Monitoring and measuring wireless network performance in the presence of *middleboxes*

Alessio Botta and Antonio Pescapé

University of Napoli Federico II (Italy) – Email: {a.botta,pescape}@unina.it

Abstract—**Monitoring and measurement of wireless networks play an important role for a number of reasons. In current networking scenarios, middleboxes (NAT, firewall, PEP, PCMS, ...) are more and more common. In this paper we present preliminary results of our ongoing work on the monitoring and measurement of the performance of real operational wireless networks in the presence of middleboxes. We analyze the 3G network of two of the major operators in Europe and the satellite network of one of the major European providers of satellite connectivity. Our results show how middleboxes can deeply change the results of an experiment: in the presence of such devices, the performance parameters of interest can obtain completely different values when measured with different tools and methodologies. To cope with these issues, we propose an approach composed of both active and passive techniques, considering several layers in a combined fashion, using different traffic profiles, and looking at both side of the considered middleboxes.**

I. INTRODUCTION

Inferring the status of real and large scale wireless networks is not always easy. Difficulties arise from well known factors: new networking environments employing unknown and proprietary protocols, variability of the measured parameters with external factors (e.g. interference from other networks in the case of wireless links), influence of the measurement tool and methodology, lack of data due to privacy reasons, etc. Beside these well known issues, a new one has come to play an increasingly critical role: the presence of network middleboxes (MBs), more and more used in real operational wireless networks. MBs are intermediate devices that sit in the middle of an end-to-end path and perform functions originally delegated to the end hosts, thus breaking the end-to-end principle. They have been progressively introduced in operational networks to overcome performance and security limitations intrinsic in the TCP/IP protocol stack. To reach this target, MBs perform functions that range from simply altering the IP addresses (as done by NAT boxes), to fully modifying the traffic properties at multiple layers of the protocol stack (as done by performance enhancement proxies).

MBs can be an issue when performing network measurements, adopting both active and passive approaches. Being normally invisible to the end hosts, MBs can easily alter the result of a measurement without being noticed by the experimenter. For example, if we are measuring the round trip time (RTT) using a passive approach (e.g. counting the time elapsed from the data packet to the ACK), while a performance enhancement proxy (PEPs) is terminating the TCP connection on both sides of an end-to-end path, we will end up measuring

the RTT of only a part of the path, depending on where the PEP is. Also, if we want to infer the maximum throughput of the network using an active approach, while the network employs a shaper that discriminates a part of the traffic, we may obtain a result that is not meaningful of the network, but rather represents the rate allowed by the shaper for that particular kind of traffic. These examples show the importance of taking into the proper account the presence of MBs when performing network measurements.

In this paper we show the results of our ongoing work on how to properly perform measurement and monitoring campaign on wireless networks in the presence of MB. Our preliminary results show how MBs can deeply change the results of an experiment. In the presence of such devices, the performance parameters of interest can obtain values that differ of orders of magnitude when measured at different layers of the protocol stack, using different traffic profiles, and in different parts of the network. To overcome these issues, faced also by other researchers when performing similar works (e.g. [1], [2], [3], as discussed in Sec. IV), we propose an approach (i) composed of both active and passive techniques; (ii) considering several layers in a combined fashion (i.e. transport and application layer); (iii) using different traffic profiles; (iv) looking at both side of the considered MB. Our approach allows on the one side, to revel the presence of such devices, and on the other side, to understand the tools most suited to the network scenario of interest.

II. REFERENCE SCENARIOS

In this paper we use the definition given in [4] for the term *middlebox*: "*any intermediary box performing functions apart from normal, standard functions of an IP router on the data path between a source host and destination host*". This definition is broad enough to accommodate a wide set of devices, such as, for example, NAT boxes, SOCKS gateways, Performance Enhancement Proxies (PEPs), traffic anonymizers, firewalls, and Packet Classifiers Markers and Schedulers (PCMSs) (see [4] for a complete list and taxonomy). Such devices have been progressively introduced in the Internet architecture to solve a number of issues, such as security and performance. In this paper – because they currently are the MBs more often causing issues on the monitoring and measurement of operational wireless networks – we focus on two particular kinds of MBs, that are PEPs and PCMSs:

a) PEP: A Performance Enhancement Proxy is a very wide-spread MB, whose aim is helping the users of a certain

network to improve his/her quality of experience. The proxying functionalities are normally implemented by terminating the TCP connections on the client side and opening new TCP connections towards the server. Operating as transparent proxies, they perform these actions without being noticed by the users. The performance enhancement features can be multiple, and depend on the network and application on which the PEP has to operate. They span over multiple layers of the protocol stack, typically from the transport to the application ones. For example, in the case of satellite networks (see Sec. III-C), the PEP can be located between the end user and the satellite link (e.g. co-located with the modem), in order to use a satellite-oriented TCP version, which is often not employed by end-users. For a broader overview PEP features we refer the reader to [5].

b) PCMS: Packet Classifiers Markers and Schedulers allow to identify and classify the packets of the different applications (or the different activities of complex applications) in order to provide a specific service for each traffic class. This is the case, for example, of devices that prioritize packets carrying voice samples of an audio conferences with respect to packets carrying other traffic typologies. PCMSs are often utilized in cellular networks to avoid users to saturate the link bandwidth with aggressive applications, such as Bittorrent (see Sec. III-B).

III. ON THE IMPACT OF MBS

In this section we report our experience regarding the influence of MBs on the monitoring and measurement of three wireless networks.

A. Impact of PEPs in a cellular network

To show how monitoring and measurement of network parameters in presence of PEPs can provide highly different results if parameters are observed in different parts of the network, in this section we analyze an operational cellular network (we call *cellular network 1*), which employs a PEP. *Cellular network 1* is from one of the major telecom operators in Europe, counting millions of users in different countries. In this activities we worked using a passive approach, on traffic traces captured at both sides of the PEP. Fig. 1 shows the network scenario we used for the tests. Speaking with the network operator, we could confirm that the PEP is located at the border of the cellular network, before the Internet access.
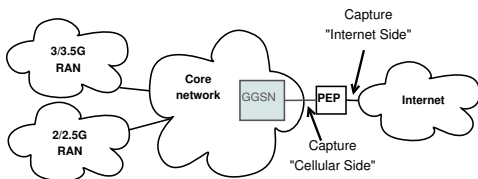

Figure 1.   Schema of the cellular network 1 with the PEP.

Fig. 2 shows the cumulative distribution function (CDF) of the average round trip time (RTT) per connection, calculated using Tstat[1] on two traffic traces captured at the two sides

[1]http://tstat.tlc.polito.it

of the PEP. The PEP of this cellular network terminates the TCP connection towards the mobile stations and opens new TCP connections towards the hosts on the Internet. As a consequence, the total end-to-end RTT can be seen as the sum of two contributions: the cellular side plus the Internet side. For this reason, it is important to take into the proper account the presence of such device when performing network measurements: it can be useful for those interested in estimating the performance as seen by the users (in fact, they only see the cellular side), while being an issue for those interested in looking at the end-to-end performance. In general, we can say that due to the presence of the PEP, the RTT measured in this network can be highly different if observed at both side of the MB.
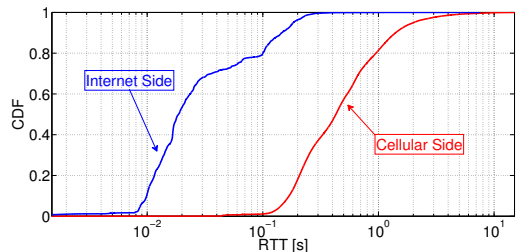

Figure 2.   Results from cellular network 1.

B. Impact of PCMSs in a cellular network

To show how performance parameters can be highly different if measured with different traffic profiles, in this section we report the results of experiments performed over the cellular network of another European 3G/4G network operator (we call it *cellular network 2*). The experiments have been conducted between a mobile station and a server under our control, using an active approach with D-ITG [6]. We generated several traffic flows, with different traffic profiles (in terms of packet sizes and inter-departure times), and filling the payload of the probing packets with both random bytes and real application payloads.
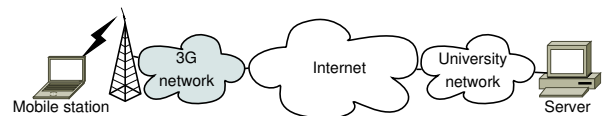

Figure 3.   Schema of the cellular network 2.

The tests revealed that the behavior of the performance parameters was different depending on the application emulated through D-ITG. For example, Fig. 4 shows the throughput (top) and jitter (bottom) over time obtained with two TCP traffic flows generated from the mobile station towards the server. The flows have the same imposed bitrate (400 Kbps), but they differ in terms of packet size and rate and payload of the packets: one is replicating an HTTP flow and the other is replicating a Bittorrent flow. As shown, the two flows obtain quite different results. In particular, the throughput of the HTTP-like flow is higher with respect to the other flow, their average values being 350Kbps and 58Kbps respectively. The

bottom plot of Fig. 4 shows that also the jitter of the two flows achieve different values over time. The average value for the HTTP-like is 32ms, while for the Bittorrent-like flow is 191ms. This is the effect of a PCMS.
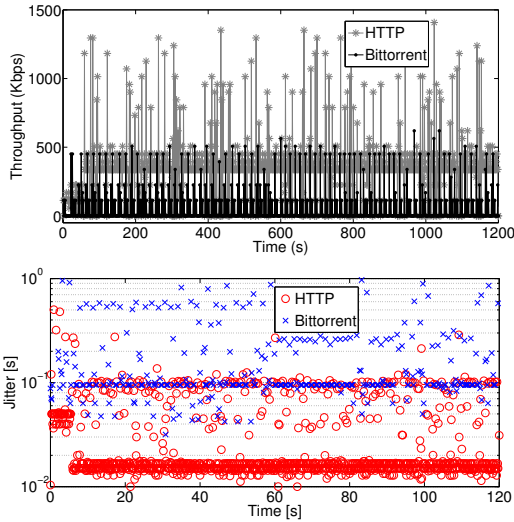

Figure 4. Results obtained on cellular network 2.

To understand how the PCMS classifies the packets we performed a set of experiments using different transport layer ports, different packet rates and sizes, and different payload types. The results revealed that this operator employs a PCMS and the main discriminating factor is the transport layer port. Concluding, we can say that the performance parameters can be highly different if measured with traffic flows having different port number and different traffic profiles.

C. Impact of PEPs in a satellite network

To show how adopting a different point of view when performing performance measurements may imply a completely different result, in this section we report the results obtained over a network testbed comprising satellite links from one of the major European providers. Fig. 5 depicts its schema: the testbed is composed of four hosts connected to the Internet each through a bidirectional satellite link, and two servers using our University broadband network. The experiments were conducted by using a combined, active/passive approach. We employed active tools such as D-ITG and iperf[2], as well as passive tools such as Tstat and tcptrace[3].

In these activities we performed a characterization of the performance in terms of parameters such as throughput, jitter, losses, and delay (RTT). Among other interesting results (not reported here for space constraints), we found that all the parameters were highly influenced by the PEP enclosed in the satellite modem provided by the operator. For example, Fig. 6 shows the RTT of a flow of TCP packets having a payload of 64 Bytes, injected at a rate of 1000pps (injected rate is 512 Kbps at application layer). In this figure we compare the

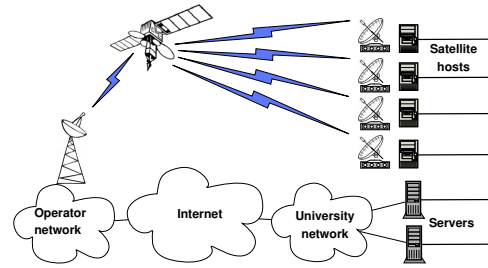[2]http://sourceforge.net/projects/iperf
[3]www.tcptrace.org

Figure 5. Schema of the satellite network.

results reported by D-ITG with those calculated by tcptrace on the same traffic. As we can see, the two tools report very different values. This is also more evident comparing the average values reported in Table I.
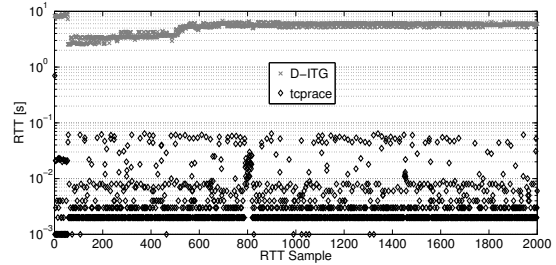

Figure 6. Results obtained on satellite network.

Note that RTT samples are calculated at different layers of the protocol stack: at the TCP layer for the top plot, at the application layer for the bottom plot. While it is normal that at different layers we see different values, such difference cannot be simply explained in this way. The difference is instead due to the fact that in such network scenario the TCP connections are terminated by the PEP, which is located in the satellite modem, connected to the host initiating the connection through a 100Mbps link. Therefore, the RTT seen at transport layer (reported by tcptrace) is only related to such local network link. On the other hand, the RTT at the application layer (reported by D-ITG) is related to the complete end-to-end path. This result shows that, when performing performance measurements, a different point of view may imply a completely different result, when MBs are present.

Table I
RTT ON THE SATELLITE LINK MEASURED WITH D-ITG AND TCPTRACE.

	D-ITG	tcptrace
PS / PR	RTT [s]	
64 / 10	0.6280	0.0528
128 / 10	0.5934	0.0475
256 / 10	0.6028	0.0344
512 / 10	0.6912	0.0423
64 / 1000	4.4296	0.0064

IV. RELATED WORK

Middleboxes (MBs) are more and more used in the Internet for a number of reasons. In [4], the authors systematically analyzed the available MBs, formally defining and classifying them in a taxonomy. The RFC discusses also potential issues related to the interconnection and intercommunication of such

devices. Among the available MBs, in this work, we focus our attention on PEP (Performance Enhancement Proxies) and PCMSs (Packet Classifiers, Markers and Schedulers) [5]. The presence of PEPs is widely spread in wireless networks and several commercial appliances are adopted [7], [8], [9], [10]. In [11], [12], [13], [14], [15], [16], [17] the use of PEPs in satellite networks is analyzed with respect to several issues (performance, security, etc.); in [18], [19], [20], [21], [22] is shown how the presence of PEPs can increase the performance of 3G/4G networks, whereas in [23] and [24] the use of PEPs is studied respectively in wi-fi and wimax scenarios. In [25], the authors recognized the wide spread use of middleboxes and, at the same time, the potential dangerous side effects (e.g. violation of important architectural principles and consequent less flexibility), whereas the interactions between TCP and MBs has been studied in [1]. As for PCMSs, in [26] the authors present an end-to-end approach for packet forwarding prioritization inference, whereas in [27] the authors propose a passive approach for the detection of discrimination mechanisms. From the monitoring and measurement point of view, one of the first works to report the issues that such middleboxes may cause in the evaluation of the network performance is [2] whereas other works [28], [16], [3] report on the issues related to the monitoring and measurement stages in presence of PEPs. With respect to [1], in our work, thanks to a combined (active and passive) analysis, we can precisely pinpoint the origin of the problem we discovered. With respect to [2], by having access to packet traces, we can dig into the causes of the issues caused by middleboxes on the estimation of network performance. Finally, with respect to [3] we analyze more scenarios and provide a methodology to cope with issues related to the monitoring and measurement stages in presence of PEPs.

V. CONCLUSION

In this paper, we have presented preliminary results from our study on how to perform monitoring and measurement of network performance in presence of middleboxes. We have reported results related to three real operational wireless networks, underlining when and how middleboxes can affect results of monitoring and measurement activities. We have learned that measuring the performance of current middlebox-rich networks requires more awareness. Finally, to reveal the presence of such devices and to cope with them, we have proposed an approach using both active and passive techniques, looking at different layers of the protocol stack, using different traffic profiles, and at both side of the middlebox. Our ongoing work is related with systematically categorize all the issues that arise from the presence of such devices, and with devising a robust methodology to perform these studies.

ACKNOWLEDGEMENTS

REFERENCES

[1] A. Medina, M. Allman, and S. Floyd, "Measuring interactions between transport protocols and middleboxes," in *ACM IMC*, 2004.

[2] M. Allman, "On the performance of middleboxes," in *ACM IMC*, 2003.

[3] W. L. Tan, F. Lam, and W. C. Lau, "An empirical study on the capacity and performance of 3g networks," *IEEE Transactions on Mobile Computing*, vol. 7, no. 6, pp. 737–750, 2008.

[4] B. Carpenter and S. Brim, "Middleboxes: Taxonomy and issues," February 2002, RFC 3234.

[5] J. Border, M. Kojo, J. Griner, G. Montenegro, and Z. Shelby, "Performance enhancing proxies intended to mitigate link-related degradations," June 2001, RFC 3135.

[6] A. Botta, A. Dainotti, and A. Pescapé, "Multi-protocol and multi-platform traffic generation and measurement," in *IEEE INFOCOM, Demo session*, 2007.

[7] http://www.viasat.com.

[8] http://www.mentat.com.

[9] http://www.bluecoat.com.

[10] http://www.hughes.com.

[11] C. Caini and R. Firrincieli, "Tcp hybla: a tcp enhancement for heterogeneous networks," *International Journal of Satellite Communications and Networking*, vol. 22, 2004.

[12] D. Velenis, D. Kalogeras, and B. S. Maglaris, "SaTPEP: A TCP Performance Enhancing Proxy for Satellite Links," in *NETWORKING*, 2002.

[13] C. Marcondes, A. Persson, M. Sanadidi, M. Gerla, R. Firrincieli, D. Beering, and G. Romaniak, "TCP in mixed Internet and geo-satellite environments: experiences and results," in *TRIDENTCOM*, 2006.

[14] C. Roseti, F. Zampognaro, and M. Luglio, "Enhancing TCP Performance over Hybrid Wireless Terrestrial-Satellite Networks," *International Conference on Advances in Satellite and Space Communications*, 2009.

[15] G. Acar and B. Evans, "Impact of Proactive Retransmissions on Forward TCP Throughput over DVB-S/S2 Railroad Satellite Links with Power Arches," in *4th Advanced Satellite Mobile Systems*, 2008.

[16] N. Ehsan, M. Liu, and R. Ragland, "Measurement based performance analysis of internet over satellite," in *SPECTS 2002*, 2002.

[17] C. Caini *et al.*, "Analysis of TCP live experiments on a real GEO satellite testbed," *Perform. Eval.*, vol. 66, no. 6, pp. 287–300, 2009.

[18] P. Rodriguez and V. Fridman, "Performance of peps in cellular wireless networks," in *International workshop on Web content caching and distribution*, 2004.

[19] M. Ivanovich, P. Bickerdike, and J. Li, "On tcp performance enhancing proxies in a wireless environment," *Communications Magazine, IEEE*, vol. 46, no. 9, pp. 76 –83, sep. 2008.

[20] C. Gomez, M. Catalan, D. Viamonte, J. Paradells, and A. Calveras, "Web browsing optimization over 2.5g and 3g: end-to-end mechanisms vs. usage of performance enhancing proxies," *Wirel. Commun. Mob. Comput.*, vol. 8, no. 2, pp. 213–230, 2008.

[21] M. C. Necker, M. Scharf, and A. Weber, "Performance of different proxy concepts in umts networks," in *Wireless Systems and Mobility in Next Generation Internet*, ser. LNCS, 2005, vol. 3427, pp. 36–51.

[22] A. Botta, A. Pescapé, G. Ventre, E. Biersack, and S. Rugel, "Performance footprints of heavy-users in 3G networks via empirical measurement," in *The 6th International Workshop on Wireless Network Measurements (WiNMee 2010)*, may 2010, pp. 330 –335.

[23] L. Munoz, M. Garcia, J. Choque, R. Aguero, and P. Mahonen, "Optimizing internet flows over ieee 802.11b wireless local area networks: a performance-enhancing proxy based on forward error correction," *Communications Magazine, IEEE*, vol. 39, no. 12, pp. 60–67, dec. 2001.

[24] K. Pentikousis, E. Piri, J. Pinola, F. Fitzek, T. Nissilä, and I. Harjula, "Empirical evaluation of VoIP aggregation over a fixed WiMAX testbed," in *TridentCom*, 2008.

[25] M. Walfish, J. Stribling, M. Krohn, H. Balakrishnan, R. Morris, and S. Shenker, "Middleboxes no longer considered harmful," in *USENIX OSDI*, 2004.

[26] G. Lu, Y. Chen, S. Birrer, F. Bustamante, and X. Li, "Popi: A user-level tool for inferring router packet forwarding priority," *Networking, IEEE/ACM Transactions on*, vol. 18, no. 1, pp. 1 –14, feb. 2010.

[27] M. B. Tariq, M. Motiwala, N. Feamster, and M. Ammar, "Detecting network neutrality violations with causal inference," in *ACM CoNEXT*, 2009.

[28] S. Lau and L. Trajković, "Analysis of traffic data from a hybrid satellite-terrestrial network," in *QSHINE*, 2007.