

IPPM Working Group
Internet-Draft
Intended status: Experimental
Expires: December 31, 2017

G. Fioccola, Ed.
M. Cociglio, Ed.
Telecom Italia
A. Sapio, Ed.
R. Sisto, Ed.
Politecnico di Torino
June 29, 2017

Multipoint Alternate Marking method for passive and hybrid performance
monitoring
draft-fioccola-ippm-multipoint-alt-mark-00

Abstract

The Alternate Marking method, as presented in [\[I-D.ietf-ippm-alt-mark\]](#), can be applied only to point-to-point flows because it assumes that all the packets of the flow measured on one node are measured again by a single second node. This document aims to generalize and expand this methodology to measure any kind of unicast flows, whose packets can follow several different paths in the network, in wider terms a multipoint-to-multipoint network. For this reason the technique here described is called Multipoint Alternate Marking.

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 [RFC 2119](#) [[RFC2119](#)].

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 <http://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 31, 2017.

Internet-Draft

Multipoint AM

June 2017

Copyright Notice

Copyright (c) 2017 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to [BCP 78](#) and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1.	Introduction	2
2.	Flow classification	3
3.	Multipoint Performance Measurement	5
3.1.	Monitoring Network	6
3.2.	Multipoint Packet Loss	7
3.3.	Network Clustering	7
3.4.	Multipoint Delay and Jitter	10
3.4.1.	Single and Double Marking measurement	10
3.4.2.	Mean Delay and Jitter	10
3.4.3.	Hash method	10
4.	Examples	10
5.	Security Considerations	11
6.	Acknowledgements	11
7.	IANA Considerations	11
8.	References	11
8.1.	Normative References	11
8.2.	Informative References	11
	Authors' Addresses	12

[1.](#) Introduction

The Alternate Marking methodology described in [[I-D.ietf-ippm-alt-mark](#)] has the property to synchronize measurements in different points maintaining the coherence of the counters. So it is possible to show what is happening in every marking period for each monitored flow. The monitoring parameters are the packet

counter and timestamps of a flow for each marking period.

There are some applications of the alternate marking method where there are a lot of monitored flows and nodes.

For instance, by considering n measurement points and n monitored flows, the order of magnitude of the packet counters for each time interval is $n \times n \times 2$ (1 per color).

Multipoint Alternate Marking aims to reduce this value and makes the performance monitoring more flexible in case a detailed analysis is not needed. It can be applied only to unicast flows.

In some circumstances it is possible to monitor a Multipoint Network by analyzing the Network Clustering, without examining in depth. In case there is packet loss or the delay is too high the filtering criteria could be specified more in order to perform a per flow detailed analysis, as described in [[I-D.ietf-ippm-alt-mark](#)].

An application could be the software defined network (SDN) paradigm where the SDN Controllers are the brains of the network and can manage flow control to the switches and routers and, in the same way, can calibrate the performance measurements depending on the necessity.

[2.](#) Flow classification

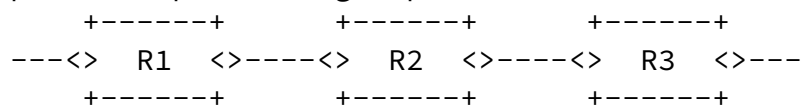
A flow is identified by all the packets having a set of common characteristics. This definition is inspired by [RFC 7011](#) [[RFC7011](#)].

As an example, by considering a flow as all the packets sharing the same source IP address or the same destination IP address, it is easy to understand that the resulting pattern will not be a point-to-point connection, but a point-to-multipoint or multipoint-to-point connection.

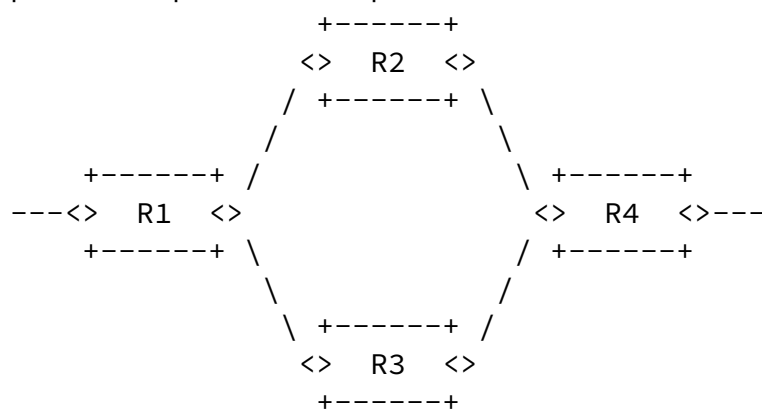
In general a flow can be defined by a set of selection rules used to match a subset of the packets processed by the network device. These rules specify a set of headers fields (Identification Fields) and the relative values that must be found in matching packets.

The choice of the identification fields directly affects the type of paths that the flow would follow in the network. In fact, it is possible to relate a set of identification fields with the pattern of the resulting graphs, as listed in Figure 1.

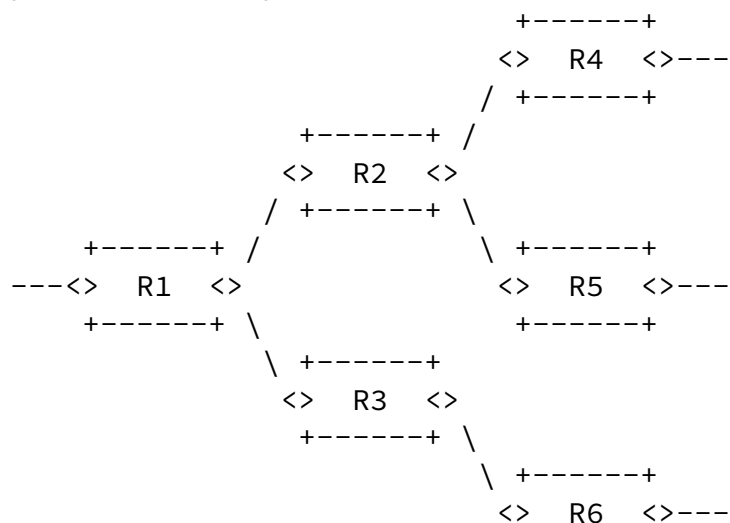
point-to-point single path



point-to-point multipath

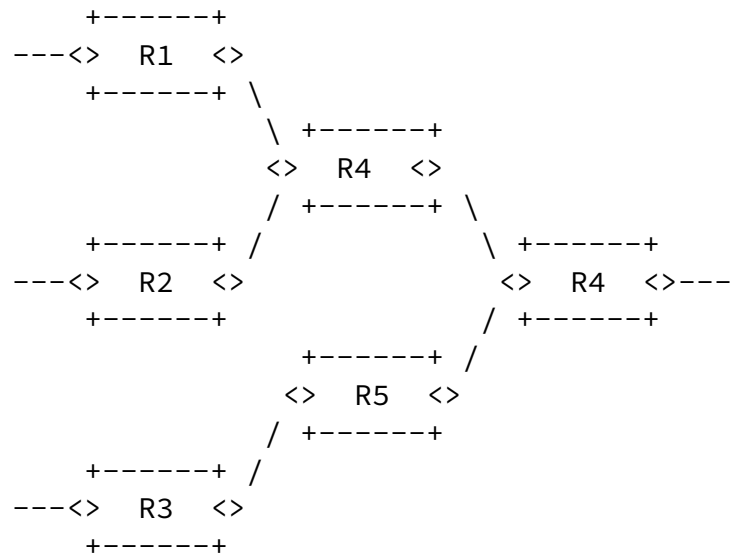


point-to-multipoint



+-----+

multipoint-to-point



multipoint-to-multipoint

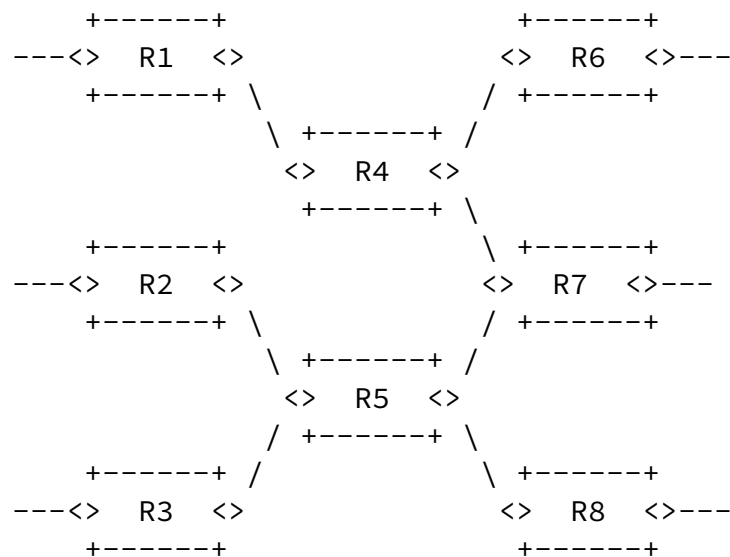


Figure 1: Flow classification

A TCP 5-tuple usually identifies flows following either a single path or a point-to-point multipath (in case of load balancing). On the

contrary, a single source address selects flows following a point-to-multipoint, while a multipoint-to-point can be the result of a matching on a single destination address. In case a selection rule and its reverse are used for bidirectional measurements, they can correspond to a point-to-multipoint in one direction and a multipoint-to-point in the opposite direction.

In this way the flows to be monitored are selected into the monitoring points using packet selection rules, that can also change the pattern of the monitored network.

The alternate marking method is applicable only to a single path (and partially to a one-to-one multipath), so the extension proposed in this document is suitable also for the most general case of multipoint-to-multipoint, which embraces all the other patterns of Figure 1.

[3.](#) Multipoint Performance Measurement

By Using the "traditional" alternate marking method only point-to-point paths can be monitored. To have an IP (TCP/UDP) flow that follows a point-to-point path we have to define, with a specific value, 5 identification fields (IP Source, IP Destination, Transport Protocol, Source Port, Destination Port).

Multipoint Alternate Marking enables the performance measurement for multipoint flows selected by identification fields without any constraints (even the entire network production traffic). It is also possible to use multiple marking points for the same monitored flow.

[3.1.](#) Monitoring Network

The Monitoring Network is deduced from the Production Network, by identifying the nodes of the graph, that are the measurement points, and the links, that are the connections between measurement points.

So a model of the monitoring network can be built according to the alternate marking method: the monitored interfaces and links are identified. Only the measurement points and links where the traffic has flowed have to be represented in the graph.

The following figure shows a simple example of a Monitoring Network graph:

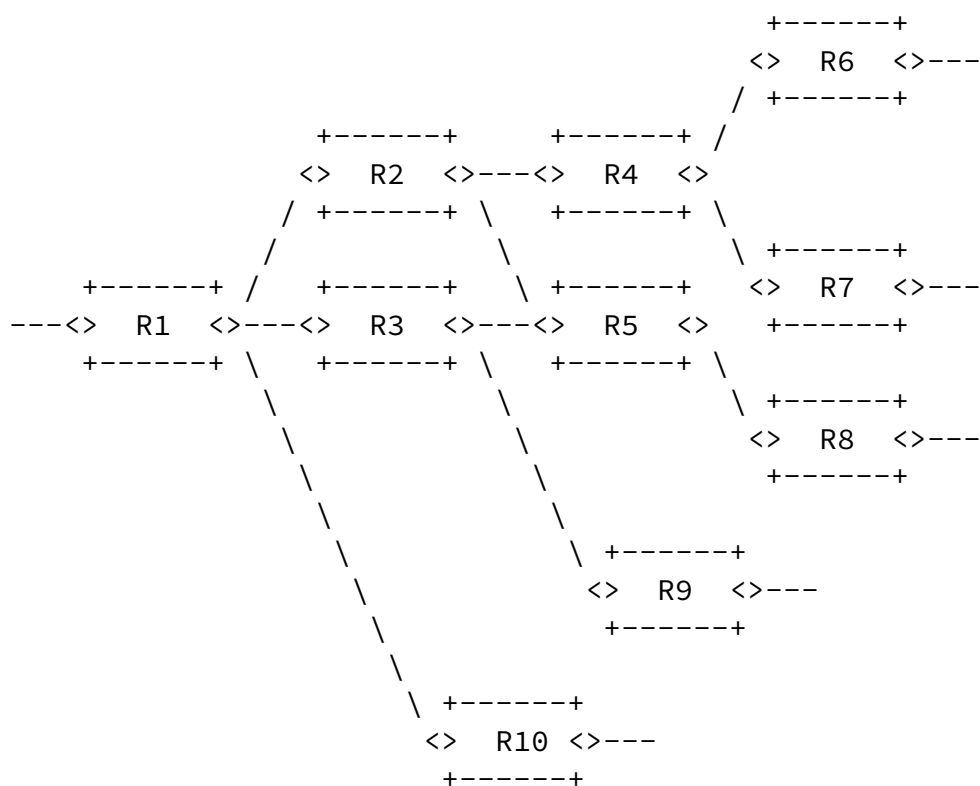


Figure 2: Monitoring Network

Each monitoring point is characterized by the packet counter that refers only to a marking period of the monitored flow.

The same is applicable also for the delay but it will be described in the following sections.

[3.2.](#) Multipoint Packet Loss

Since all the packets of the considered flow leaving the network have previously entered the network, the number of packets counted by all the input nodes is always greater or equal than the number of packets

counted by all the output nodes.

And in case of no packet loss occurring in the marking period, if all the input and output points of the network domain to be monitored are measurement points, the number of packets is the same on all the ingress interfaces and on all the egress interfaces. The intermediate measurement points have only the task to split the measurement.

It is possible to define the Network Packet Loss (for 1 flow, for 1 period): <<In a packet network, the number of lost packets is the number of packets counted by the input nodes minus the number of packets counted by the output nodes>>. This is true for every packet flow in each marking period.

The Monitored Network Packet Loss with n input nodes and m output nodes is given by:

$$PL = (PI_1 + PI_2 + \dots + PI_n) - (PO_1 + PO_2 + \dots + PO_m)$$

where:

PL is the Network Packet Loss (number of lost packets)

PI_i is the Number of packets flowed through the i -th Input node in this period

PO_j is the Number of packets flowed through the j -th Output node in this period

[3.3.](#) Network Clustering

The previous Equation can determine the number of packets lost globally in the monitored network, exploiting only the data provided by the counters in the input and output nodes.

In addition it is also possible to leverage the data provided by the other counters in the network to converge on the smallest identifiable subnetworks where the losses occur. These subnetworks are named Clusters.

A Cluster is a subnetwork of the entire Monitoring Network graph that

still satisfies the packet loss equation where PL in this case is the number of packets lost in the Cluster.

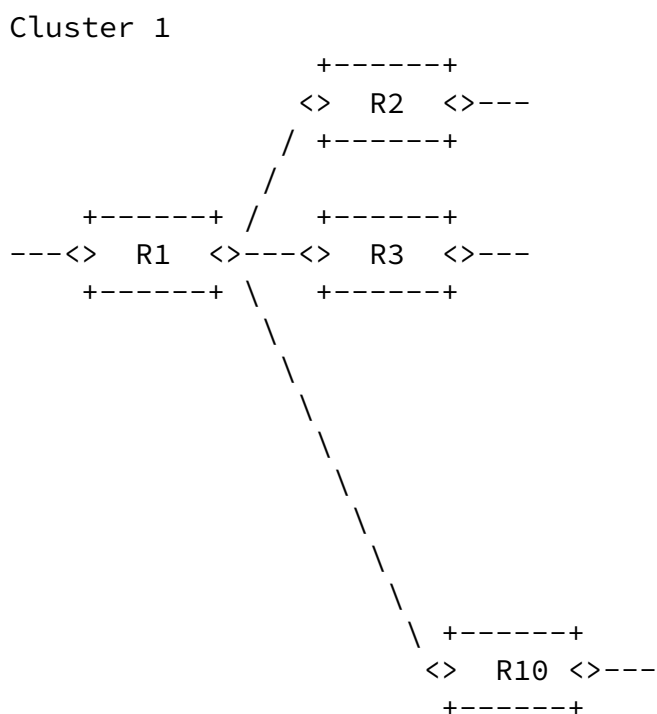
For this reason a Cluster should contain all the arcs emanating from its input nodes and all the arcs terminating at its output nodes. This ensures that we can count all the packets (and only those) exiting an input node again at the output node, whatever path they follow.

In a completely monitored network (a network where every network interface is monitored), each network device corresponds to a Cluster and each physical link corresponds to two Clusters (one for each direction).

Clusters can have different sizes depending on flow filtering criteria adopted.

Moreover, sometimes Clusters can be simplified; for example when two monitored interfaces are divided by a single router (one is the input interface and the other is the output interface and the router has only these two interfaces), instead of counting exactly twice, upon entering and leaving, in this case it is possible to consider a single measurement point.

In our monitoring network graph example it is possible to identify 4 Clusters:



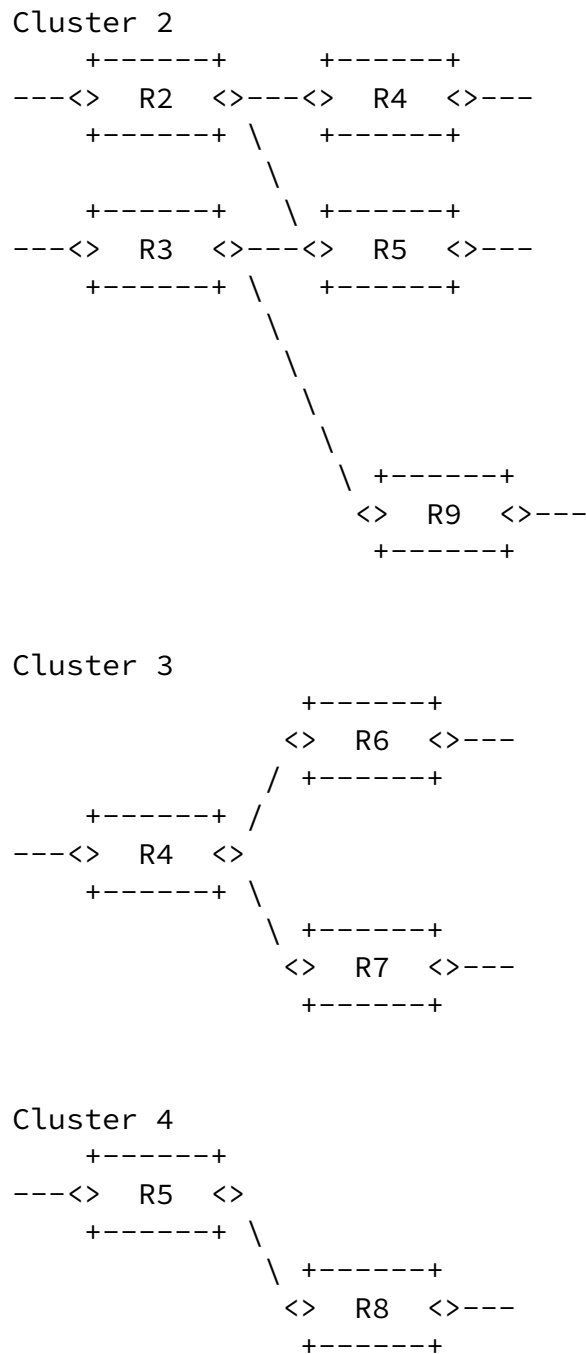


Figure 3: Clusters example

There are Clusters with more than 2 nodes and two-nodes Clusters. In the two-nodes Clusters the loss is on the link (Cluster 4). In more-than-2-nodes Clusters the loss is on the Cluster but we cannot know in which link (Cluster 1, 2, 3).

Obviously, by combining some Clusters in a new connected subnetwork

(called Super Cluster) the Packet Loss Rule is still true.

[3.4.](#) Multipoint Delay and Jitter

The same line of reasoning can be applied to Delay and Jitter.

[3.4.1.](#) Single and Double Marking measurement

Delay and Jitter measurements relative to a picked packet (both single and double marked) cannot be performed in the Multipoint scenario, since they would not be representative of the entire flow. The packets can follow different paths with various delays and in general it is very difficult to recognize a marked packet in a multipoint-to-multipoint path.

[3.4.2.](#) Mean Delay and Jitter

Mean delay and jitter measurements can also be generalized to the case of multipoint flows. It is possible to compute the average one-way delay of packets, in one block, in a cluster or in the entire monitored network.

The average latency can be measured as the difference between the weighted averages of the mean timestamps of the sets of output and input nodes.

[3.4.3.](#) Hash method

[RFC 5475](#) [[RFC5475](#)] introduces sampling and filtering techniques for IP Packet Selection.

Hash selection methodologies can work in a multipoint-to-multipoint path and can be used both coupled to mean delay or stand alone.

[4.](#) Examples

There are three application fields where it may be useful to take into consideration the Multipoint Alternate Marking:

- o VPN: The IP traffic is selected on IP source basis in both directions. At the end point WAN interface all the output traffic

is counted in a single flow. The input traffic is composed by all the other flows aggregated for source address. So, by considering n end-points, the monitored flows are n (each flow with 1 ingress point and $(n-1)$ egress points) instead of $n*(n-1)$ flows (each flow, with 1 ingress point and 1 egress point);

- o Mobile Backhaul: LTE traffic is selected, in the Up direction, by the ENodeB source address and, in Down direction, by the ENodeB destination address because the packets are sent from the Mobile

Packet Core to the ENodeB. So the monitored flow is only one per ENodeB in both directions;

- o OTT(Over The Top) services: The traffic is selected, in the Down direction by the source addresses of the packets sent by OTT Servers. In the opposite direction (Up) by the destination IP addresses of the same Servers. So the monitoring is based on a single flow per OTT Servers in both directions.

[5.](#) Security Considerations

tbc

[6.](#) Acknowledgements

tbc

[7.](#) IANA Considerations

tbc

[8.](#) References

[8.1.](#) Normative References

[I-D.ietf-ippm-alt-mark]

Fioccola, G., Capello, A., Cociglio, M., Castaldelli, L., Chen, M., Zheng, L., Mirsky, G., and T. Mizrahi, "Alternate Marking method for passive and hybrid performance monitoring", [draft-ietf-ippm-alt-mark-05](#) (work in progress), June 2017.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", [BCP 14](#), [RFC 2119](#), DOI 10.17487/RFC2119, March 1997, <<http://www.rfc-editor.org/info/rfc2119>>.

8.2. Informative References

[RFC5475] Zseby, T., Molina, M., Duffield, N., Niccolini, S., and F. Raspall, "Sampling and Filtering Techniques for IP Packet Selection", [RFC 5475](#), DOI 10.17487/RFC5475, March 2009, <<http://www.rfc-editor.org/info/rfc5475>>.

Fioccola, et al. Expires December 31, 2017 [Page 11]

Internet-Draft Multipoint AM June 2017

[RFC7011] Claise, B., Ed., Trammell, B., Ed., and P. Aitken, "Specification of the IP Flow Information Export (IPFIX) Protocol for the Exchange of Flow Information", STD 77, [RFC 7011](#), DOI 10.17487/RFC7011, September 2013, <<http://www.rfc-editor.org/info/rfc7011>>.

Authors' Addresses

Giuseppe Fioccola (editor)
Telecom Italia
Via Reiss Romoli, 274
Torino 10148
Italy

Email: giuseppe.fioccola@telecomitalia.it

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

Email: mauro.cociglio@telecomitalia.it

Amedeo Sapio (editor)
Politecnico di Torino
Corso Duca degli Abruzzi, 24
Torino 10129
Italy

Email: amedeo.sapio@polito.it

Riccardo Sisto (editor)
Politecnico di Torino
Corso Duca degli Abruzzi, 24
Torino 10129
Italy

Email: riccardo.sisto@polito.it