6MAN Working Group   G. Fioccola
Internet-Draft   T. Zhou
Intended status: Standards Track   Huawei
Expires: December 24, 2021   M. Cociglio
                                                     Telecom Italia
                                                             F. Qin
                                                       China Mobile
                                                            R. Pang
                                                       China Unicom
                                                      June 22, 2021

IPv6 Application of the Alternate Marking Method
draft-ietf-6man-ipv6-alt-mark-07

Abstract

   This document describes how the Alternate Marking Method can be used
   as a passive performance measurement tool in an IPv6 domain. It
   defines a new Extension Header Option to encode Alternate Marking
   information in both the Hop-by-Hop Options Header and Destination
   Options Header.

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, 2021.

Copyright Notice

Table of Contents

1. Introduction

   [RFC8321] and [RFC8889] 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 to as the Alternate
   Marking Method.

   This document defines how the Alternate Marking Method can be used to
   measure performance metrics in IPv6. The rationale is to apply the
   Alternate Marking methodology to IPv6 and therefore allow detailed
   packet loss, delay and delay variation measurements both hop-by-hop
   and end-to-end to exactly locate the issues in an IPv6 network.

   The Alternate Marking is an on-path telemetry technique and consists
   in synchronizing the measurements in different points of a network by

switching the value of a marking bit and therefore divide the packet
flow into batches. Each batch represents a measurable entity
unambiguously recognizable by all network nodes along the path. By
counting the number of packets in each batch and comparing the values
measured by different nodes, it is possible to precisely measure the
packet loss. In a similar way the alternation of the values of the
marking bits can be used as a time reference to calculate the delay
and delay variation. The Alternate Marking operation is further
described in Section 5.

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

[I-D.fioccola-v6ops-ipv6-alt-mark] summarizes 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 (type-
length-value) that can be encoded in the Options Headers (Hop-by-Hop
or Destination) for the purpose of the Alternate Marking Method
application in an IPv6 domain. While the case of Segment Routing
Header (SRH), defined in [RFC8754], is also discussed, it is valid
for all the types of Routing Header (RH).

The threat model for the application of the Alternate Marking Method
in an IPv6 domain is reported in Section 6. As for all the on-path
telemetry technique, the only definitive solution is that this
methodology MUST be applied in a controlled domain and therefore the
application to untrusted domain is NOT RECOMMENDED.

1.1. Terminology

This document uses the terms related to the Alternate Marking Method
as defined in [RFC8321] and [RFC8889].

1.2. 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.

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.

[I-D.fioccola-v6ops-ipv6-alt-mark] analyzed and discussed all the
available possibilities and the drawbacks:

   Reusing existing Extension Header for Alternate Marking leads to a
   non-optimized implementation;

   Using the IPv6 destination address to encode the Alternate Marking
   processing is very expensive;

   Using the IPv6 Flow Label for Alternate Marking conflicts with the
   utilization of the Flow Label for load distribution purpose
   ([RFC6438]).

In the end, [I-D.fioccola-v6ops-ipv6-alt-mark] demonstrated that a
new Hop-by-Hop or a new Destination Option was the best approach.

The approach for the Alternate Marking application to IPv6 specified
in this memo is compliant with [RFC8200]. It involves the following
operations:

o 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). The intermediate nodes and destination node MUST only
   read the marking values of the option without modifying the Option
   Header.

o 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 and the measurement can be done only
   for the nodes configured to read the Option. As further discussed
   in Section 4, the presence of the hop-by-hop option should not
   affect the traffic throughput both on nodes that do not recognize
   this option and on the nodes that support it. However it is
   important to mention that there is a difference between the theory
   and the implementation and it can happen that packets with hop-by-
   hop option could also be skipped or processed in the slow path.
   While some proposals are trying to address this problem
   ([I-D.peng-v6ops-hbh], [I-D.hinden-6man-hbh-processing]), these
   aspects are out of the scope for this document.

o 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. However, as said, routers
will examine this option if properly 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 Section 5.3, goes in this direction and it is used to
identify a monitored flow.

The FlowMonID is different from the Flow Label field of the IPv6
Header ([RFC6437]). The Flow Label field in the IPv6 header is used
by a source to label sequences of packets to be treated in the
network as a single flow and, as reported in [RFC6438], it can be
used for load-balancing/equal cost multi-path (LB/ECMP). 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. Also, since the Flow Label is pseudo-
random, there is always a finite probability of collision. Those
reasons make the definition of the FlowMonID necessary for IPv6.
Indeed, the FlowMonID is designed and only used to identify the
monitored flow. Flow Label and FlowMonID within the same packet are
totally disjoint, have different scope, identify different flows, and
are intended for different use cases.

The rationale for the FlowMonID is further discussed in Section 5.3.
This 20 bit field allows easy and flexible identification of the
monitored flow and enables a finer granularity and improved
measurement correlation. An important point that will be discussed
in Section 5.3.1 is the uniqueness of the FlowMonID and how to allow
disambiguation of the FlowMonID in case of collision.

The following section highlights an important requirement for the
application of the Alternate Marking to IPv6. The concept of the
controlled domain is explained and it is considered an essential
precondition, as also highlighted in Section 6.

2.1. Controlled Domain

[RFC8799] introduces the concept of specific limited domain solutions
and, in this regard, it is reported the IPv6 Application of the
Alternate Marking Method as an example.

IPv6 has much more flexibility than IPv4 and innovative applications
have been proposed, but for a number of reasons, such as the
policies, options supported, the style of network management and
security requirements, it is suggested to limit some of these
applications to a controlled domain. This is also the case of the
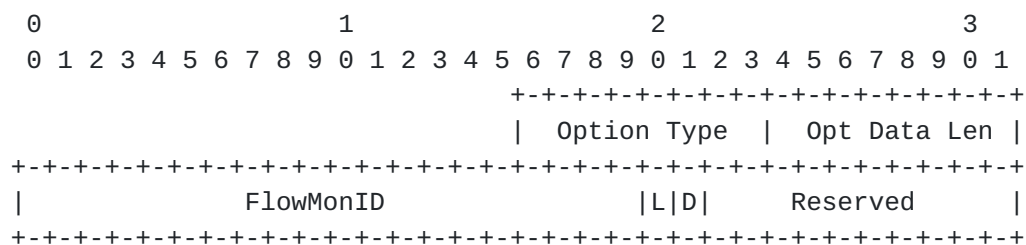Alternate Marking application to IPv6 as assumed hereinafter.

Therefore, the IPv6 application of the Alternate Marking Method MUST
NOT be deployed outside a controlled domain. It is RECOMMENDED that
an implementation can be able to reject packets that carry Alternate
Marking data and are entering or leaving the controlled domains. The
security considerations clarify this requirement and are reported in
Section 6.

3. Definition of the AltMark Option

The definition of a new TLV for the Options Extension Headers,
carrying the data fields dedicated to the Alternate Marking method,
is reported below.

3.1. Data Fields Format

The following figure shows the data fields format for enhanced
Alternate Marking TLV. This AltMark data can be encapsulated in the
IPv6 Options Headers (Hop-by-Hop or Destination Option).

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
                                +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
                                |  Option Type  |  Opt Data Len |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |              FlowMonID               |L|D|     Reserved      |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

where:

o Option Type: 8 bit identifier of the type of Option that needs to
   be allocated. Unrecognized 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).
   In this way, since the three high-order bits of the AltMark Option

      are set to 000, it means that nodes can simply skip this Option if
      they do not recognize and that the data of this Option do not
      change en route, indeed the source is the only one that can write
      it.

   o Opt Data Len: 4. It is the length of the Option Data Fields of
      this Option in bytes.

   o FlowMonID: 20 bits unsigned integer. The FlowMon identifier is
      described in Section 5.3. As further discussed below, it has been
      picked 20 bit since it is a reasonable value and a good compromise
      in relation to the chance of collision if it is set pseudo
      randomly by the source node or set by a centralized controller.

   o L: Loss flag for Packet Loss Measurement as described in
      Section 5.1;

   o D: Delay flag for Single Packet Delay Measurement as described in
      Section 5.2;

   o 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 it is 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.

   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. But, 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. A new
type of Routing Header, referred as SRH, has been defined for SRv6.
Like any other use case of IPv6, Hop-by-Hop and Destination Options
are useable when SRv6 header is present. Because SRv6 is implemented
through a Segment Routing Header (SRH), Destination Options before
the Routing Header are processed by each destination in the route
list, that means, in case of SRH, by every SR node that is identified
by the SR path. More details about the SRv6 application are
described in [I-D.fz-spring-srv6-alt-mark].

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

o Destination Option not preceding a Routing Header => measurement
   only by node in Destination Address.

o Hop-by-Hop Option => every router on the path with feature
   enabled.

o Destination Option preceding a Routing Header => every destination
   node in the route list.

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 (Operations,
Administration, and Maintenance). 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.

It is important to highlight that the definition of the Hop-by-Hop
Options in this document is designed to minimize throughput impact
both on nodes that do not recognize the Option and on node that
support it. Indeed, the three high-order bits of the Options Header
defined in this draft are 000 and, in theory, as per [RFC8200] and
[I-D.hinden-6man-hbh-processing], this means "skip if do not
recognize and data do not change en route". [RFC8200] also mentions
that the nodes only examine and process the Hop-by-Hop Options header
if explicitly configured to do so. For these reasons, this HbH
Option should not affect the throughput. However, in practice, it is
important to be aware for the implementation that the things may be
different and it can happen that packets with Hop-by-Hop are forced
onto the slow path, but this is a general issue, as also explained in
[I-D.hinden-6man-hbh-processing].

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 field 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. The source node is the only
one that marks the packets to create the batches, while the
intermediate nodes only read the marking values and identify the
packet batches. 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.

Both fixed number of packets and fixed timer can be used by the
source node to create packet batches. But, as also explained in
[RFC8321], using a fixed timer for the switching offers better
control over the method, indeed the length of the batches can be
chosen large enough to simplify the collection and the comparison of
the measures taken by different network nodes. In the implementation
the counters can be sent out by each node to the controller that is
responsible for the calculation. It is also possible to exchange

   this information by using other on-path techniques. But this is out
   of scope for this document.

   Packets with different L values may get swapped at batch boundaries,
   and in this case, it is required that each marked packet can be
   assigned to the right batch by each router. 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 mathematical formula on timing aspects,
   explained in section 3.2 of [RFC8321], must be satisfied and it takes
   into considerations the different causes of reordering such as clock
   error and network delay. The assumption is to define the available
   counting interval where to get stable counters and to avoid these
   issues. Specifically, if the effects of network delay are ignored,
   the condition to implement the methodology is that the clocks in
   different nodes MUST be synchronized to the same clock reference with
   an accuracy of +/- B/2 time units, where B is the fixed time duration
   of the block. In this way each marked packet can be assigned to the
   right batch by each node. Usually the counters can be taken in the
   middle of the batch period to be sure to take still counters. In a
   few words this implies that the length of the batches MUST be chosen
   large enough so that the method is not affected by those factors.
   The length of the batches can be determined based on the specific
   deployment scenario.

   L bit=1   ----------+   +-----------+   +----------
                       | | | |
   L bit=0   +-----------+   +-----------+
           Batch n   ...   Batch 3   Batch 2   Batch 1
         <---------> <---------> <---------> <---------> <--------->

                            Traffic Flow
           ===>
   L bit   ...1111111111 0000000000 1111111111 00000000000 111111111...
           ===>

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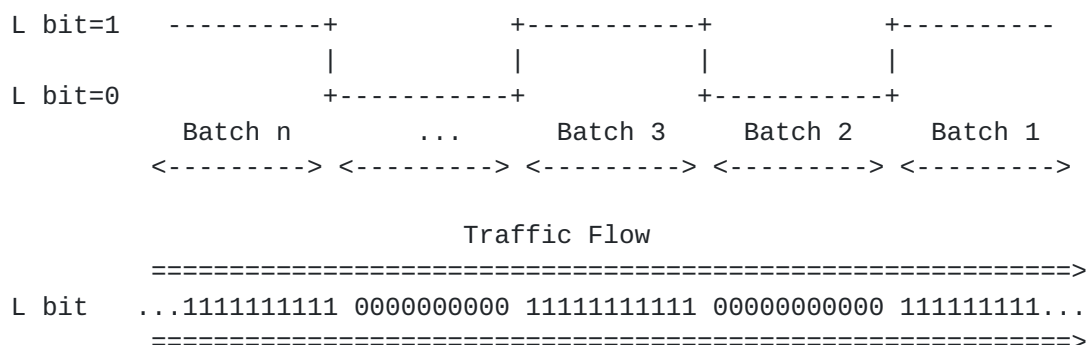   using L bit

   It is worth mentioning that the length of the batches is considered
   stable over time in the previous figure. In theory, it is possible
   to change the length of batches over time and and among different
   flows for more flexibility. But, in practice, it could complicate
   the correlation of the information.

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. But 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. Regarding the rule for
       the selection of the packet to be double-marked, the same
       considerations in Section 5.1 apply also here and the double-
       marked packet can be chosen within the available counting
       interval that is not affected by factors such as clock errors.
       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.

   In summary the approach with double marking is better than the
   approach with single marking. Moreover the two approaches can also
   be combined to have even more information and statistics on delay.

   Similar to what said in Section 5.1 for the packet counters, in the
   implementation the timestamps can be sent out to the controller that
   is responsible for the calculation or could also be exchanged using
   other on-path techniques. But this is out of scope for this
   document.

```
L bit=1   ----------+               +-----------+               +----------
                    |          |          |          |
L bit=0             +-----------+           +-----------+

D bit=1        +           +           +           +           +
               |          |          |          |          |
D bit=0   ------+----------+----------+----------+-----------+-----

                              Traffic Flow
          ============================================================>
L bit   ...1111111111 0000000000 11111111111 00000000000 111111111...

D bit   ...0000010000 0000010000 00000100000 00001000000 000001000...
          ============================================================>
```
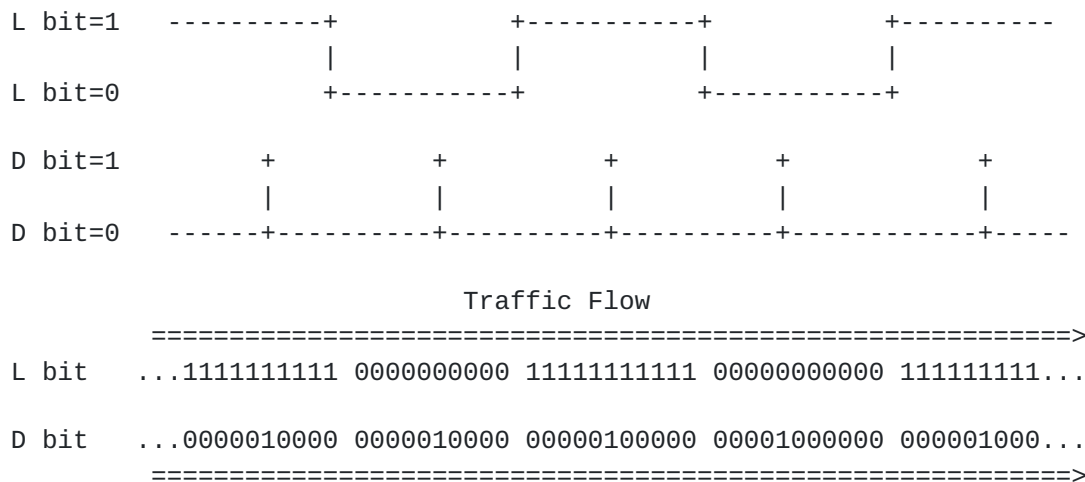
          Figure 2: Double-Marking Methodology using L bit and D bit

   Likewise 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:

   o 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.

   o 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.

   o 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 set in two ways:

      * It can be uniformly assigned by the central controller. Since
        the controller knows the network topology, it can set the value
        properly to avoid or minimize ambiguity and guarantee the
        uniqueness.

      * It can be algorithmically generated by the source node, that can
        set it pseudo-randomly with some chance of collision. This
        approach cannot guarantee the uniqueness of FlowMonID but it may
        be preferred for local or private networks, where the conflict
        probability is small due to the large FlowMonID space.

   The value of 20 bits has been selected for the FlowMonID since it is
   a good compromise and implies a low rate of ambiguous FlowMonIDs that
   can be considered acceptable in most of the applications. Indeed
   with 20 bits the number of combinations is 1048576.

   if the FlowMonID is set by the source node, the intermediate nodes
   can read the FlowMonIDs from the packets in flight and act
   accordingly. While, if the FlowMonID is set by the controller, both
   possibilities are feasible for the intermediate nodes which can learn
   by reading the packets or can be instructed by the controller.

   When all values in the FlowMonID space are consumed, the centralized
   controller can keep track and reassign the values that are not used
   any more by old flows, while if the FlowMonID is pseudo randomly
   generated by the source, conflicts and collisions are possible.

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.
   Indeed, by 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 1206 flows. 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 [RFC8889].

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.

There are documents on the control plane mechanisms of Alternate
Marking, e.g. [I-D.ietf-idr-sr-policy-ifit],
[I-D.chen-pce-pcep-ifit].

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 by the
source node 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. As
already discussed in Section 4, it is RECOMMENDED that the AltMark
Option does not affect the throughput and therefore the user
experience.

Harm to the measurement: Alternate Marking measurements could be
harmed by routers altering the fields of the AltMark Option (e.g.
marking of the packets, FlowMonID) or by a malicious attacker adding
AltMark Option to the packets in order to consume the resources of
network devices and entities involved. As described above, the
source node is the only one that writes the Option Header while the
intermediate nodes and destination node only read it without
modifying the Option Header. But, for example, an on-path attacker
can modify the flags, whether intentionally or accidentally, or
insert deliberately a new option to the packet flow or delete the
option from the packet flow. The consequent effect could be to give
the appearance of loss or delay or invalidate the measurement by
modifying option identifiers, such as FlowMonID. The malicious
implication can be to cause actions from the network administrator
where an intervention is not necessary or to hide real issues in the
network. Since the measurement itself may be affected by network
nodes intentionally altering the bits of the AltMark Option or
injecting Options headers as a means for Denial of Service (DoS), the
Alternate Marking MUST be applied in the context of a controlled
domain, where the network nodes are locally administered and this
type of attack can be avoided.

The flow identifier (FlowMonID) composes the AltMark Option together
with the two marking bits (L and D). As explained in Section 5.3.1,
there is a chance of collision if the FlowMonID is set pseudo
randomly and a solution exist. In general this may not be a problem
and a low rate of ambiguous FlowMonIDs can be acceptable, since this
does not cause significant harm to the operators or their clients and
this harm may not justify the complications of avoiding it. But, for
large scale measurements where it is possible to monitor a big number
of flows, the disambiguation of the FlowMonID field is something to
take into account.

The privacy concerns also need to be analyzed even if the method only
relies on information contained in the Option Header without any
release of user data. Indeed, from a confidentiality perspective,
although AltMark Option does not contain user data, the metadata can
be used for network reconnaissance to compromise the privacy of users
by allowing attackers to collect information about network

performance and network paths. AltMark Option contains two kind of
metadata: the marking bits (L and D bits) and the flow identifier
(FlowMonID).

   The marking bits are the small information that is exchanged
   between the network nodes. Therefore, due to this intrinsic
   characteristic, network reconnaissance through passive
   eavesdropping on data-plane traffic is difficult. Indeed an
   attacker cannot gain information about network performance from a
   single monitoring point. The only way for an attacker can be to
   eavesdrop on multiple monitoring points at the same time, because
   they have to do the same kind of calculation and aggregation as
   Alternate Marking requires, and, after that, it can finally gain
   information about the network performance, but this is not
   immediate.

   The FlowMonID field is used in the AltMark Option as identifier of
   the monitored flow. It represents a more sensitive information
   for network reconnaissance and may allow a flow tracking type of
   attack because an attacker could collect information about network
   paths.

Furthermore, in a pervasive surveillance attack, the information that
can be derived over time is more. But the application of the
Alternate Marking to a controlled domain helps to mitigate all the
above aspects of privacy concerns.

At the management plane, attacks can be set up by misconfiguring or
by maliciously configuring AltMark Option. Thus, AltMark Option
configuration MUST be secured in a way that authenticates authorized
users and verifies the integrity of configuration procedures.
Solutions to ensure the integrity of AltMark Option are outside the
scope of this document.

As stated above, the precondition for the application of the
Alternate Marking is that it MUST be applied in specific controlled
domains, thus confining the potential attack vectors within the
network domain. [RFC8799] analyzes and discusses the trend towards
network behaviors that can be applied only within a limited domain.
This is due to the specific set of requirements especially related to
security, network management, policies and options supported which
may vary between such limited domains. A limited administrative
domain provides the network administrator with the means to select,
monitor and control the access to the network, making it a trusted
domain. In this regard it is expected to enforce policies at the
domain boundaries to filter both external packets with AltMark Option
entering the domain and internal packets with AltMark Option leaving
the domain. Therefore the trusted domain is unlikely subject to

hijacking of packets since packets with AltMark Option are processed
and used only within the controlled domain.

Additionally, it is to be noted that the AltMark Option is carried by
the Options Header and it may have some impact on the packet sizes
for the monitored flow and on the path MTU, since some packets might
exceed the MTU. However the relative small size (48 bit in total) of
these Option Headers and its application to a controlled domain help
to mitigate the problem.

It is worth mentioning that the security concerns may change based on
the specific deployment scenario and related threat analysis, which
can lead to specific security solutions that are beyond the scope of
this document. As an example, the AltMark Option can be used as Hop-
by-Hop or Destination Option and, in case of Destination Option,
multiple domains may be traversed by the AltMark Option that is not
confined to a single domain. In this case, the user, aware of the
kind of risks, may still want to use Alternate Marking for telemetry
and test purposes but the inter-domain links need to be secured
(e.g., by IPsec) in order to avoid external threats.

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]. Also,
the time, which is distributed to the network nodes through the time
protocol, is centrally taken from an external accurate time source,
such as an atomic clock or a GPS clock, and by attacking the time
source it can be possible to compromise the integrity of the
measurement as well. There are security measures that can be taken
to mitigate the GPS spoofing attacks and a network administrator
should certainly employ solutions to secure the network domain.

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 assignment 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, Stewart
Bryant, Christopher Wood, Yoshifumi Nishida, Tom Herbert, Stefano
Previdi, Brian Carpenter, Eric Vyncke, Greg Mirsky, 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.chen-pce-pcep-ifit]
           Chen, H., Yuan, H., Zhou, T., Li, W., Fioccola, G., and Y.
           Wang, "Path Computation Element Communication Protocol
           (PCEP) Extensions to Enable IFIT", draft-chen-pce-pcep-
           ifit-02 (work in progress), February 2021.

[I-D.fioccola-v6ops-ipv6-alt-mark]
           Fioccola, G., Velde, G. V. D., 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.fz-spring-srv6-alt-mark]
          Fioccola, G., Zhou, T., and M. Cociglio, "Segment Routing
          Header encapsulation for Alternate Marking Method", draft-
          fz-spring-srv6-alt-mark-00 (work in progress), January
          2021.

[I-D.hinden-6man-hbh-processing]
          Hinden, R. M. and G. Fairhurst, "IPv6 Hop-by-Hop Options
          Processing Procedures", draft-hinden-6man-hbh-
          processing-00 (work in progress), December 2020.

[I-D.ietf-idr-sr-policy-ifit]
          Qin, F., Yuan, H., Zhou, T., Fioccola, G., and Y. Wang,
          "BGP SR Policy Extensions to Enable IFIT", draft-ietf-idr-
          sr-policy-ifit-01 (work in progress), February 2021.

[I-D.peng-v6ops-hbh]
          Peng, S., Li, Z., Xie, C., Qin, Z., and G. Mishra,
          "Processing of the Hop-by-Hop Options Header", draft-peng-
          v6ops-hbh-03 (work in progress), January 2021.

[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>.

[RFC6438] Carpenter, B. and S. Amante, "Using the IPv6 Flow Label
          for Equal Cost Multipath Routing and Link Aggregation in
          Tunnels", RFC 6438, DOI 10.17487/RFC6438, November 2011,
          <https://www.rfc-editor.org/info/rfc6438>.

[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>.

   [RFC8799] Carpenter, B. and B. Liu, "Limited Domains and Internet
              Protocols", RFC 8799, DOI 10.17487/RFC8799, July 2020,
              <https://www.rfc-editor.org/info/rfc8799>.

   [RFC8889] Fioccola, G., Ed., Cociglio, M., Sapio, A., and R. Sisto,
              "Multipoint Alternate-Marking Method for Passive and
              Hybrid Performance Monitoring", RFC 8889,
              DOI 10.17487/RFC8889, August 2020,
              <https://www.rfc-editor.org/info/rfc8889>.

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