Real-time video adaptation in virtualised 5G networks

Salva-Garcia, Pablo; Alcaraz-Calero, Jose M.; Wang, Qi; Barros, Maria; Gavras, Anastasius

Published in:
Proceedings of the 44th Annual IEEE Conference on Local Computer Networks

DOI:
10.1109/LCN44214.2019.8990815

Published: 13/02/2020

Real-Time Video Adaptation in Virtualised 5G Networks

Pablo Salva-Garcia*, Jose M. Alcaraz-Calero*, Qi Wang* Maria Barros† and Anastasius Gavras†
* University of the West of Scotland, United Kingdom
†Eurescom GmbH, Germany

***Abstract*—Video applications are expected to increasingly dominate the traffic of mobile networks in the 5G era, and thus a real-time adaptation of these high resource demanding network applications is crucial in optimising the overall 5G networks. In this manuscript, we leverage Virtual Network Function (VNF) techniques to implement a video adaptation service that automatically adapts the quality of video transmissions depending on the status of the network. Furthermore, Network Function Virtualisation (NFV) techniques are employed here to simplify, optimise and speed up the deployment process of the aforementioned video adapter service (vAdapter), and therefore, allowing its on-demand deployment in a flexible way. We design, implement and test the scheme in a realistic virtualised 5G testbed. Empirical results focus on the scalability evaluation and performance as well as demonstrates a significant bandwidth reduction without compromising the final user's video quality expectations.**

***Index Terms*—Real-time; Multimedia; adaptive; SDN; NFV; 5G; Network Management; QoS.**

I. Introduction

Video traffic has dominated the traffic in global mobile networks, and the trend is that such dominance will continue and grow over the coming years, with the increasing popularity of bandwidth-demanding video applications such as Virtual/Augmented Reality, and novel visual applications enabled by 5G such as remote surgery for eHealth use cases. It is high time to design, deploy and test adaptive networking applications such as the presented video adapter schema to allow multimedia real-time services to occurs. Software Defined Networking (SDN) and Network Function Virtualization (NFV) techniques presented nowadays in novel virtualised architectures pave the way to investigate, develop and efficiently deploy alternatives to ensure successful network transmissions even when the network becomes congested and unstable.

These applications typically adopt high-definition (HD) or ultra-high-definition (UHD) resolutions and frame rates, generating continuously heavy traffic load to the network and leading to congestion and thus overall performance downgrading in the 5G system. Meanwhile, existing video tools can hardly recognise 5G traffic or cope with the latest video codec (e.g., H.265/HEVC and its scalable extension SHVC [1]) in the virtualised 5G environment. Therefore, it is essential to design and develop a new video adapter that is capable of parsing the 5G traffic and processing the video traffic encoded by the latest codec, preferably in an autonomous way, and empirically test and evaluate its functionality and performance, which is the main contribution of this work. Section II describes the state of the art in 5G video adaptation techniques. Section III presents the virtualised 5G infrastructure testbed. Section IV focuses on presenting and analysing the experimental results. Finally, conclusions and future work activities are drawn in Section V.

II. Background and Related Work

The H.265 video codec standard can reduce bandwidth requirements up to 50% without decreasing the quality throughout the encoding process [2]. Recent improvements in the codec standard [3] show its potential to overpass it is predecessor H.264 codec [4] with respect its usage in the network. H.265 also supports scalable video encoding that consists of multiple layered video flows, where a base layer comprises the basic version of the video while on or more enhancement layers incorporate additional data to enhance the video quality. In this way, a Media Aware Network Element (MANE) with scalable H.265 capabilities would be able to adapt a video streaming by a selective dropping of its enhancement layers.

Not just video coding requirements but also 5G requirements must be addressed for any solution that aims to adapt video flows traversing this kind of networks: Embrace virtualization and softwarisation technologies to provide scalability and on-demand deployments, deal with traffic in multi-carrier and mobility scenarios for 5G video streaming; support nested encapsulation demands imposed by both core network and edge segments of the 5G multi-tenant networks; and perform filtering in the inner layers of overlay networks. To the best of our knowledge, little existing work has considered both, video-coding and 5G network requirements. For instance, Authors in [5] propose a video adaptation technology based on a variable bitrate transcoding but there is no information about the video codec, in [6] is presented a 4K scalable H.265 video streaming without highlight the streaming protocol nor the video adaptation technique used, in [7] it is approached a media-aware network element with scalable H.265 capabilities, but scalability and flexibility features, which are being fostered in 5G implementations, are not achieved. Other studies such as [8], [9] do not give information about the video codec or multi-tenancy support. As a direct consequence, there is a lack of scalable network video optimization systems that are simultaneously able to deal with traffic encapsulation demands imposed by both core network and edge segments of the 5G multi-tenant networks and perform filtering in the inner layers of overlay networks.

III. Virtualized 5G Infrastructure

An empirical virtualized Long Term Evolution (LTE) based infrastructure has been designed and deployed. The infrastructure also presents several 5G features, which allow a realistic analysis of video traffic in LTE-based 5G networks.

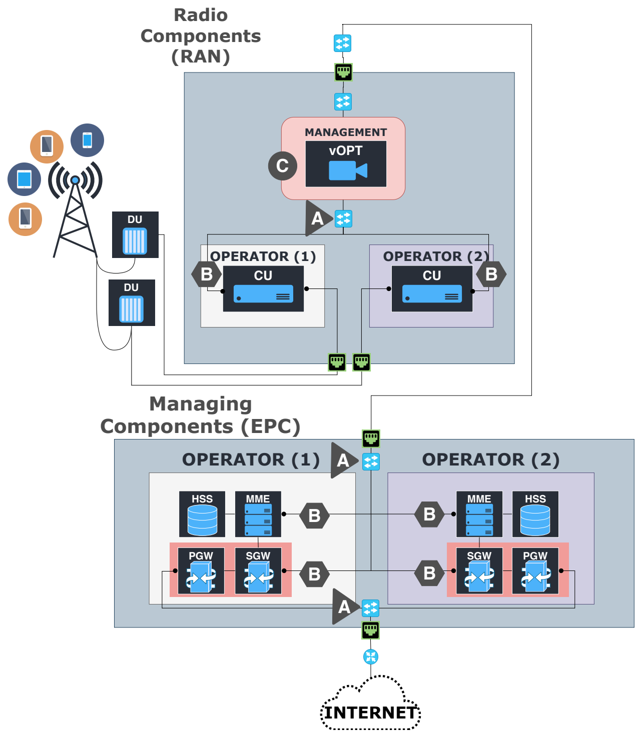

Fig. 1: Scenario for a real-time video filtering in virtualized and multi-tenant 5G Infrastructure

A. 5G Infrastructure

Figure 1 illustrates a virtualised architecture where although different operators are sharing physical resources, each one represents a completely different tenant administrative domain. The infrastructure is composed of 10 computers with Ubuntu 16.04 and provisioned with OpenStack. our scenario relies on Neutron and OpenDayLight as Software-Defined Networking (SDN) controller. In turn, OpenDayLight uses OpenFlow and OVSDB for controlling the Open Virtual Switch (OVS) v2.9 and the data path of the Virtual Machines (VMs).

To allow functional decoupling between 5G Distributed Units (DU) and Centralised Units (CUs), an evolution of the OpenAirInterface, the Mosaic5G [10] project, has been deployed. It is noted that even when the core network is still using LTE-based terminology for its components, all of them have been fully virtualised and run in VNFs. Therefore, providing a realistic LTE-based 5G infrastructure for research and evaluation of traffic across all segments of the network.

Figure 1 shows different network segments that traffic crosses. In each of the network segments, traffic can be encapsulated by using different protocols for multi-tenancy and/or mobility support. It has been labelled with *A* for those that allow tenant isolation by using OpenStack and are being encapsulated in protocols such as VLAN, VXLAN or GRE. Labelled with *B* are the data-plane control points where GTP encapsulation is used for mobility support. Finally, points marked with a *C* label are those where our proposed vAdapter (or stack of vAdapters) is deployed for dealing with any nested-encapsulated traffic.

B. Video Traffic in 5G

The proposed vAdapter is able to deal with the scalable extension of the H.265/HEVC video codec. The approach of this study is based on real-time streaming video, thus Real-Time Transport Protocol (RTP) has been selected for streaming video flows. RTP allows for packetisation of the H265 bitstream using the Network Abstraction Layers (NAL) of each RTP packet, thus supporting HEVC streaming for video on demand, among other applications. Those NALs are a sequence of data units that contain both the payload and high-level information to allow parsing the main properties. The vAdapter uses such high-level information for detecting and applying the enforced policies to the specific video flows. In this particular case (video adapter service), our vAdapter is not only able to operate at the application layer, where the video protocols are placed, but also deal with the complex hierarchical encapsulation that can be found throughout a 5G architecture.

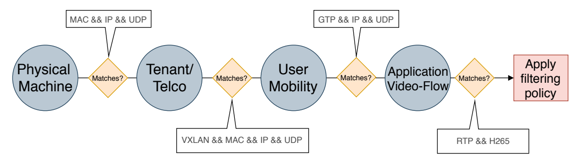

Fig. 2: Hierarchical Encapsulation in 5G Networks.

An example of nested encapsulation is depicted in Figure 2. Our video adapter service deals with protocols belonging to the physical machine's communication, protocols used for tenant isolation, protocols that allow user mobility, and finally, those protocols belonging to the application layer before applying any video adaptation. Each protocols stack provides information about the specific network segment by which the traffic passes, therefore, each of them must be carefully inspected to make sure that an adaptation is going to be placed over the targeted video flow without compromising the proper functioning of the rest of the network traffic. To do so, the video adapter solution proposed in this paper leverages and adapts the existing filtering mechanism BPF [11], that provides an efficient way of filtering packets in the kernel space, to filter such complex and nested encapsulated media traffic in 5G Networks.

IV. Empirical Results

This section aims to demonstrate the suitability of the proposed vAdapter service when deployed in a realistic 5G network. Furthermore, scalability and flexibility of the service are also tested by deploying several vAdapter as Virtual Network Functions (VNFs), i.e., vAdapterVNFs. Finally, end-to-end results are evaluated.

Bitstreams encoded in SHVC were sent through the real deployment presented in section III-A. Different Packet Captures (PCAP) files were obtained from such streaming between the video streamer (an LTE-based smartphone) and the Serving Gateway (SG). This testbed implementation uses those PCAP files, which contain a real and nested encapsulated 5G packet structure, for testing the system behaviour when increasing both, the number of simultaneous video flows traversing the testbed and the number of vAdapters instantiated.

A. Implementation of the 5G Testbed

In a realistic 5G scenario, thousands of users will use a video streaming service simultaneously. Our proposed vAdapter has been designed to optimise each of those video streams regarding the experience perceived by the user. Therefore, massive video traffic must be properly inspected and a consequent set of traffic policies applied per every single flow. It is not feasible to handle such large quantities of policies and massive traffic with just one vAdapter. Table I, shows the number of simultaneous video flows able to inspect with no packet loss and acceptable delay/jitter values, 128. A scalable approach is therefore presented where thousands of video flows are managed simultaneously. Our solution benefits from leveraging NFV and cloud computing technologies to deploy dynamically and on-demand, VNFs in the format of distributed vAdapterVNFs.

TABLE I: Summary statistics when 128 video-flows are simultaneously handled by one vAdapter.

Percentage of Packet Loss (%)	0.00%
Average of Delay (ms)	0.06109873
Average of Jitter (ms)	0.013810244

Each vAdapterVNF deals with a subset of the rules according to a network segmentation addressed in a particular RAN. A first filtering agent acts as a load balancer redirecting the traffic quickly to the pertinent vAdapterVNF responsible for handling a subset of rules. The network segmentation and forwarding in the load balancer can be achieved with just one rule per deployed subsequent vAdapterVNF, inspecting the inner IP packet of the encapsulated traffic. This scalable approach enables the deployment of additional vAdapterVNFs according to the network conditions. Fig. 3 depicts the infrastructure used in our testbed. 12 vAdapterVNFs were needed to cope with 1536 flows generated, each one holding up to 128 filtering rules, one per flow. As empirically demonstrated in Table I, 128 is the advisable number of complex rules per vAdapterVNF without incurring packet loss or unacceptable delays/jitter.

B. Test Results

Following results are produced by measuring the concerned metrics (latency etc.) from point A to B in Figure 3, or in other words, from the moment that video flows reach the physical machine (where our vAdapters are running as VNFs) to the point they leave the computer. Therefore, load balancer delay, network delay and video adaptation delay introduced by each vAdapterVNF are also considered.

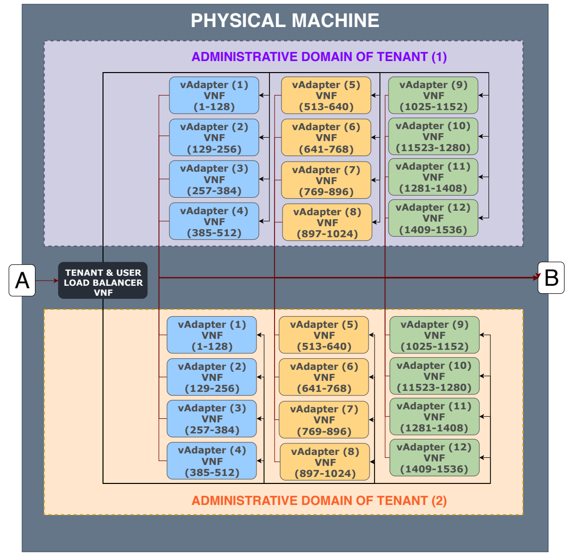

Fig. 3: Implementation of the performance evaluation testbed.

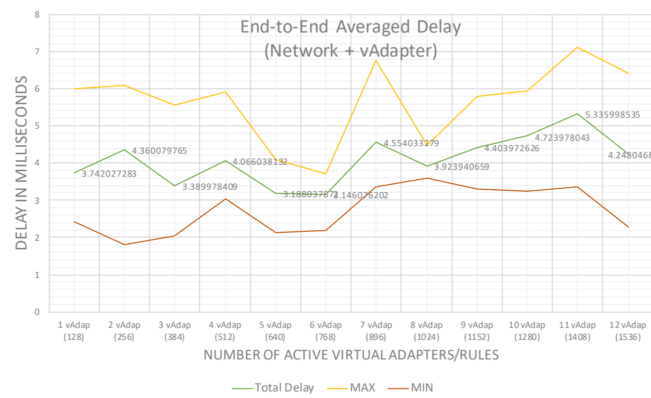

Fig. 4: E2E delay according to the number of vAdapterVNFs

From our first experiment shown in Figure 4, it can be seen that the total delay was nearly constant even when the number of vAdapterVNFs was increasing. There is not a direct correlation between the number of video services used and the delay introduced in the system. Thus, it can be concluded that our proposal is scalable since the addition of a new vAdapterVNF in our system does not affect the end-to-end delay. To be precise, the delay was on average about 4ms and reaching maximums of around 7ms was the worst cases. It is important to highlight that every vAdapterVNF included in the system was fully working by processing hundreds of video flows and dropping enhancement video layers, which have the least impact on the perceived quality of the video.

Figure 5 provides the information to differentiate the delay introduced by the proposed vAdapterVNF versus the one produced by the network to divert the traffic across the VMs. As can be seen, networking delays were constant (about 3.5ms) while the time consumed by the vAdapters is around 0.5ms. Therefore, it can be reaffirmed that the increasing number of vAdapters did not introduce a proportional delay, and there was no correlation between the number of services and the time consumed by such processes.

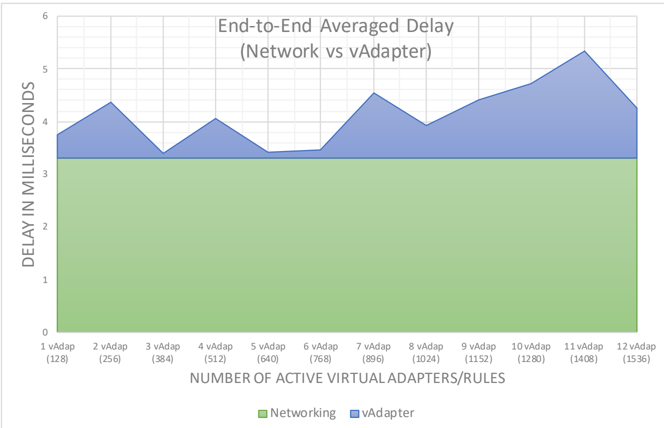

Fig. 5: Network vs vAdapter End-to-end delay

Finally, to evaluate the accumulative bandwidth that could be saved every second during the streaming, this study has conducted an experiment as shown in Figure 6. In this particular case, a set of scalable videos were streamed and adapted by several vAdapters (128 video flows per each vAdapter). Each video flow was composed by a base layer with FHD spatial resolution and an enhancement layer in UHD. A selective dropping of the enhancement layer per each video flow produced a significant amount of bandwidth saved. This experiment has been conducted by sending scalable videos of 3.8Mbytes in a bitrate of 1,010,000bps, with a stream duration of 30.57s. In the most stressful scenario, 1536 video flows were being processed by the 12 vAdapters in the system. As can be seen in Figure 6, at the end of the transmission about 2,981Mbytes (close to 3Gbytes) were saved. From the network point of view, possible congestion has been mitigated; from the users' point of view, the perceived quality remained stable since there were no uncontrolled packet loss but just a downgrading of the video resolution.

V. Conclusions

This manuscript has presented the design and implementation of a new video adapter that is compliant with the latest video codec and can be flexibly deployed in virtualised 5G networks as a VNF. Results gathered from a realistic 5G testbed have confirmed the capabilities of the presented video adapter in dissecting and processing 5G video traffic over the novel virtualised and multi-tenanted infrastructure. Furthermore, scalability results have proved the ability of the video adapter service to act as a complete stack of VNFs to handle massive video traffic effectively. In addition, promising results have been obtained in the saving of bandwidth during the adaptation process and, consequently, it has been demonstrated how to alleviate network congestion without compromising the quality of the services offered.

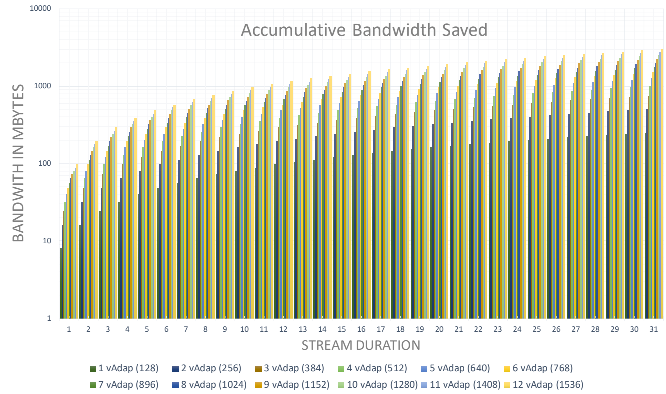

Fig. 6: Accumulative bandwidth saved per second

Acknowledgment

This work was funded in part by the European Commission Horizon 2020 5G-PPP Programme under Grant Agreement Number H2020-ICT-2016-2/761913 (SliceNet: End-to-End Cognitive Network Slicing and Slice Management Framework in Virtualised Multi-Domain, Multi-Tenant 5G Networks).

References

[1] J. M. Boyce, Y. Ye, J. Chen, and A. K. Ramasubramonian, "Overview of SHVC: Scalable extensions of the high efficiency video coding standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 26, no. 1, pp. 20–34, Jan. 2016.

[2] J. R. Ohm, G. J. Sullivan, H. Schwarz, T. K. Tan, and T. Wiegand, "Comparison of the coding efficiency of video coding standards-including high efficiency video coding (HEVC)," *IEEE Transactions on Circuits and Systems for Video Technology*, 2012.

[3] S. Radicke, J. U. Hahn, Q. Wang, and C. Grecos, "A parallel HEVC intra prediction algorithm for heterogeneous CPU+GPU Platforms," *IEEE Transactions on Broadcasting*, 2016.

[4] T. Wiegand, G. J. Sullivan, G. Bjøntegaard, and A. Luthra, "Overview of the H.264/AVC video coding standard," *IEEE Transactions on Circuits and Systems for Video Technology*, 2003.

[5] M.-A. Kourtis, H. Koumaras, G. Xilouris, and F. Liberal, "An NFV-based video quality assessment method over 5G small cell networks," *IEEE Multimedia*, 2017.

[6] O. Awobuluyi, J. Nightingale, Q. Wang, and J. M. Alcaraz-Calero, "Video quality in 5G networks: Context-Aware QoE management in the SDN control plane," in *2015 IEEE International Conference on Computer and Information Technology; Ubiquitous Computing and Communications; Dependable, Autonomic and Secure Computing; Pervasive Intelligence and Computing*, Oct. 2015, pp. 1657–1662.

[7] P. Salva-Garcia, J. M. Alcaraz-Calero, R. M. Alaez, E. Chirivella-Perez, J. Nightingale, and Q. Wang, "5G-UHD: Design, prototyping and empirical evaluation of adaptive Ultra-High-Definition video streaming based on scalable H.265 in virtualised 5G networks," *Computer Communications*, vol. 118, pp. 171–184, 2018.

[8] C. Wang, Z. Lin, S. Yang, and P. Lin, "Mobile edge computing-enabled channel-aware video streaming for 4G LTE," in *2017 13th International Wireless Communications and Mobile Computing Conference (IWCMC)*, Jun. 2017, pp. 564–569.

[9] C. Ge, N. Wang, G. Foster, and M. Wilson, "Toward QoE-Assured 4K Video-on-Demand delivery through mobile edge virtualization with adaptive prefetching," *IEEE Trans. Multimedia*, vol. 19, no. 10, pp. 2222–2237, Oct. 2017.

[10] C. Commons, "Mosaic5G." [Online]. Available: http://mosaic-5g.io/

[11] S. McCanne and V. Jacobson, "The bsd packet filter: A new architecture for user-level packet capture," in *Proceedings of the USENIX Winter 1993 Conference Proceedings on USENIX Winter 1993. San Diego, California*. USENIX Association, 1993, pp. 2–2.