6MAN Working Group
Internet-Draft
Intended status: Standards Track
Expires: December 24, 2020

G. Fioccola
T. Zhou
Huawei
M. Cociglio
Telecom Italia
F. Qin
China Mobile
R. Pang
China Unicom

June 22, 2020

IPv6 Application of the Alternate Marking Method
draft-ietf-6man-ipv6-alt-mark-01

Abstract

This document describes how the Alternate Marking Method can be used as the passive performance measurement tool in an IPv6 domain and reports implementation considerations. It proposes how to define a new Extension Header Option to encode alternate marking technique and both Hop-by-Hop Options Header and Destination Options Header are considered.

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 BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

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 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 December 24, 2020.

Table of Contents

1. Introduction 2
2. Alternate Marking application to IPv6 3
3. Definition of the AltMark Option 4
3.1. Data Fields Format 4
4. Use of the AltMark Option 5
5. Alternate Marking Method Operation 7
5.1. Packet Loss Measurement 7
5.2. Packet Delay Measurement 8
5.3. Flow Monitoring Identification 9
5.3.1. Uniqueness of FlowMonID 10
5.4. Multipoint and Clustered Alternate Marking 10
5.5. Data Collection and Calculation 11
6. Security Considerations 11
7. IANA Considerations 12
8. Acknowledgements 12
9. References 12
9.1. Normative References 13
9.2. Informative References 13
Authors' Addresses 14

1. Introduction

[RFC8321] and [I-D.ietf-ippm-multipoint-alt-mark] describe a passive performance measurement method, which can be used to measure packet loss, latency and jitter on live traffic. Since this method is based on marking consecutive batches of packets, the method is often

referred as Alternate Marking Method.

The Alternate Marking Method has become mature to be implemented and encoded in the IPv6 protocol and this document defines how it can be used to measure packet loss and delay metrics in IPv6.

The format of the IPv6 addresses is defined in [RFC4291] while [RFC8200] defines the IPv6 Header, including a 20-bit Flow Label and the IPv6 Extension Headers. The Segment Routing Header (SRH) is defined in [RFC8754].

[I-D.fioccola-v6ops-ipv6-alt-mark] reported a summary on the possible implementation options for the application of the Alternate Marking Method in an IPv6 domain. This document, starting from the outcome of [I-D.fioccola-v6ops-ipv6-alt-mark], introduces a new TLV that can be encoded in the Options Headers (both Hop-by-Hop or Destination) for the purpose of the Alternate Marking Method application in an IPv6 domain. The case of SRH ([RFC8754]) is also discussed, anyway this is valid for all the types of Routing Header (RH).

2. Alternate Marking application to IPv6

The Alternate Marking Method requires a marking field. As mentioned, several alternatives have been analysed in [I-D.fioccola-v6ops-ipv6-alt-mark] such as IPv6 Extension Headers, IPv6 Address and Flow Label.

In consequence to the previous document and to the discussion within the community, it is possible to state that the only correct and robust choice that can actually be standardized would be the use of a new TLV to be encoded in the Options Header (Hop-by-Hop or Destination Option).

This approach is compliant with [RFC8200] indeed the Alternate Marking application to IPv6 involves the following operations:

- The source node is the only one that writes the Option Header to mark alternately the flow (for both Hop-by-Hop and Destination Option).

- In case of Hop-by-Hop Option Header carrying Alternate Marking bits, it is not inserted or deleted, but can be read by any node

along the path. The intermediate nodes may be configured to support this Option or not. Anyway this does not impact the traffic since the measurement can be done only for the nodes configured to read the Option.

- In case of Destination Option Header carrying Alternate Marking bits, it is not processed, inserted, or deleted by any node along the path until the packet reaches the destination node. Note that, if there is also a Routing Header (RH), any visited destination in the route list can process the Option Header.

Hop-by-Hop Option Header is also useful to signal to routers on the path to process the Alternate Marking, anyway it is to be expected that some routers cannot process it unless explicitly configured.

The optimization of both implementation and scaling of the Alternate Marking Method is also considered and a way to identify flows is required. The Flow Monitoring Identification field (FlowMonID), as introduced in the next sections, goes in this direction and it is used to identify a monitored flow.

Note that the FlowMonID is different from the Flow Label field of the IPv6 Header ([RFC8200]). Flow Label is used for application service, like load-balancing/equal cost multi-path (LB/ECMP) and QoS. Instead, FlowMonID is only used to identify the monitored flow. The reuse of flow label field for identifying monitored flows is not considered since it may change the application intent and forwarding behaviour. Furthermore the flow label may be changed en route and this may also violate the measurement task. Those reasons make the definition of the FlowMonID necessary for IPv6. Flow Label and FlowMonID within the same packet have different scope, identify different flows, and associate different uses.

An important point that will also be discussed in this document is the the uniqueness of the FlowMonID and how to allow disambiguation of the FlowMonID in case of collision. [RFC6437] states that the Flow Label cannot be considered alone to avoid ambiguity since it could be accidentally or intentionally changed en route for compelling operational security reasons and this could also happen to the IP addresses that can change due to NAT. But the Alternate

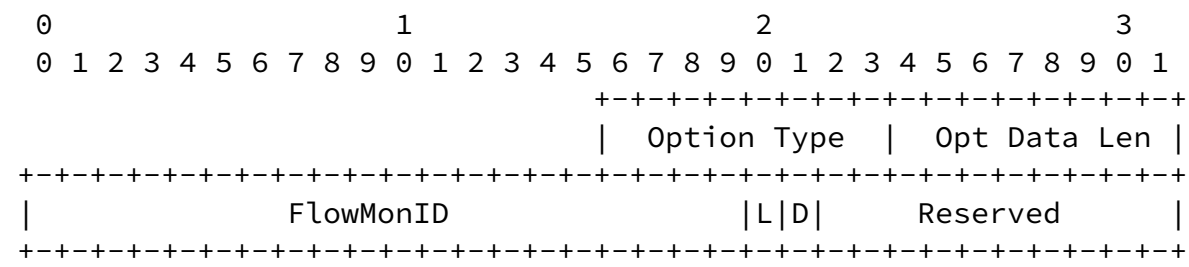
Marking is usually applied in a controlled domain, which would not have NAT and there is no security issue that would necessitate rewriting Flow Labels. So, for the purposes of this document, both IP addresses and Flow Label should not change in flight and, in some cases, they could be considered together with the FlowMonID for disambiguation.

3. Definition of the AltMark Option

The desired choice is to define a new TLV for the Options Extension Headers, carrying the data fields dedicated to the alternate marking method.

3.1. Data Fields Format

The following figure shows the data fields format for enhanced alternate marking TLV. This AltMark data is expected to be encapsulated in the IPv6 Options Headers (Hop-by-Hop or Destination Option).



where:

- Option Type: 8 bit identifier of the type of Option that needs to be allocated. Unrecognised Types MUST be ignored on receipt. For Hop-by-Hop Options Header or Destination Options Header, [RFC8200] defines how to encode the three high-order bits of the Option Type field. The two high-order bits specify the action that must be taken if the processing IPv6 node does not recognize the Option Type; for AltMark these two bits MUST be set to 00 (skip over this Option and continue processing the header). The third-highest-order bit specifies whether or not the Option Data can change en route to the packet's final destination; for AltMark the value of this bit MUST be set to 0 (Option Data does not change en route).

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

- FlowMonID: 20 bits unsigned integer. The FlowMon identifier is described hereinafter.

- L: Loss flag for Packet Loss Measurement as described hereinafter;

- D: Delay flag for Single Packet Delay Measurement as described hereinafter;

- Reserved: is reserved for future use. These bits MUST be set to zero on transmission and ignored on receipt.

4. Use of the AltMark Option

The AltMark Option is the best way to implement the Alternate Marking method and can be carried by the Hop-by-Hop Options header and the Destination Options header. In case of Destination Option, it is processed only by the source and destination nodes: the source node inserts and the destination node removes it. While, in case of Hop-by-Hop Option, it may be examined by any node along the path, if explicitly configured to do so. In this way an unrecognized Hop-by-Hop Option may be just ignored without impacting the traffic.

So it is important to highlight that the Option Layout can be used both as Destination Option and as Hop-by-Hop Option depending on the Use Cases and it is based on the chosen type of performance measurement. In general, it is needed to perform both end to end and hop by hop measurements, and the alternate marking methodology allows, by definition, both performance measurements. Anyway, in many cases the end-to-end measurement is not enough and it is required also the hop-by-hop measurement, so the most complete choice is the Hop-by-Hop Options Header.

IPv6, as specified in [RFC8200], allows nodes to optionally process Hop-by-Hop headers. Specifically the Hop-by-Hop Options header is not inserted or deleted, but may be examined or processed by any node along a packet's delivery path, until the packet reaches the node (or each of the set of nodes, in the case of multicast) identified in the

Destination Address field of the IPv6 header. Also, it is expected that nodes along a packet's delivery path only examine and process the Hop-by-Hop Options header if explicitly configured to do so.

The Hop-by-Hop Option defined in this document is designed to take advantage of the property of how Hop-by-Hop options are processed. Nodes that do not support this Option SHOULD ignore them. This can mean that, in this case, the performance measurement does not account for all links and nodes along a path.

Another application that can be mentioned is the presence of a Routing Header, in particular it is possible to consider SRv6. SRv6 leverages the Segment Routing header which consists of a new type of routing header. Like any other use case of IPv6, Hop-by-Hop and Destination Options are useable when SRv6 header is present. Because SRv6 is a routing header, Destination Options before the routing header are processed by each destination in the route list.

In summary, it is possible to list the alternative possibilities:

- Destination Option => measurement only by node in Destination Address.
- Hop-by-Hop Option => every router on the path with feature enabled.
- Destination Option + SRH => every node that is an identity in the SR path.

In general, Hop-by-Hop and Destination Options are the most suitable ways to implement Alternate Marking.

It is worth mentioning that new Hop-by-Hop Options are not strongly recommended in [RFC7045] and [RFC8200], unless there is a clear justification to standardize it, because nodes may be configured to ignore the Options Header, drop or assign packets containing an Options Header to a slow processing path. In case of the AltMark data fields described in this document, the motivation to standardize a new Hop-by-Hop Option is that it is needed for OAM. An intermediate node can read it or not but this does not affect the

packet behavior. The source node is the only one that writes the Hop-by-Hop Option to mark alternately the flow, so, the performance measurement can be done for those nodes configured to read this Option, while the others are simply not considered for the metrics.

In addition to the previous alternatives, for legacy network it is possible to mention a non-conventional application of the Destination Option for the hop by hop usage. [RFC8200] defines that the nodes along a path examine and process the Hop-by-Hop Options header only if Hop-by-Hop processing is explicitly configured. On the other hand, using the Destination Option for hop by hop action would cause worse performance than Hop-by-Hop. The only motivation for the hop by hop usage of Destination Options can be for compatibility reasons but in general it is not recommended.

5. Alternate Marking Method Operation

This section describes how the method operates. [RFC8321] introduces several alternatives but in this section the most applicable methods are reported and a new fied is introduced to facilitate the deployment and improve the scalability.

5.1. Packet Loss Measurement

The measurement of the packet loss is really straightforward. The packets of the flow are grouped into batches, and all the packets within a batch are marked by setting the L bit (Loss flag) to a same value. The source node can switch the value of the L bit between 0 and 1 after a fixed number of packets or according to a fixed timer, and this depends on the implementation. By counting the number of packets in each batch and comparing the values measured by different network nodes along the path, it is possible to measure the packet loss occurred in any single batch between any two nodes. Each batch represents a measurable entity unambiguously recognizable by all network nodes along the path.

It is important to mention that for the application of this method there are two elements to consider: the clock error between network nodes and the network delay. These can create offsets between the batches and out-of-order of the packets. The consequence is that it

is necessary to define a waiting interval where to get stable

counters and to avoid these issues. In addition this implies that the length of the batches MUST be chosen large enough so that it is not affected by those factors.

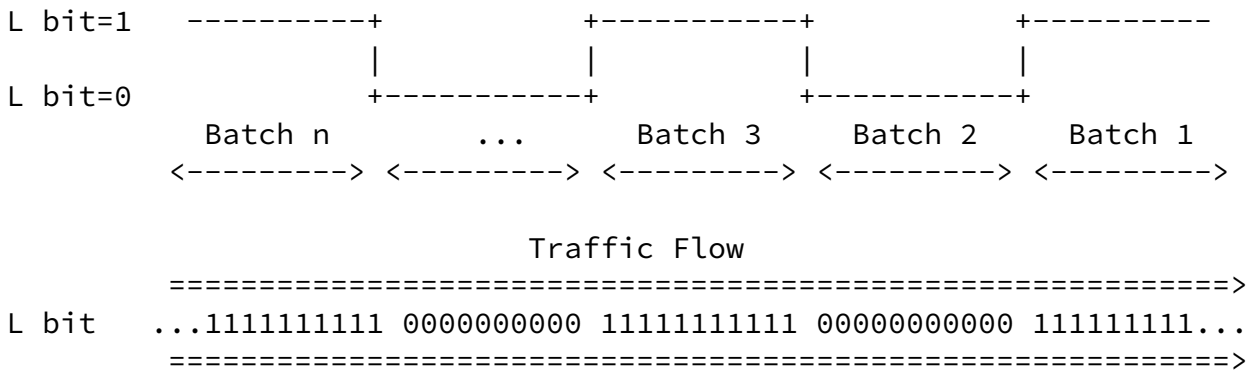


Figure 1: Packet Loss Measurement and Single-Marking Methodology using L bit

5.2. Packet Delay Measurement

The same principle used to measure packet loss can be applied also to one-way delay measurement. Delay metrics MAY be calculated using the two possibilities:

1. Single-Marking Methodology: This approach uses only the L bit to calculate both packet loss and delay. In this case, the D flag MUST be set to zero on transmit and ignored by the monitoring points. The alternation of the values of the L bit can be used as a time reference to calculate the delay. Whenever the L bit changes and a new batch starts, a network node can store the timestamp of the first packet of the new batch, that timestamp can be compared with the timestamp of the first packet of the same batch on a second node to compute packet delay. Anyway this measurement is accurate only if no packet loss occurs and if there is no packet reordering at the edges of the batches. A different approach can also be considered and it is based on the concept of the mean delay. The mean delay for each batch is calculated by considering the average arrival time of the packets for the relative batch. There are limitations also in this case indeed, each node needs to collect all the timestamps and calculate the average timestamp for each batch. In addition the information is limited to a mean value.

2. Double-Marking Methodology: This approach is more complete and uses the L bit only to calculate packet loss and the D bit (Delay

flag) is fully dedicated to delay measurements. The idea is to use the first marking with the L bit to create the alternate flow and, within the batches identified by the L bit, a second marking is used to select the packets for measuring delay. The D 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. The most efficient and robust mode is to select a single double-marked packet for each batch, in this way there is no time gap to consider between the double-marked packets to avoid their reorder. If a double-marked packet is lost, the delay measurement for the considered batch is simply discarded, but this is not a big problem because it is easy to recognize the problematic batch and skip the measurement just for that one. So in order to have more information about the delay and to overcome out-of-order issues this method is preferred.

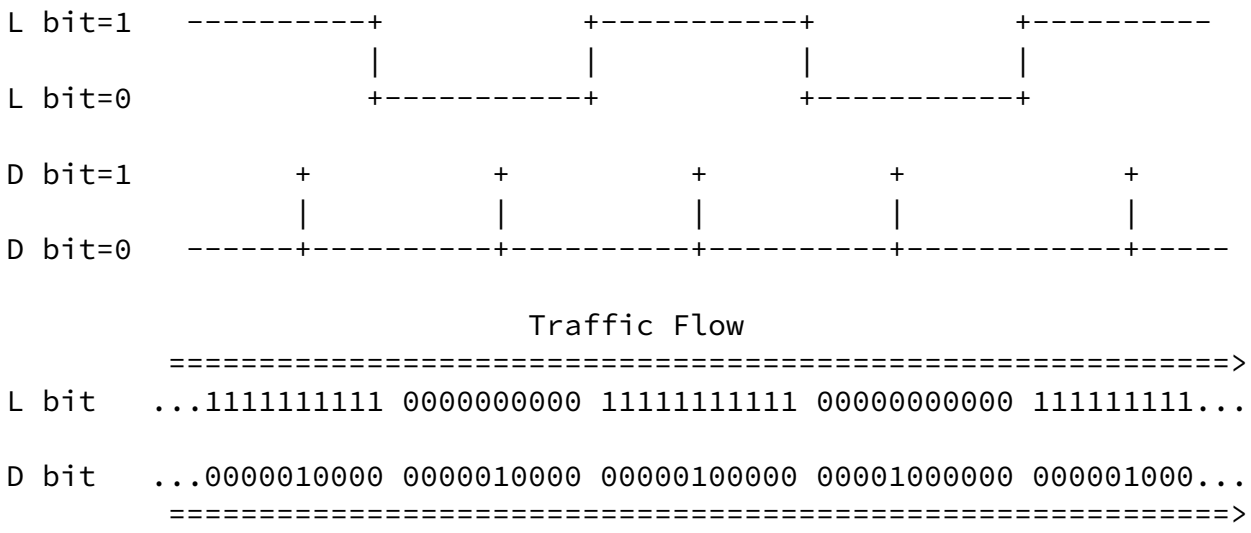


Figure 2: Double-Marking Methodology using L bit and D bit

Similar to packet delay measurement (both for Single Marking and Double Marking), the method can also be used to measure the inter-arrival jitter.

5.3. Flow Monitoring Identification

The Flow Monitoring Identification (FlowMonID) 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 identifier 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 FlowMon identifier field is to uniquely identify a monitored flow within the measurement domain. The field is set at the source node. The FlowMonID can be uniformly assigned by the central controller or algorithmically generated by the source node. The latter approach cannot guarantee the uniqueness of FlowMonID but it may be preferred for local or private network, where the conflict probability is small due to the large FlowMonID space.

5.3.1. Uniqueness of FlowMonID

It is important to note that if the 20 bit FlowMonID is set independently and pseudo randomly there is a chance of collision. So, in some cases, FlowMonID could not be sufficient for uniqueness.

In general the probability of a flow identifier uniqueness correlates to the amount of entropy of the inputs. For instance, using the well-known birthday problem in probability theory, if the 20 bit FlowMonID is set independently and pseudo randomly without any additional input entropy, there is a 50% chance of collision for just 1206 flows. For a 32 bit identifier the 50% threshold jumps to 77,163 flows and so on. So, for more entropy, FlowMonID can either be combined with other identifying flow information in a packet (e.g. it is possible to consider the hashed 3-tuple Flow Label, Source and Destination addresses) or the FlowMonID size could be increased.

This issue is more visible when the FlowMonID is pseudo randomly generated by the source node and there needs to tag it with additional flow information to allow disambiguation. While, in case of a centralized controller, the controller should set FlowMonID by

considering these aspects and instruct the nodes properly in order to guarantee its uniqueness.

5.4. Multipoint and Clustered Alternate Marking

The Alternate Marking method can also be extended to any kind of multipoint to multipoint paths, and the network clustering approach allows a flexible and optimized performance measurement, as described in [I-D.ietf-ippm-multipoint-alt-mark].

The Cluster is the smallest identifiable subnetwork of the entire Network graph that still satisfies the condition that the number of packets that goes in is the same that goes out. With network clustering, it is possible to use the partition of the network into clusters at different levels in order to perform the needed degree of detail. So, for Multipoint Alternate Marking, FlowMonID can identify in general a multipoint-to-multipoint flow and not only a point-to-point flow.

5.5. Data Collection and Calculation

The nodes enabled to perform performance monitoring collect the value of the packet counters and timestamps. There are several alternatives to implement Data Collection and Calculation, but this is not specified in this document.

6. Security Considerations

This document aims to apply a method to perform measurements that does not directly affect Internet security nor applications that run on the Internet. However, implementation of this method must be mindful of security and privacy concerns.

There are two types of security concerns: potential harm caused by the measurements and potential harm to the measurements.

Harm caused by the measurement: Alternate Marking implies modifications on the fly to an Option Header of IPv6 packets but this must be performed in a way that does not alter the quality of service experienced by the packets and that preserves stability and performance of routers doing the measurements. The advantage of the Alternate Marking method is that the marking bits are the only

information that is exchanged between the network nodes. Therefore, network reconnaissance through passive eavesdropping on data-plane traffic does not allow attackers to gain information about the network performance. Moreover, Alternate Marking should usually be applied in a controlled domain and this also helps to limit the problem.

Harm to the Measurement: Alternate Marking measurements could be harmed by routers altering the marking of the packets or by an attacker injecting artificial traffic. Since the measurement itself may be affected by network nodes along the path intentionally altering the value of the marking bits of IPv6 packets, the Alternate Marking should be applied in the context of a controlled domain, where the network nodes are locally administered and this type of attack can be avoided. Indeed the source and destination addresses are within the controlled domain and therefore it is unlikely subject

to hijacking of packets, because it is possible to filter external packets at the domain boundaries. In addition, an attacker cannot gain information about network performance from a single monitoring point; it must use synchronized monitoring points at multiple points on the path, because they have to do the same kind of measurement and aggregation as Alternate Marking requires.

The privacy concerns of network measurement are limited because the method only relies on information contained in the Option Header without any release of user data. Although information in the Option Header is metadata that can be used to compromise the privacy of users, the limited marking technique seems unlikely to substantially increase the existing privacy risks from header or encapsulation metadata.

The Alternate Marking application described in this document relies on an time synchronization protocol. Thus, by attacking the time protocol, an attacker can potentially compromise the integrity of the measurement. A detailed discussion about the threats against time protocols and how to mitigate them is presented in [RFC7384].

7. IANA Considerations

The Option Type should be assigned in IANA's "Destination Options and Hop-by-Hop Options" registry.

This draft requests the following IPv6 Option Type assignments from the Destination Options and Hop-by-Hop Options sub-registry of Internet Protocol Version 6 (IPv6) Parameters (https://www.iana.org/assignments/ipv6-parameters/).

Hex Value	Binary Value act chg rest	Description	Reference
TBD	00 0 tbd	AltMark	[This draft]

8. Acknowledgements

The authors would like to thank Bob Hinden, Ole Troan, Tom Herbert, Stefano Previdi, Brian Carpenter, Eric Vyncke, Ron Bonica for the precious comments and suggestions.

9. References

9.1. Normative References

[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>.

9.2. Informative References

[I-D.fioccola-v6ops-ipv6-alt-mark]
Fioccola, G., Velde, G., Cociglio, M., and P. Muley, "IPv6 Performance Measurement with Alternate Marking Method", draft-fioccola-v6ops-ipv6-alt-mark-01 (work in progress), June 2018.

[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", draft-ietf-ippm-multipoint-alt-mark-09 (work in progress), March 2020.

[RFC4291] Hinden, R. and S. Deering, "IP Version 6 Addressing Architecture", RFC 4291, DOI 10.17487/RFC4291, February 2006, <https://www.rfc-editor.org/info/rfc4291>.

[RFC6437] Amante, S., Carpenter, B., Jiang, S., and J. Rajahalme, "IPv6 Flow Label Specification", RFC 6437, DOI 10.17487/RFC6437, November 2011, <https://www.rfc-editor.org/info/rfc6437>.

[RFC7045] Carpenter, B. and S. Jiang, "Transmission and Processing of IPv6 Extension Headers", RFC 7045, DOI 10.17487/RFC7045, December 2013, <https://www.rfc-editor.org/info/rfc7045>.

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

[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", RFC 8321, DOI 10.17487/RFC8321, January 2018, <https://www.rfc-editor.org/info/rfc8321>.

[RFC8754] Filsfils, C., Ed., Dukes, D., Ed., Previdi, S., Leddy, J., Matsushima, S., and D. Voyer, "IPv6 Segment Routing Header (SRH)", RFC 8754, DOI 10.17487/RFC8754, March 2020, <https://www.rfc-editor.org/info/rfc8754>.

Authors' Addresses

Giuseppe Fioccola
Huawei
Riesstrasse, 25
Munich 80992
Germany

Email: giuseppe.fioccola@huawei.com

Tianran Zhou
Huawei
156 Beiqing Rd.
Beijing 100095
China

Email: zhoutianran@huawei.com

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

Email: mauro.cociglio@telecomitalia.it

Fengwei Qin
China Mobile
32 Xuanwumenxi Ave.
Beijing 100032
China

Email: qinfengwei@chinamobile.com

Ran Pang
China Unicom
9 Shouti South Rd.
Beijing 100089

China

Email: pangran@chinaunicom.cn