colored flow identified by the L marking bit, a second marked D marking bit is used to select the packets for measuring delay. The D marking bit creates a new set of marked packets that are fully identified over the network, so that a network node can store the timestamps of these packets; these timestamps can be compared with the timestamps of the same packets on a second node to compute packet delay values for each packet.

Likewise to packet delay measurement, the on-path jitter can be measured by measuring multiple blocks.

5.3. Measurement Type

For different measurement requirements, there are End-to-End measurement type and Hop-by-Hop measurement type.

With the End-to-End measurement type, it can measure the forwarding performance between source node and end node when the traffic passes through the measurement domain. The performance of each intermediate node or link is not cared about. Therefore, when using the End-to-End measurement type, only the source node and end node need to collect performance data and report data to controller.

With the Hop-by-Hop measurement type, each node along the path which has enabled performance measurement SHOULD collect performance data and report data to the controller when the traffic passes through the measurement domain.

Compared to the End-to-End measurement type, the Hop-by-Hop measurement type can more accurately locate the network packet loss and delay in position.

The measurement type is determined by the position of Flow Monitor Option in the IPv6 Extension Header. The Flow Monitor Option can be encapsulated in Hop-by-Hop Options Header or Destination Options Header. When it is encapsulated in the Hop-by-Hop Options Header,

each node along the path will deal with it. That is Hop-by-Hop measurement. When the Flow Monitor Option is encapsulated in the Destination Options Header, it means End-to-End measurement.

5.4. Two-way Flow Measurement

As described in [I-D.draft-ietf-ippm-rfc8321bis] the source node needs to virtually split traffic flows into consecutive blocks according to some methods, such as configuring an access-control list (ACL) for each of the monitored flows. But, if we want to measure bidirectional forwarding performance of monitored flows on the specified path, we need to configure ACLs associated monitored flows on the source node and end node at the same time. This will increase the configuration and maintenance workload. And this work is more complex, such as source IP addresses in the source node configuration need to be transformed as destination IP addresses in the end node, and other characteristics are similar.

Therefore, this document provides a two-way flow measurement method. It generates reverse flow measurement configuration dynamically in the end node according to the forward flow.

Two-way flow performance measurement is implemented as follows:

1. The source node configures ACLs for monitored flows that need bidirectional flow measurement.

2. When the source node receives the corresponding monitored flow, it encapsulates Flow Monitor Option into the IPv6 Extension Header, and sets the F field to 1.

3. When the end node receives the monitored flow which F field has been set to 1, it analysis the information of positive monitored flow, changes the source and destination information, dynamically generates ACLs with the characteristics of reverse monitored flows, and distributes configuration on end node.

4. At the same time, the end node assigns FlowMonID for reverse monitored flows, and reports the new reserve FlowMonID, the NodeMonID of the end node and the reverse flow information to controller.

5. When the end node receives the reserve monitored flow, the end node encapsulates Flow Monitor Option into IPv6 Extension Header, sets F field to 0, uses the FlowMonID and NodeMonID of end node, and fills other fields of Flow Monitor option according to the end node's configuration.

6. All nodes along the reserve path enabled performance measurement collect performance data, report to controller according Flow Monitor option in the packet header.

5.5. Data Collection and Report

Each node which participates in performance measurement collects performance data, records packet counts, received timestamps, sent timestamps, FlowMonID, NodeMonID and other related information specified by Flow Measure Type bitmap, and reports to the controller. For the source node, it needs to report characteristic information of monitored flow additionally.

The network nodes report to controller by Telemetry technique. The period of report can be the measurement period filled in the P field of Flow Monitor Option, can also be specified in the Telemetry subscription, or is designated by local configuration. This document does not limit the specific method.

5.6. Function Extension Consideration

5.6.1. The Use of Ext FM Type Bitmap

At present, the performance measurement is commonly attention to network packet loss, delay and jitter. However, with the expanding of network applications, other network performance parameters begin to be concerned, such as out-of-order rate. When network failure, controller wants to be able to obtain more abundant information, and in order to locate fault point quickly requires all nodes along the path to report current queue depth, input and output interface name, and so on.

By defining bits of Ext FM Type field in the Flow Monitor Option and carrying additional information in the monitored flows, the measurement function can be extended in the future.

For example, when the measurement period is small, in order to measure the out of order rate more accurately, the ingress node can specify the sequence number for the monitoring packet and carry it in the flow monitor option. Assume that bit0 of Ext FM Type is defined as an out-of-order measurement mark. When the source node receives monitored flow, it sets bit0 to indicate to count out-of-order packets. At the same time, it fills in additional information after Ext FM Type bitmap with ordinal Sequence parameters. After extension, the Flow Monitor Option package format is as follows:

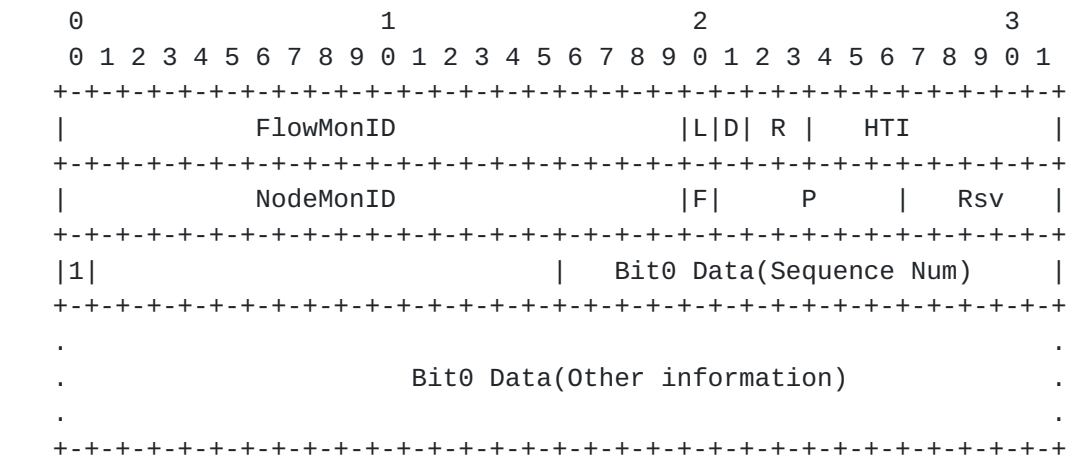


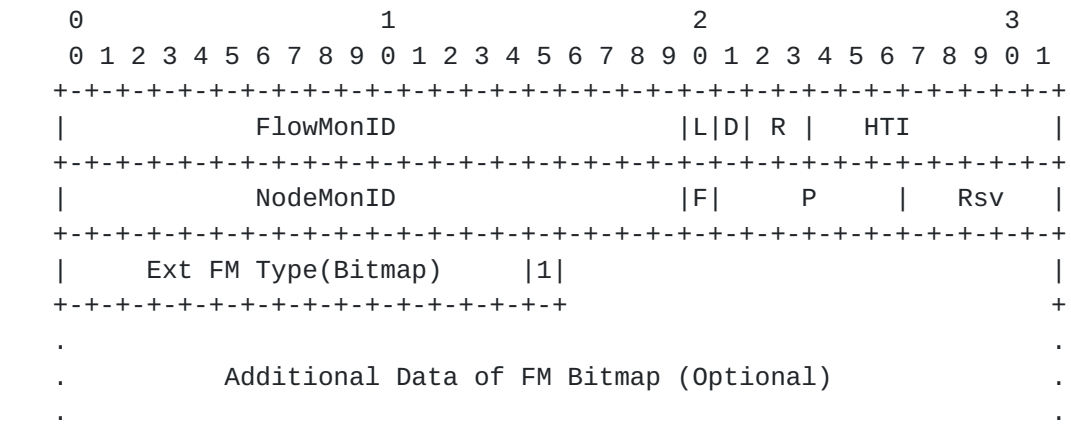
Figure 2: Use Bit0 For Out-of-order Measurement

Using the same method, the other bits of Ext FM Type field can be extended. Additional information is optional, whether it is carried is decided by the specified extension function.

5.6.2. Bitmap Extension

The Ext FM Type field has 16 bit, so 16 measurement functions can be extended. For general applications, the bitmap is enough. In order to reduce the effect on forwarding performance, it is also not recommended too much measurement processes at the same time.

However, considering functionality to be expanded in the future, bit15 is reserved, used to break the bitmap limit of 16. If bit15 is set to 1, it indicates carrying the extension bitmap. By default, bit15 is zero. For the performance of the data plane, it is also not recommended to define optional additional data too long.



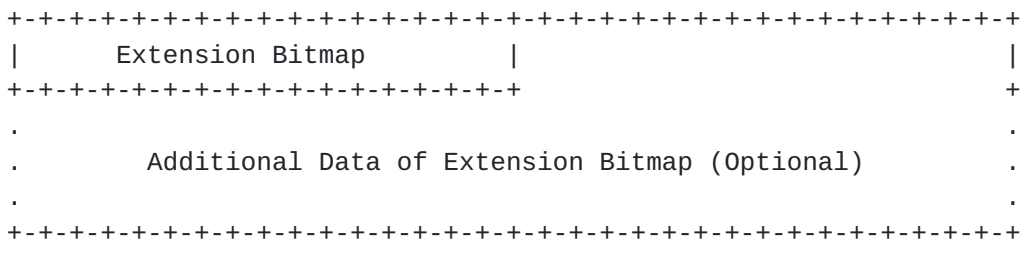


Figure 3: Extension Bitmap Format

Based on the previous out-of-order measurement example, for example, after the bits of Ext FM Type have been exhausted, use bit2 of Extension Bitmap to expand FM type. Flow Monitor Option package format is as shown below:

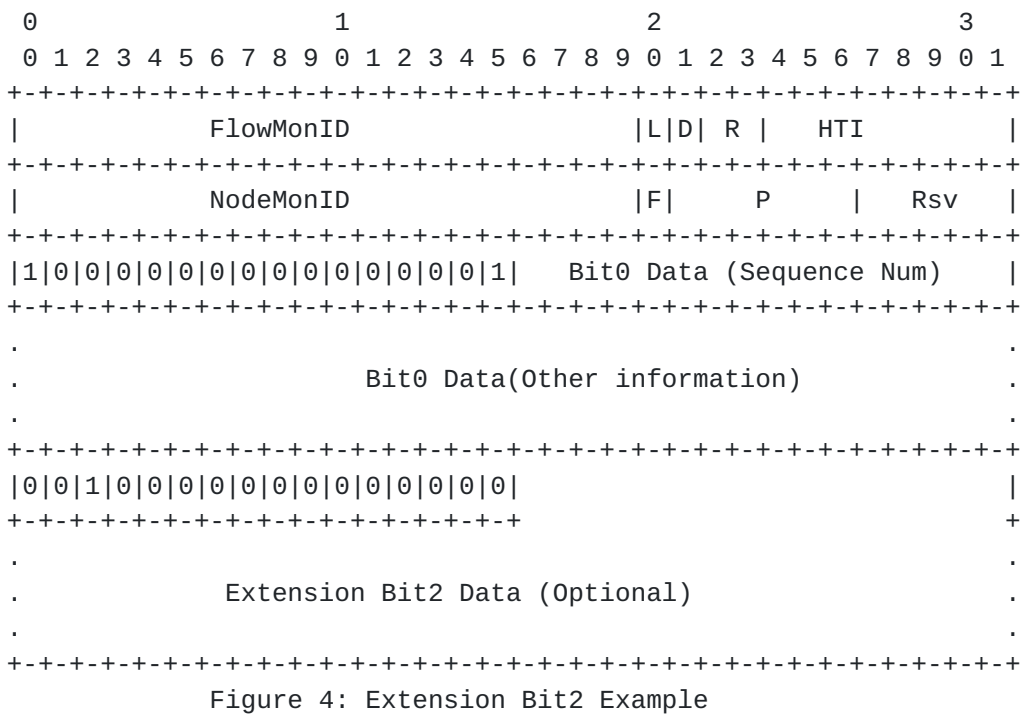


Figure 4: Extension Bit2 Example

6. IANA Considerations

The Flow Monitor Option Type should be assigned in IANA.

7. Security Considerations

The potential security threats of Alternate-Marking method have been described in detail in Section 10 of [I-D.draft-ietf-ippm-

rfc8321bis]. The performance measurement method described in this document does not introduce additional new security issues.

8. References

8.1. Normative References

[I-D.ietf-6man-ipv6-alt-mark] Fioccola, G., Zhou, T., Cociglio, M., Qin, F., and R. Pang, "IPv6 Application of the Alternate Marking Method", draft-ietf-6man-ipv6-alt-mark-16 (work in progress), July 2022.

[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>.

[RFC8200] Deering, S. and R. Hinden, "Internet Protocol, Version 6(IPv6) Specification", STD 86, RFC 8200, DOI 10.17487/RFC8200, July 2017, <https://www.rfc-editor.org/info/rfc8200>.

[I-D.draft-ietf-ippm-rfc8321bis] Fioccola, G., Ed., Cociglio, M., Mirsky, G., Mizrahi, T., Zhou, T., "Alternate-Marking Method", draft-ietf-ippm-rfc8321bis-03 (work in progress), July 2022.

[I-D.draft-ietf-ippm-rfc8889bis] Fioccola, G., Ed., Cociglio, M., Sapio, A., and Sisto, R., Zhou, T., " Multipoint Alternate-Marking Clustered Method", draft-ietf-ippm-rfc8889bis-03(work in progress), July 2022.

8.2. Informative References

TBD

9. Acknowledgments

The authors would like to thank the following for their valuable contributions of this document:

TBD

Authors' Addresses

Haojie Wang
China Mobile
China

Email: wanghaojie@chinamobile.com

Yisong Liu
China Mobile
China

Email: liuyisong@chinamobile.com

Changwang Lin
New H3C Technologies
China

Email: linchangwang.04414@h3c.com

Xiao Min
ZTE Corporation
China

Email: xiao.min2@zte.com.cn

Greg Mirsky
Ericsson

Email: gregimirsky@gmail.com