Multipoint Alternate Marking method for passive and hybrid performance monitoring

draft-fioccola-ippm-multipoint-alt-mark-02

London, Mar 2018, IETF 101

Giuseppe Fioccola (Telecom Italia)
Mauro Cociglio (Telecom Italia)
Amedeo Sapio (Politecnico di Torino)
Riccardo Sisto (Politecnico di Torino)

OAM RFC 8321 Applications

There are three documents that define how to use two bits long field to perform marking method in BIER, SFC and NVO3:

draft-ietf-bier-pmmm-oam

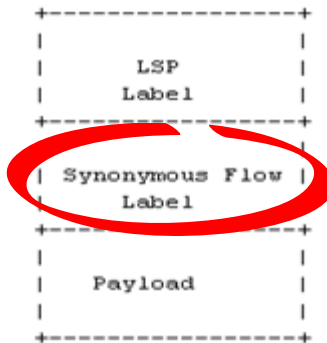draft-mirsky-sfc-pmamm

draft-fmm-nvo3-pm-alt-mark

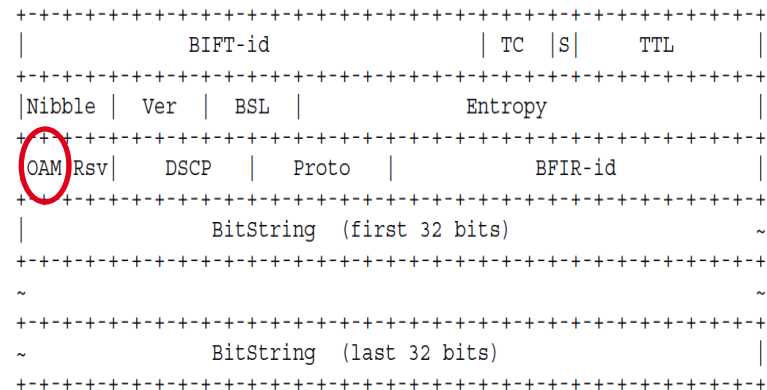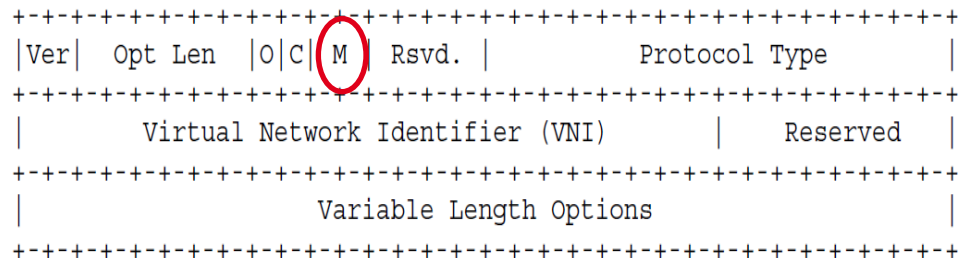Marking Method Application to MPLS RFC6374:

draft-ietf-mpls-rfc6374-sfl

Bier
```
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                 BIFT-id              | TC  |S|     TTL        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|Nibble |  Ver  |  BSL  |              Entropy                  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|OAM|Rsv|   DSCP    |  Proto  |           BFIR-id               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                 BitString  (first 32 bits)                  ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                                                              ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                 BitString  (last 32 bits)                    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

MPLS
Label
Stack
with SFL
```
+-------------------+
|                   |
|       LSP         |
|      Label        |
+-------------------+
|                   |
| Synonymous Flow   |
|      Label        |
+-------------------+
|                   |
|     Payload       |
|                   |
+-------------------+
```

Geneve
```
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|Ver|   Opt Len   |O|C| M | Rsvd. |        Protocol Type        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|        Virtual Network Identifier (VNI)       |   Reserved   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                   Variable Length Options                   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Nsh
```
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|Ver|O|R|    TTL    |    Length   | M |R|R|MD Type|    Proto    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
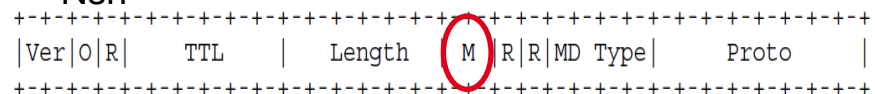
- There is also an alternate marking variation in QUIC:

draft-trammell-quic-spin

RFC 8321: What's next?

✓ There are some performance measurements applications where a lot of flows and nodes have to be monitored.

✓ The idea is to generalize and expand PM methodologies to measure any kind of unicast flows: in general multipoint-to-multipoint.

✓ A new framework can be introduced: **Multipoint Alternate Marking**

- It adds flexibility to PM because it can reduce the order of magnitude of the packet counters for large networks.
- It allows an SDN Orchestrator to supervise, control and manage PM in large networks.

Multipoint Alternate Marking: Cluster Packet Loss

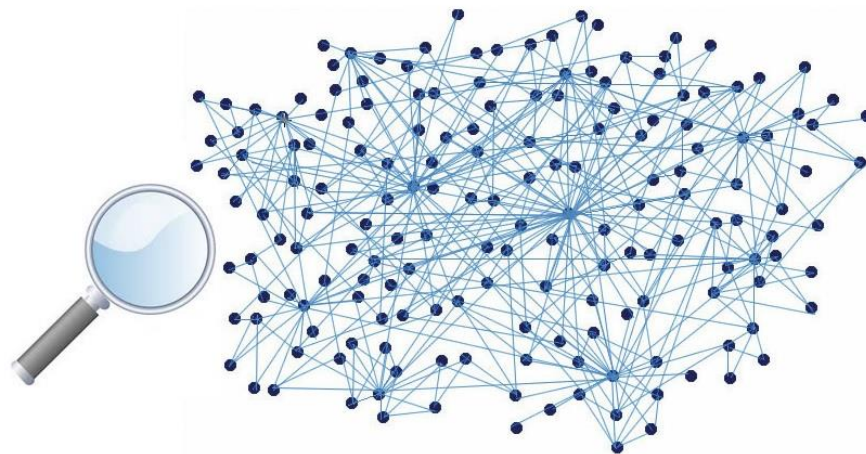The monitoring network can be considered as a whole or can be split in Clusters

Clusters are the smallest subnetworks, maintaining the packet loss property for each subnetwork

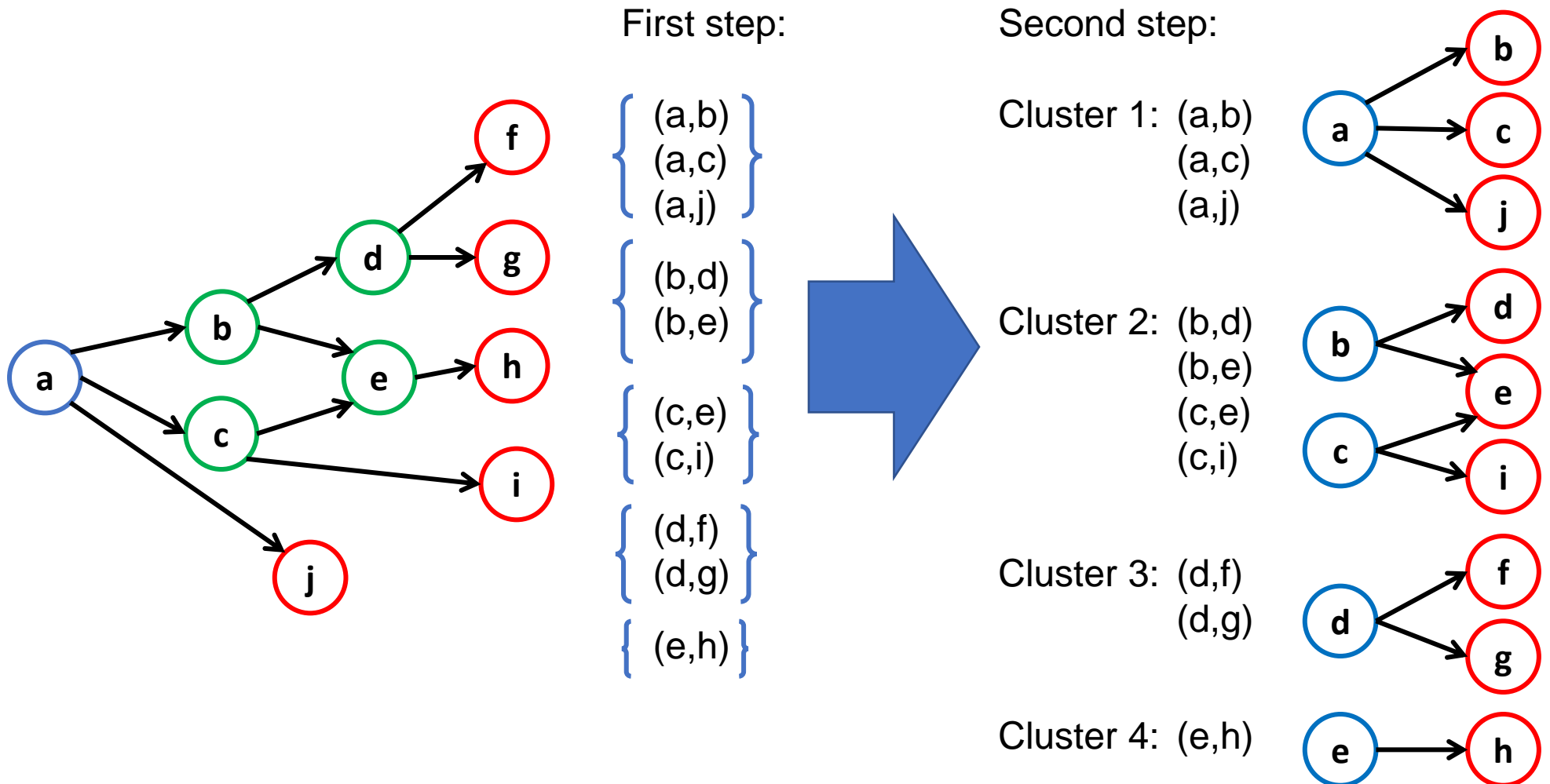A possible algorithm for Cluster partition is a two-step algorithm:

1. Group the links where there is the same starting node;

2. Join the grouped links with at least one ending node in common.

Clusters are, in general, group-to-group segments.

They can also be combined in new connected subnetworks at different levels depending on the detail we want to achieve.

A simple Algorithm for Cluster partition

First step:

$\{$ (a,b) (a,c) (a,j) $\}$

$\{$ (b,d) (b,e) $\}$

$\{$ (c,e) (c,i) $\}$

$\{$ (d,f) (d,g) $\}$

$\{$ (e,h) $\}$

Second step:

Cluster 1: (a,b) (a,c) (a,j)

Cluster 2: (b,d) (b,e) (c,e) (c,i)

Cluster 3: (d,f) (d,g)

Cluster 4: (e,h)

Multipoint Alternate Marking:
Delay measurement with RFC 5475 + RFC 8321

Mean delay and delay variation measurements can also be generalized to the case of multipoint flows and clusters.
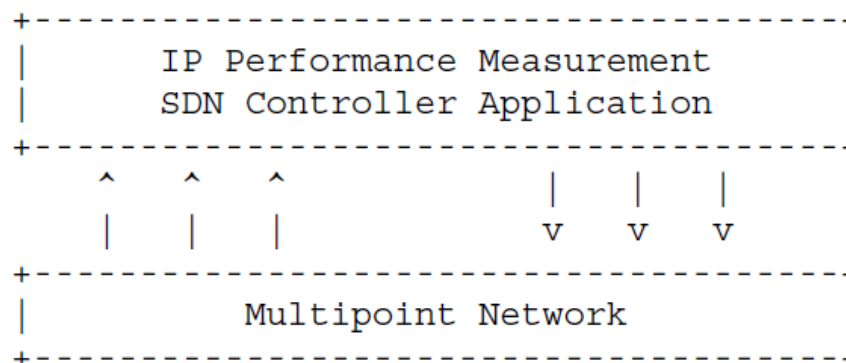
Single packets delay and delay variation cannot be performed for multipoint paths with double marking but can be done with Hash Selection described in RFC 5475: it gives a way to select the same packets in every monitoring point of a network.

There are two possible alternatives:

- Basic Hash: Alternate Marking splits the continuous flow in batches of packets and anchor the samples so this simplifies the correlation of the hashing packets along the path. But using Basic Hash, the number of samples depends on packet rate

- Dynamic Hash: In a marking period it is possible to select a number of samples «almost» constant with an iterative algorithm that statistically converges at the end of a marking period

Use Case: Multipoint Alternate Marking in an SDN scenario

- The IP Performance Measurement SDN Controller Application can orchestrate and calibrate the level of detail in network monitoring data by configuring measurement points roughly or meticulously to allow an optimized monitoring.

- Two ways to calibrate: Flow Filtering and Cluster Zooming

- Using Network Clustering approach it is possible to monitor a Multipoint Network. We can start without examining in depth, and in case there is packet loss or the delay is too high, the filtering criteria and clusters partition can be specified in different ways to perform a more detailed analysis.

- A FSM (Finite State Machine) can be programmed such that each state represents a composition of clusters (see draft-sambo-netmod-yang-fsm).

```
+-------------------------------------------+
|         IP Performance Measurement        |
|         SDN Controller Application        |
+-------------------------------------------+
   ^    ^    ^                |    |    |
   |    |    |                v    v    v
+-------------------------------------------+
|              Multipoint Network           |
+-------------------------------------------+
```

Summary and Next Steps

This document adds a new point of view to the alternate marking method:

- A Controller can calibrate Performance Measurements. It can start with the entire Network.

- In case of necessity, the filtering criteria could be specified more in order to perform a Cluster or a point-to-point flow detailed analysis.

- Hashing technique helps to perform a better (also per flow) delay and delay variation analysis.

Ask for working group adoption

Inputs and Comments always welcome