MEDIA AWARE NETWORK ELEMENT DATA PLANE PERFORMANCE EVALUATION

Marius VOCHIN[1], Eugen BORCOCI[2], Gustavo CARNEIRO[3]

ALICANTE is a FP7 research project proposing an architecture in which, the CAN layer offers enhanced content-aware connectivity services to the upper layers for the efficient transport of media services. ALICANTE integrates, in an efficient way, and in the same tree, three separate transport mechanisms for multicast traffic—IP multicast, inter-domain tunnelling, and peer-to-peer - with QoS and dynamic adaptation support, which is something unique, not found in related solutions. It also allows independent domains to collaborate in forming a single inter-domain tree, for incremental deployment. While it uses legacy QoS technologies, such as core IP/MPLS/Diffserv in the core routers, it proposes a new edge router, called MANE (Media Aware Network Element), CAN capable. This paper evaluates a modular MANE implementation solution used to implement flow classification, MPLS encapsulation and decapsulation, separation between VCANs, and QoS enforcement with Linux TC (traffic control) mechanism. Based on data plane performance measurements, it is shown that the implementation imposes minor overheads over existing routing infrastructure and scalability is achievable.

Keywords: content-aware networking, network aware applications, quality of services, multimedia distribution, Future Internet, multicast, media-aware network element, scalability.

1. Introduction

Emerging FI architectures should be able to fully support and facilitate all kinds of current and future media-oriented applications. ALICANTE is a FP7 research project proposing an architecture in which, service distribution is achieved through a mesh of intelligent media-centric home gateways ("Home-Boxes"), which utilize novel Content-Aware Network (CAN) overlays, developed on top of the existing transport infrastructures. For many emerging applications, multicast is the most natural and efficient media distribution method.

[1] Eng., Telecommunication Dept., University POLITEHNICA of Bucharest, Romania, e-mail: marius.vochin@elcom.pub.ro
[2] Prof., Telecommunication Dept., University POLITEHNICA of Bucharest, Romania, e-mail: eugen.borcoci@elcom.pub.ro
[3] PhD eng., Instituto de Engenharia e Sistemas de Computadores do PORTO, e-mail: gjc@inescporto.pt

A CAN & network multicast architecture and design are proposed in [1]. The main design decisions and trade-offs listed, and a multicast architecture is proposed, from the point of view of CAN management, intra-domain multicast, inter-domain multicast, peer-to-peer multicast, and finally the Multicast Bridge, which is the data plane of the multicast subsystem.

It is observed that ALICANTE integrates, in an efficient way, three separate transport mechanisms for multicast traffic; in the same tree, one can have IP multicast, inter-domain tunnelling, and integration with HB peer-to-peer. The integration of all the different transport methods in the same end-to-end multicast tree is something unique in ALICANTE. Furthermore, QoS support is integrated with the layered coding scheme adopted by ALICANTE, allowing in-network elements to drop low priority media layers first, when network problems are detected. For inter-domain multicast, ALICANTE employs an innovative technique to allow IP multicast packets to cross inter-domain links without necessarily always resorting to encapsulation, leveraging content-awareness features built into the ALICANTE architecture. The ALICANTE control plane allows independent domains to collaborate in forming a single inter-domain tree, allowing incremental deployment [2].

The work of this paper is a part of the effort inside of an European FP7 ICT research project, “MediA Ecosystem Deployment Through Ubiquitous Content-Aware Network Environments”, ALICANTE, [3][4]. An innovative architecture is proposed, for a “Networked Media Ecosystem”, supporting flexible cooperation between providers, operators, and users. Three interworking environments are defined: Network Environment (NE) including Network Providers, Service Environment (SE) including Content and Service Providers and User Environment (UE) including all End-Users [4]. The architecture validation and implementation were deployed in a large-scale international pilot.

The main objectives of this paper are to validate the functional capabilities and to measure the forwarding performance and other relevant scalability factors of a MANE implementation in a real testbed.

The paper is organized as follows: Section 2 presents some of the related work existent in the field. The ALICANTE multicast bridge architecture and its functionality are defined in Section 3. Section 4 presents multicast data plane performance and scalability related measurements. Conclusions, open issues, and future work are presented in Section 5.

2. Related work

The multicast transport service has received more attention in the last years, in the context of increasing of group communication needs and also of real time flows and content/media distribution to large groups of users like IPTV, Video on demand (VoD), peer-to-peer (P2P). The traditional IP multicast (IPM) despite its two decade age [5], is not however globally deployed [6] [7] [8] due to problems related to group management, needed router capabilities, inter-domain issues and QoS problems. On the other hand, Overlay Multicast (OM) [9] [10] has received increased attention in the last decade, including tree based and also P2P solutions. OM has lower efficiency and speed but it is fast implementable (it is not relying on network layer multicast capabilities). A hybrid approach (IPM + OM) can be an attractive trade-off, both in terms of scalability, efficiency and flexibility. One can benefit of intra-domain IPM where it exists, and use OM outside the IP multicast area. Therefore, in ALICANTE, a solution for a multicast hybrid infrastructure is proposed. The multicast services are offered to the upper layers by the VCANs. The solution is based on hybrid multicast (H-Mcast) framework, QoS capable, where the IP intra-domain multicast (if available) is combined with inter-domain overlay multicast.

3. Multicast Bridge

The multicast data plane module in the MANE can be seen as a box which receives a packet as input and re-transmits the same packet a number of times to multiple destinations. The complexity with CAN layer multicast happens at the CAN Manager level; the MANE receives instructions of how to multiply and forward multicast traffic. These instructions are issued by the CAN Manager, but delivered by the IntraNRM module to the MANE, which then uses them to build a multicast forwarding table.

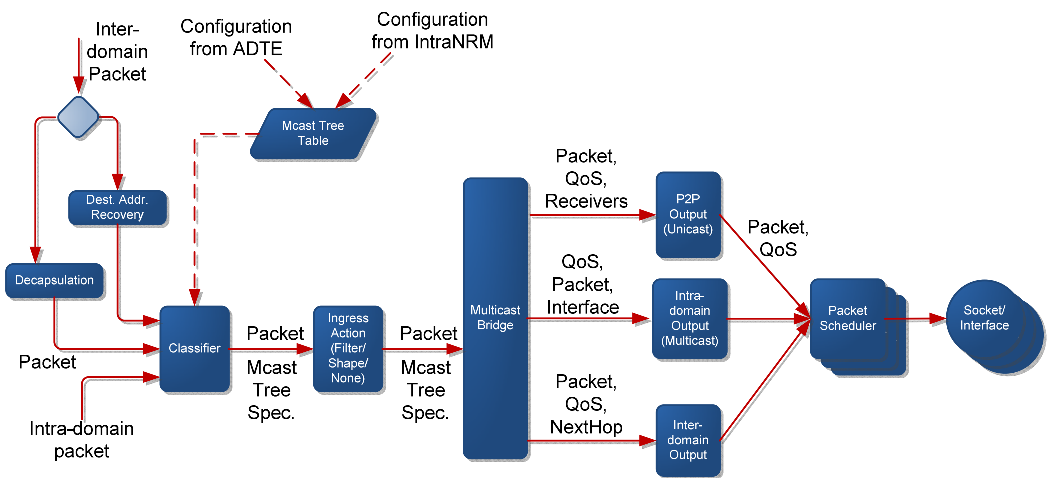


Fig. 1 MANE Multicast Bridge architecture

Fig. 1 shows the general architecture of the multicast data plane of the MANE. It can receive packets from either the local domain (we call that an intra-domain packet) or from an inter-domain link (inter-domain packet). In the latter case, the packet to be processed may be encapsulated in a UDP tunnel; therefore, the encapsulation must be removed before the packet is further processed. Once the IP multicast packet is obtained, a classification function is applied that uses the source and destination IP addresses tuple as a key to a lookup table, the SLS/Forwarding Table. The use of the source and destination IP addresses as key can be explained by the fact that, in ALICANTE, the multicast trees are SSM, i.e., the source IP together with the IP multicast group (destination) uniquely identify a multicast tree. As a result of the table lookup, the QoS specification that applies to the multicast tree is obtained, and it determines the action to apply to packets as they arrive. Typically, intra-domain packets are filtered or shaped to ensure that they do not exceed the traffic envelope indicated by the SLS negotiated with the SP. The action to apply, and respective parameters, is therefore configurable by the SP. The next stage of packet processing, if it survives the filtering or after it is delayed by shaping, is to enter the Multicast Bridge, which is the module responsible for retransmitting the packet a number of times, as dictated by the multicast tree topology. There are three main output functions that can be invoked:

1. Intra-domain Output: this function consists in transmitting the multicast packet through a MANE interface that is connect to a multicast-enabled domain;

2. Inter-domain Output: this function consists in transmitting the multicast packet through an inter-domain link towards a peer MANE in a different domain. Thus, UDP encapsulation is required;

3. P2P Output: this output function is responsible for transmitting the packet a number of times as unicast for the HBs that requested it (except the ones receiving via P2P).

In most cases, some form of packet scheduling is required before each packet is really transmitted. For this reason, a Packet Scheduler module instance exists associated with each output link, to ensure that, in case of network congestion, packets with lower priority are dropped before the packets with higher priority are considered for dropping.

4. Experiments

In Fig. 2, on U.P.B. premises we built a testbed topology comprising three IP domains containing 5 HBs, eight MANEs, and seven MPLS core routers:

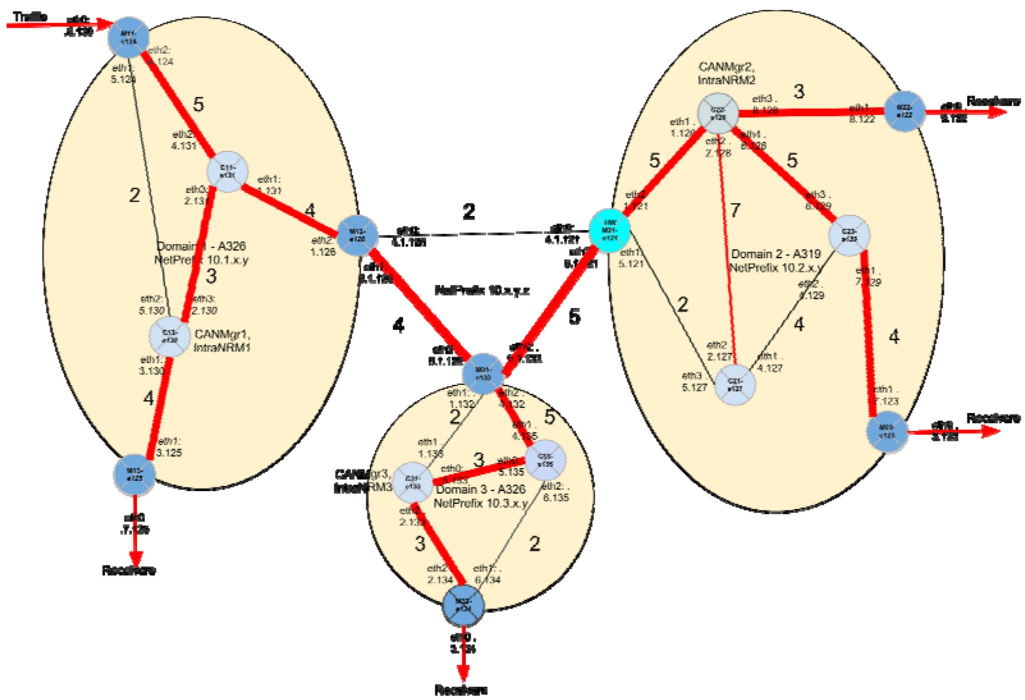


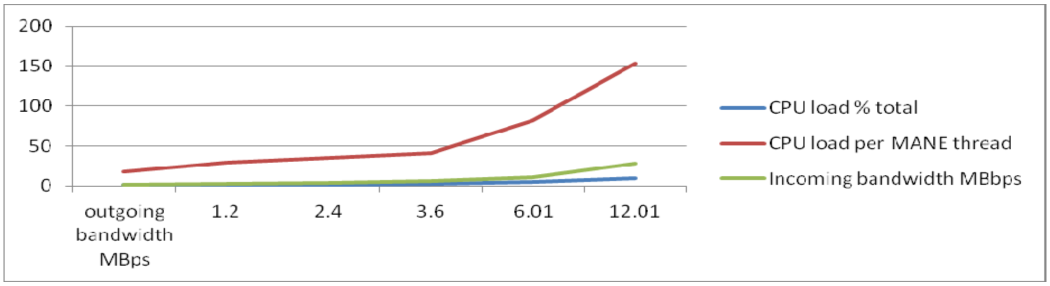
Fig. 2 Multicast validation tree in U.P.B. testbed

All the elements in the topology run Debian Squeeze x64 on quad CPU Intel Xeon W3520, hyper-threading enabled, 4GB RAM memory, Gigabit Ethernet Intel Pro/1000 82574L and Realtek 8111 PCIx Ethernet cards. In [11] we measured performance in four different routing configurations: using standard IP, using kernel MANE MPLS, then using user MANE IP and MPLS. For the RTT tests, we used ping with large and small packets. The MANE implementations bring a minor increase in end to end transit time and a slight increase in the standard deviation of RTT. For throughput measurements, test traffic was generated and measured in user-space using the Iperf traffic generator. The data rates achieved with the MANE implementations are comparable to plain Linux data rates, but packets per second throughput offered by kernel MANE implementation is slightly less. User MANE implementation offers considerable less throughput; part of this difference can be accounted by the current implementation of all modules in user space.

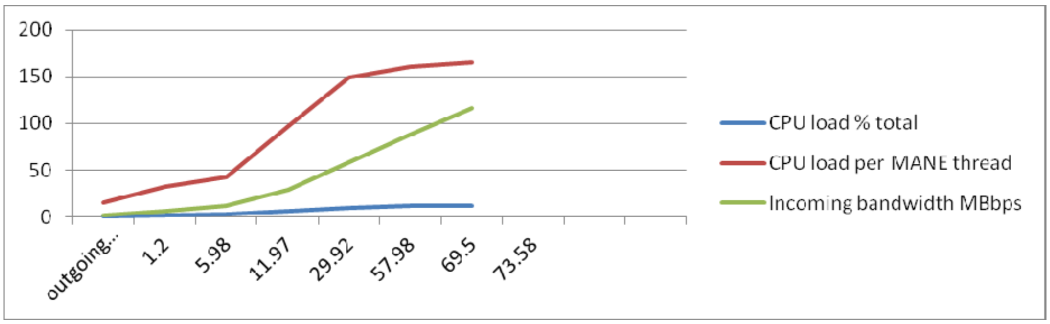
MANE multicast bridge data plane performance was tested in U.P.B. physical testbed. MANE software was run on Linux machine e126 and multicast test traffic was sent with Ixia IXExplorer traffic generator, with one hybrid multicast tree installed on all three testbed domains.

The following 4 tests were conducted:

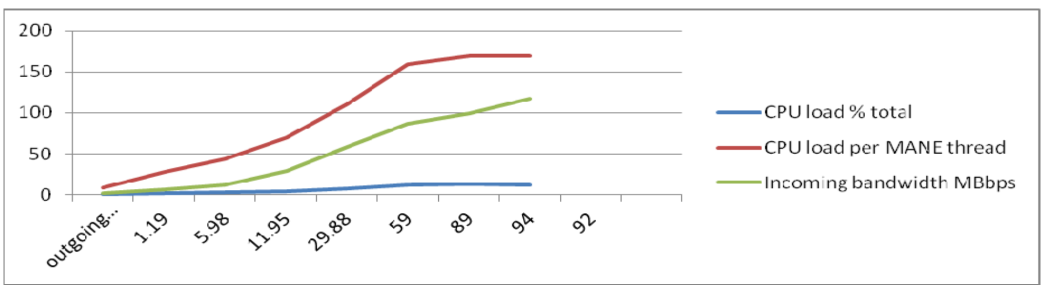
1. Traffic was generated with fixed packet size of 500b, and a bandwidth of 1,2,3,5,10,25,50,75,100% of maximum IXExplorer capacity.



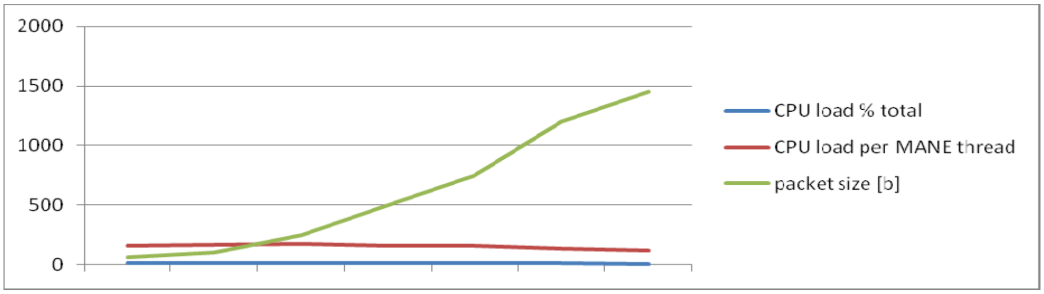
2. Traffic was generated with fixed packet size of 1000b, and a bandwidth of 1,5,10,25,50,75,100% of maximum IXExplorer capacity.



3. Traffic was generated with fixed packet size of 1450b, and a bandwidth of 1,5,10,25,50,75,85,100% of maximum IXExplorer capacity.



4. Traffic was generated with an incoming packet rate of 50% of maximum IXExplorer capacity.



It can be seen that for small packet sizes the throughput achieved is less than 12MBps, but the throughput increases with the packet size, almost achieving gigabit speed for packets close to the MTU. The Bridge forwarding capability is around 81000 pps with no loss.

5. Conclusions and future work

The presented media aware network architecture, aims at virtualizing network resources for the purpose of offering higher QoS to media flows. The MANE (Media Aware Network Element) is an edge router that has a central role in implementing the separation between networks, by classifying incoming traffic and distributing it to appropriate MPLS paths inside each domain.

We implemented the MANE using off-the-shelf hardware, using Click modular router to interface the components: classification, routing, adaptation, multicast, MPLS FEC association, encapsulation and decapsulation [12]. Our preliminary implementation shows a modest increase in processing overhead when compared with traditional IP/MPLS processing. It was shown that a MANE system can treat 1000 flows without introducing an important performance penalty [11]. Multicast bridge performance was also investigated, showing no significant penalties for medium to large data packets.

Forwarding performance could be improved with a multi-threading implementation of the software. It can be seen that total CPU load is relatively low, even when the core that runs the MANE software is used at full capacity.

In the future, we aim at developing the MANE in two directions: adding deep packet inspection functionality, to assist in classification of traffic not yet associated with a VCAN, and integration with high speed network processing cards, to target operation at line speed.

Performance improvements regarding QoS would be possible by using an hierarchical filter structure that permits hashing. Both these directions aim at creating a MANE that can be deployed in the field by service providers.

Acknowledgment

This work was supported in part by project POSDRU/89/1.5/S/62557 and in part by the EC in the context of the ALICANTE project (FP7-ICT-248652).

R E F E R E N C E S

[1] *Radu Iorga, António Pinto, Eugen Borcoci, Gustavo Carneiro, Radu Miruta, Tania Calcada*, “Management Driven Hybrid Multicast Framework for Content Aware Networks”, IEEE Communications Magazine 52(1): 2014, pp158-165

[2] *Eugen Borcoci, Gustavo Carneiro, Radu Iorga*, “Hybrid Multicast Management in a Content Aware Multidomain Network”, The Third International Conference on Advances in Future Internet, AFIN 2011.

[3] *Eugen Borcoci, Daniel Negru, Christian Timmerer*, "A Novel Architecture for Multimedia Distribution Based on Content-Aware Networking", Third International Conference on Communication Theory, Reliability, and Quality of Service, CTRQ 2010, pp.162-168.

[4] FP7 ICT project, “MediA Ecosystem Deployment Through Ubiquitous Content-Aware Network Environments”, ALICANTE, No248652, http://www.ict-alicante.eu/.

[5] *S.Deering*, IETF RFC 1112 “Host Extensions for IP Multicasting”, 1989.

[6] *C. Diot et.al*, “Deployment Issues for the IP Multicast Service and Architecture”, IEEE Network, **Vol. 14**, No. 1, 2000, pp. 78–88.

[7] *H. Asaeda et.al.*, “Architecture for IP Multicast Deployment: Challenges and Practice”, IEICE Transactions on Communications, **Vol. E89-B**, No. 4, 2006, pp. 1044–1051.

[8] *E.W.Biersack*, “Where is Multicast Today?” ACM SIGCOMM Computer Communication Review, **Vol. 35**, No. 5, 2005, pp. 83–84.

[9] *B.Jamin and L.Zhang*, “Host Multicast: A Framework for Delivering Multicast to End Users”, Proceedings of the IEEE INFOCOM 2002, **Vol. 3**, 2002, pp. 1366–1375.

[10] *L.Lao, J.-H Cui, M. Gerla, D.Maggiorini*, “A Comparative Study of Multicast Protocols: Top, Bottom, or In the Middle?”, Proceedings of the IEEE INFOCOM 2005, **Vol. 4**, Miami, USA, 2005, pp. 2809–2814.

[11] *Marius Vochin, Eugen Borcoci, Dragoş S. Niculescu, Mihai Stanciu*, “Scalability Analysis of a Media Aware Network Element”, The 2012 European Signal Processing Conference (EUSIPCO-2012).

[12] *Dragoş S. Niculescu, Mihai Stanciu, Marius Vochin, Eugen Borcoci*, “Implementation of a Media Aware Network Element for Content Aware Networks”, Fourth International Conference on Communication Theory, Reliability, and Quality of Service, CTRQ 2011.