Multipoint Alternate Marking method for passive and hybrid performance monitoring

draft-fioccola-ippm-multipoint-alt-mark-04

Montreal, Jul 2018, IETF 102

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

Document changes: -02 to -03 to -04

We got a lot of comments after London meeting
- Thanks to Al, Qin, Tal, Rachel, Sudhin, Chongfeng

These latest revisions include a better description of some points:
- The scope of multipoint alternate marking with respect to the base alternate marking approach (RFC8321);
- An improvement to the description of the multipoint approach;
- A step-by-step example of the algorithm for Cluster partition;
- More details on delay measurement methods applicable for multipoint paths;
- The classification between multipoint paths basis and single packet basis delay measurement;
- A new Section on "Timing Aspects" has been included. This extends the corresponding Section of RFC8321.

Multipoint Alternate Marking: RFC8321 development

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

✓ The idea is to generalize and expand point-to-point RFC8321 to measure any kind of <u>unicast flows (not to multicast)</u>: 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.
- It enables dynamic performance monitoring and the possibility to set the desired performance measurement depending on the needs
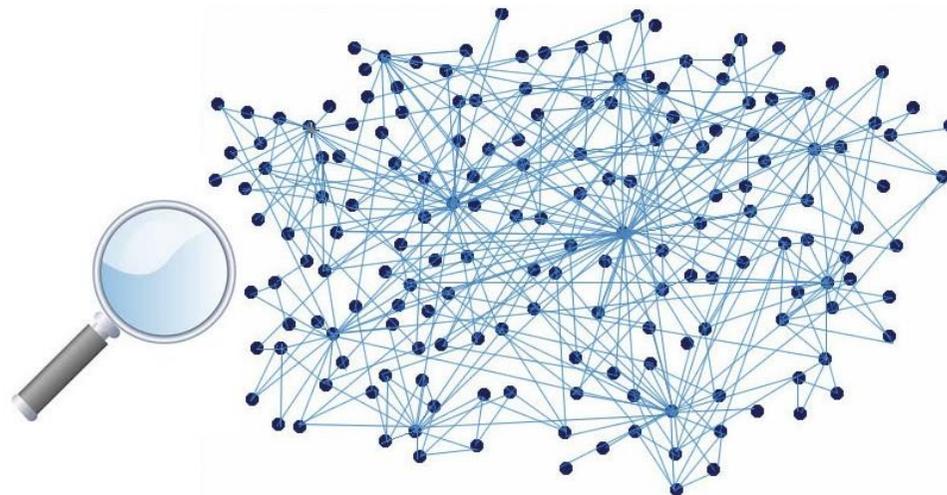
Multipoint Alternate Marking: Cluster Packet Loss

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

Clusters are the smallest subnetworks (in general group-to-group segments), maintaining the "packet loss property" for each subnetwork

They can also be combined in new connected subnetworks at different levels depending on the detail we want to achieve.
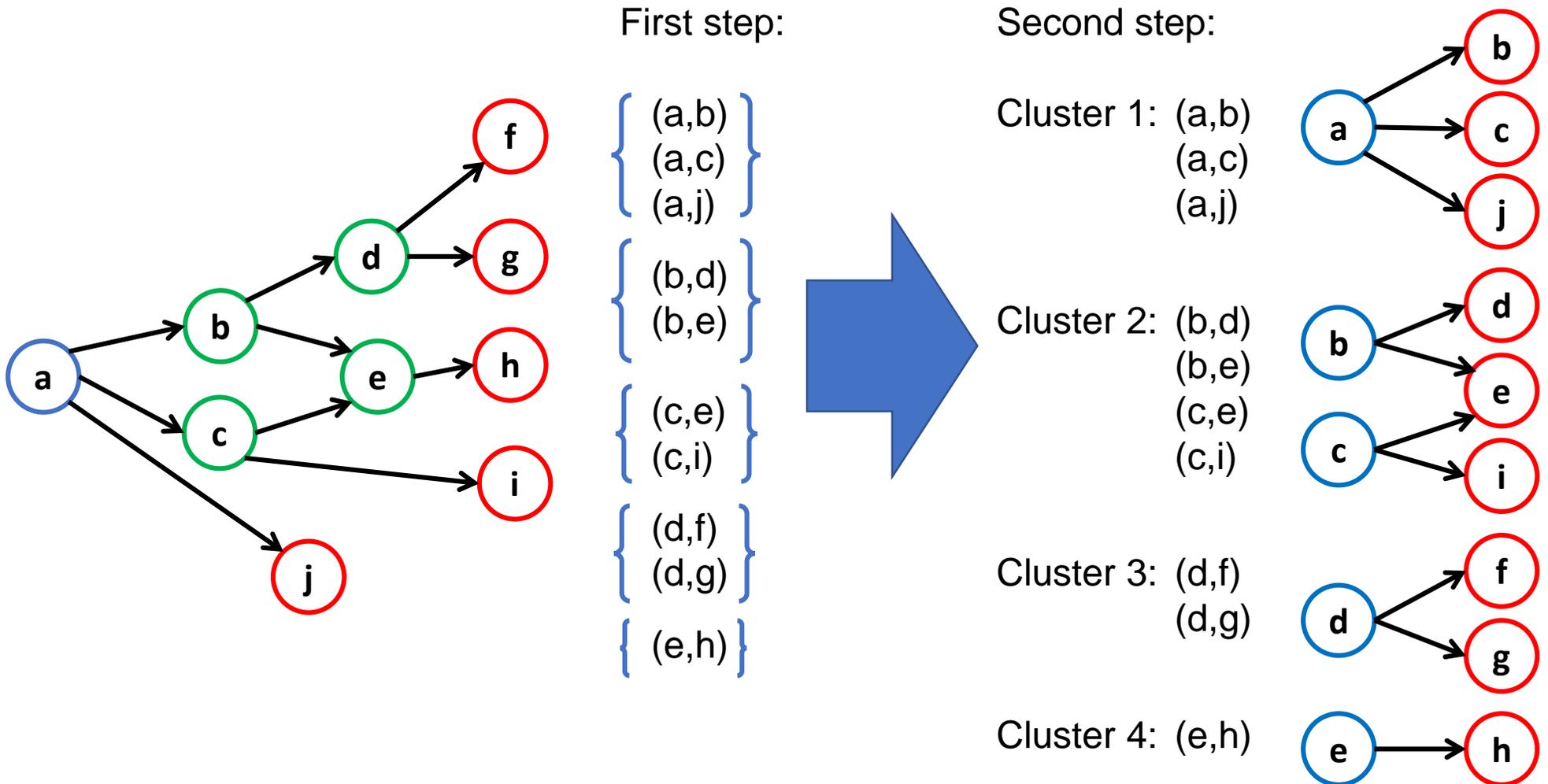
- *Without network clustering*, it is possible to apply alternate marking only for all the network or per single flow.

- *With network clustering*, it is possible to use the network clusters partition at different levels to perform the needed degree of detail.

A simple Algorithm for Cluster partition

A possible algorithm for Cluster partition is a two-step algorithm:
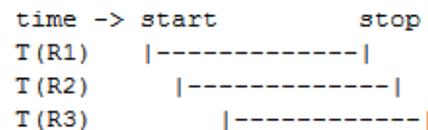1. Group the links where there is the same starting node;
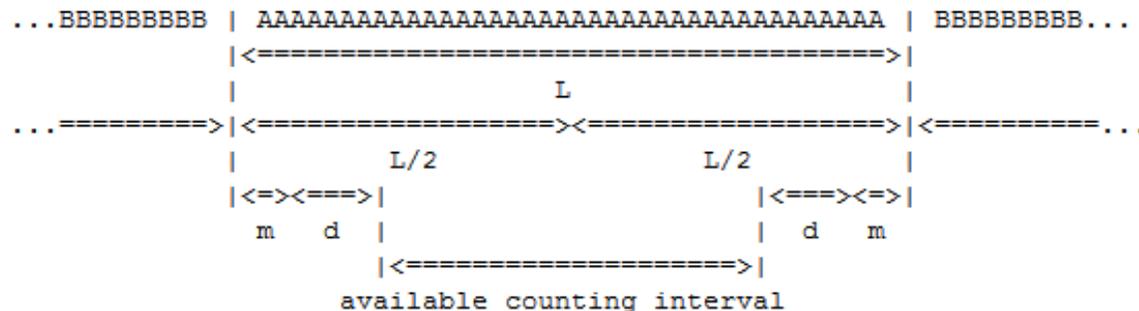2. Join the grouped links with at least one ending node in common.

Timing Aspects

This extends the corresponding Section in RFC8321

When we expand to multipoint-to-multipoint flows, we have to consider that all source nodes mark the traffic and the source measurement intervals can be of different lengths and with different offsets

```
time -> start            stop
T(R1)    |-------------|
T(R2)       |-------------|
T(R3)          |------------|
```

The misalignment between the marking source routers gives an additional constraint and the value of m is added to d (that already includes clock error and network delay).

```
...BBBBBBBBB | AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA | BBBBBBBBB...
            |<=====================================>|
            |                   L                   |
...=========>|<=================><=================>|<=========...
            |        L/2                 L/2        |
            |<=><===>|                      |<===><=>|
              m   d  |                      |  d   m
                     |<===================>|
                     available counting interval
```

The condition that must be satisfied is that the available counting interval is > 0, and that means: L - 2m - 2d > 0

Delay Measurement: multipoint paths basis and single packet basis

This classification has been introduced to distinguish between the two possible ways of measurement

- **multipoint path basis measurement**: the delay value is representative of an entire multipoint path (e.g. whole multipoint network, a cluster or a combination of clusters).
  - mean delay: The average latency can be measured as the difference between the mean timestamps of the sets of output and input nodes.

- **single packet basis measurement**: the multipoint path is used just to easily couple packets between inputs and output nodes of a multipoint path:
  - single marking based on the first/last packet of the interval would not work,
  - double marking/multiplexed marking works only for point-to-multipoint flows with limitations,
  - hashing is a more general solution (clusters simplify the coupling of the samples from a topological point of view, as well as marking method anchor the samples to a specific period and facilitates their correlation).
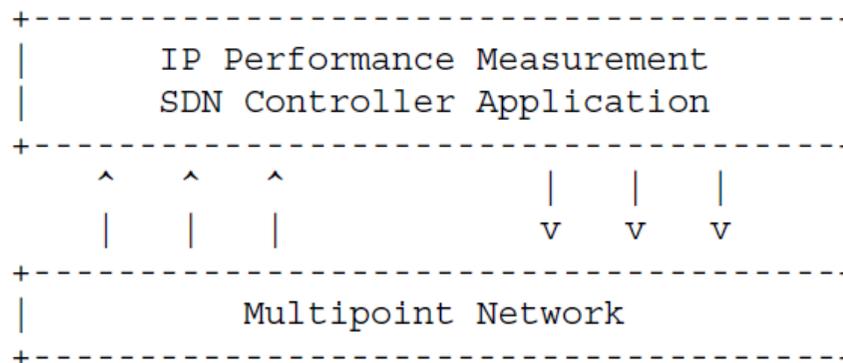
Delay measurement with RFC 5475 + RFC 8321

Single packet basis delay and delay variation measurement 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

A new point of view to the alternate marking method:

- A Controller can calibrate Performance Measurements.

- Dynamic Performance Measurement is introduced.

- Hashing technique helps to perform a better analysis.

Ask for working group adoption.

Inputs and Comments always welcome