End-to-end quality aware optimization for multimedia clouds

Dimitris Athanasopoulos*, Ilias Politis*, Asimakis Lykourgiotis*, Christos Tselios† and Tasos Dagiuklas*
*Hellenic Open University, Patras, 26335, Greece, Email: {dathanasopoulos, ipolitis, alykourgiotis, dagiuklas}@eap.gr
†University of Patras, Rio 26500, Greece, Email: tselios@ece.upatras.gr

Abstract—This article presents a novel, end-to-end, quality-aware optimization framework for multimedia clouds, where path selection mechanisms are exploited in conjunction with media optimization in order to support multimedia delivery in a quality-aware manner. As wireless data traffic worldwide is characterized by exponential growth, with the most prominent part being multimedia services, consumers get in the challenging position to compete for the limited wireless network resources. Cloud technologies and especially Software-Defined Networking is the perfect candidate technology in order to provide an elastic, dynamic provisioning of network resources that adapt to a highly changing environment, where application requirements and data volumes vary over time. This work combines the selection of the optimum path in the core network with quality-aware media adaptation based on the current conditions of the wireless access network. Thus the proposed framework achieves efficient network resources utilization in an end-to-end fashion.

I. Introduction

As multimedia and real-time applications are migrating on-line, new network architectures and novel paradigms for services and networks convergence, need to be described. Evidently, both Quality of Service (QoS) and Quality of Experience (QoE) need to be optimised and maintained by content providers and network operators [1], [2]. Currently, there are technological solution that could be deployed towards this end, including Content Delivery Network (CDN), Data Center Virtualization (DCV) and bulk data processing. Although the requirements for processing power, resource management and required bandwidth for such solutions to be utilised and deployed in large scales is considerable, Cloud infrastructure have recently arisen as the preferred solution.

This new environment, which is based on the software and hardware virtualization, has advantages from both the multimedia content and the network providers. The former will be driven to offer applications on demand, massively scalable and always on, in order to satisfy users with heterogeneous capabilities and characteristics, while following a pay-as-you-grow business model to minimise the costs. The latter, will exploit the users' ability to store, process and adapt multimedia data in the cloud and instantiate cloud functionalities that users may access with a minimum latency and user experience requirements [3], [4]. However, existing cloud computing technologies are not optimised for multimedia and video streaming services. Currently, the available computing infrastructures are striving to cope with the demand of multiple video flow handling in terms of encoding, processing and streaming.

Software-defined networks (SDNs) are a new framework of networking which abstracts underlying network infrastructure and provides access through software and applications allowing significant changes on network behaviour [5], [6]. This introduces the ability of applying business logic to network behaviour in a dynamic way and facilitates data flow control through constant (if not real-time) modification, since thousands of virtual machines and sizeable applications can be deployed. The availability of open APIs for management and data plane control, like OpenFlow [7], OpenStack [8] or OpenNaaS [9], provide an additional degree of integration of Network Functions Virtualization (NFV) and cloud infrastructure.

This paper presents a novel optimization framework, which combines SDN's dynamic video flow adaptation via a proposed QoS-aware path switching scheme, along with a media aware optimization algorithm that selectively shapes the video traffic according to the access network conditions. The proposed framework is evaluated based on test-bed experimentations using video sequences with various coding and content characteristics. The obtained results in terms of network throughput and perceived video quality are compared against a non-optimised set-up that acts as the baseline scenario.

The rest of the paper is structured as follows. Section II describes the proposed SDN path selection algorithm, while in Section III the media aware optimisation scheme is briefly discussed. Section IV includes the experimental set-up and an evaluation on the obtained results. The paper concludes in Section V.

II. SDN based video streaming optimization

SDN is able to provide high flexibility, unprecedented programmability and dynamic reconfigurability of the network resources. Along with the ever increasing demand for video traffic over wireless and wired All-IP networks, renders SDN the most suitable candidate technology to provide elasticity and dynamic provisioning of network resources in order to cope with the varying video streaming characteristics.

To achieve dynamic and just-in-time decision for QoS provision over the SDN network, a real-time network monitoring and analytics tool was designed. The tool transparently and efficiently collects link-based (i.e., maximum theoretical bandwidth, topology information, etc.) and traffic-based metrics
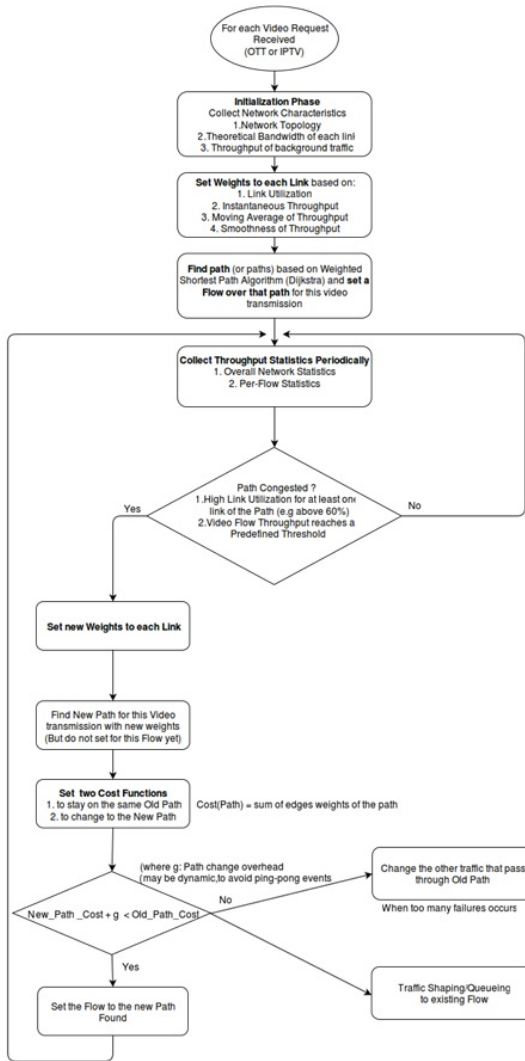
Fig. 1. Proposed SDN based video streaming optimization and QoS-aware path selection algorithm

(i.e., packet loss, path delay, link throughout, link utilization, etc.) from the network. Moreover, a flow-based routing algorithm is proposed for differentiating and prioritizing scalable video coding (SVC) base layer flows, over the SDN. Therefore the video traffic flows are differentiated between base layer streams and enhancement layer streams and then prioritized according to their impact on the perceptual quality of the received video. The proposed algorithm of Fig. 1 determines appropriate paths between one or more sources to one or more destinations, which satisfy the constraints of the video flows marked as important.

III. MEDIA AWARE PROXY

MAP is defined as a transparent user-space module responsible for low-delay adaptation and filtering of scalable video streams [10]. In conjunction with the QoE controller, which models QoE and informs MAP about the predicted level of the
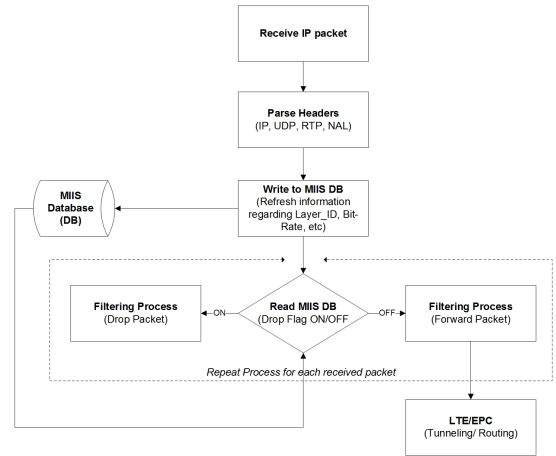

Fig. 2. Algorithmic representation of Media Aware Proxy

perceived quality, the latter is able to either drop or forward packets that carry specific layers of a stream to the receiving video users. In details, MAP is a network function based on the Media-Aware Network Element (MANE) standard [11], [12]. Briefly, MANE can be considered as either a middle box or an application layer getaway capable of aggregating or thinning RTP streams by selectively dropping packets that have the less significant impact on the user's video experience. As such, MANE has been proposed as an intermediate system that is capable to receive and de-packetise RTP traffic in order to customise the encapsulated network abstraction layer units, according to client's and access network's requirements. Within the context of the proposed media aware architecture, MAP's role is twofold. Firstly, it acts as a central point of decision in order to overcome networking limitations imposed by firewalls and Network Address Translation (NAT) protocol that are extensively used in real life networks. Secondly, it receives and parses RTP streams and customises the streaming according to the video client's requirements and network conditions, based on the QoE aware decision engine of the QoE controller.

In particular, MAP, which is designed to run in Linux kernel level in order to ensure minimum impact on the end-to-end delay, acts as a transparent proxy of the mobile client and it is responsible to parse the packets that are destined for all mobile users over multiple ports. Each received packet is forwarded in a queue and its header is parsed by an RTP parser process in order to identify the embedded video related information, without changing the header's fields. The above-mentioned algorithmic process is shown in Fig. 2.

IV. FRAMEWORK EVALUATION

A. Experimental set-up

The proposed optimization framework was evaluated based on an experimental test-bed which included a core network and an emulated LTE Enhanced Packet Core (EPC) part, as shown in Fig. 3. Specifically, the core network part consists of several OpenFlow switches that allow their switch flow
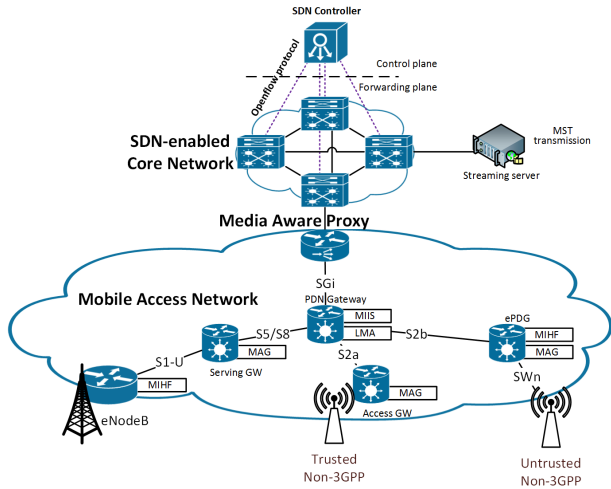
Fig. 3. Experimental test-bed

TABLE I. ENCODING AND STREAMING PARAMETERS

Video Sequence	Park Crowd	Touch Rush
No of Frames	1080 frames	
Intra period	8 frames	
QP base-enhancement	36-30	
Frame rate	30 frames per second	
Resolution	704×576 pixels	
Media Transmission Unit	1500 bytes	

table to be controlled remotely by the SDN controller. The scalable video streams were transmitted over the SDN network where different priorities where assigned to each of the flows, according to the proposed algorithm of Fig. 1. For the purposes of this research the test-bed included FloodLight v1.0 [15] as the SDN controller, OpenvSwitch (OVS) v.2.0.2 [16] for the OpenFlow switches and the OpenFlow 1.3 protocol [17] as the communication infrastructure between the SDN controller and the Openflow switches.

Moreover, the developed real-time monitoring and analysis SDN tool is independent of both the platform and the controller, while it operates on top of every SDN controller, responsible for providing dynamic QoS path allocation and path switching mechanisms to each of the video flows, according to pre-defined priorities. This could be achieved by accurate and near real-time traffic measurements across the entire SDN network, along with a number of different techniques including, flow marking, monitoring, manipulation and video throughput estimation mechanisms. It is worth mentioning that the proposed application tool provides all the above features by triggering native OpenFlow 1.3 protocol messages from the SDN controller to the OpenFlow switches, such as "*OFP_MULTIPART_REQ*" for accurate traffic measuring and "*OFP_FLOW_MOD*" for flow marking and manipulation.

Additionally, MAP is located between the core network and the Packet Data Network (PDN) gateway and manages the incoming traffic and filters the video streams according to the access network conditions. The emulated EPC part of the test-bed, comprises of virtual machines according to the EPC topology and support basic mobility management and monitoring functionalities.

B. Video Sequences

For the purposes of this study four common real-world captured stereo video test sequences ("park_joy", "croud_run", "touchdown_pass" and "rush_field_cuts"), were used [13]. All sequences had spatial resolution of 704×576 pixels at a frame rate of 30 frames per second. The first two and the last two sequences were merged together, in order to form longer video sequences. Thus the experiments could last longer, providing adequate time for the evaluating of the proposed optimization framework. The new sequences have been named as "Park Crowd" and "Touch Rush", respectively. The encoding characteristics, as well as, the streaming parameters are summarised in Table I

The H.264/SVC encoding was performed using the encoder provided by Vanguard Software tool [14] configured to create two layers, one base layer and one enhancement layer, using Medium Grain Scalability, while the group of pictures (GOP) size was 8 frames long. The SVC was the favourable choice for the particular experiments, as it allows video content to be delivered over multiple paths according to the SDN path selection algorithms and the heterogeneous wireless channels with a manageable bit rate. The particular encoder can natively compensate for packet losses up to 5% through a collection of patent protected error concealment schemes. Each video frame was encapsulated in a single network abstraction layer unit (NALU), which in turn was encapsulated into a single Real-Time Transport Protocol (RTP) packet with a maximum payload size of 1400 bytes, so as to avoid IP fragmentation. The generated video streams are transmitted through the network using the UDP/IP protocol stack. Additionally, the Parameter Sets (PS) packets are transmitted to the client through a separate TCP/IP channel in order to avoid any losses.

C. Results

The video sequences used throughout the experiments have a different behaviour in terms of throughput variations, as shown in Fig. 4. The solid line, indicates the goodput generated at the server for both sequences, while the dashed line represents the wireless medium saturation throughput for the two cases. Since both video sequences have different average and maximum throughput values, the saturation throughput was set at a different value for each. Hence, the resulted packet loss could be measured to be the same for both sequences and in the range of 5%.

Particular interest has been given to the prioritization of the base layer. Specifically, the proposed SDN flow control and path selection mechanism, assigned higher priority to the base layer, ensuring that the relevant RTP packets would flow over uncongested paths along the SDN network. In parallel, the MAP at the EPC level would selectively drop enhancement layer packets, in order to guarantee the error free delivery of the base layer. The impact of both methodologies on the
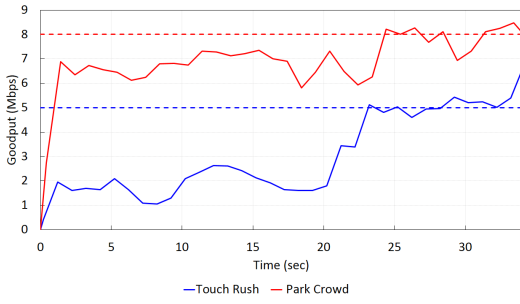
Fig. 4. Transmission data rate of both video sequences versus the wireless saturation throughput (dashed line)
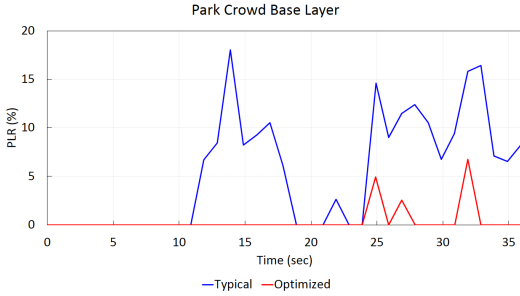

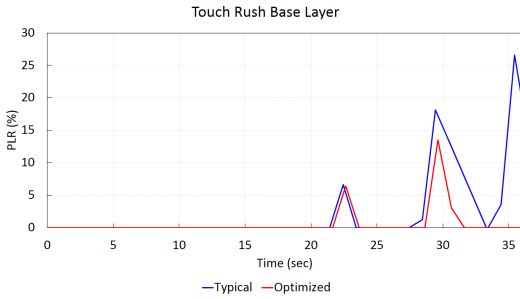Fig. 5. Base layer PLR for "Park Crowd" sequence

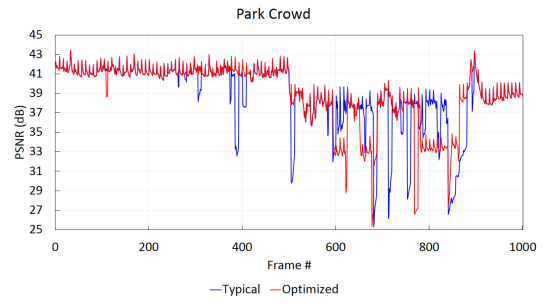Fig. 6. Base layer PLR for "Touch Rush" sequence

Fig. 7. PSNR Score for "Park Crowd" Video Sequence
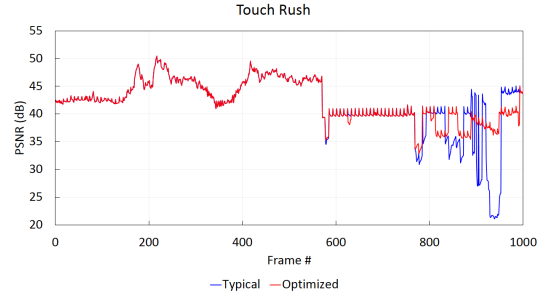
Fig. 8. PSNR Score for "Touch Rush" Video Sequence
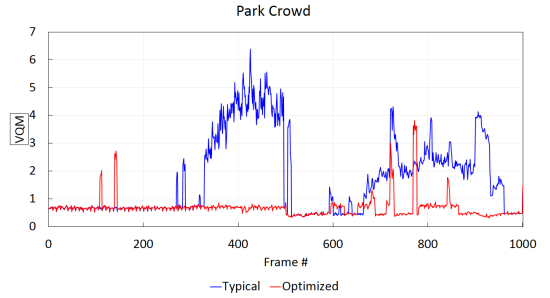
Fig. 9. VQM Score for "Park Crowd" Video Sequence

measured packet loss of the base layer packets is illustrated in Figs 5 and 6.

The performance of the proposed optimization framework is compared in terms of perceived video quality, against a baseline scenario where video packets are delivered from the SDN and the access networks as best-effort traffic. Fig. 7 and Fig. 8 compare the measured Peak Signal to Noise Ratio (PSNR) per frame, for both cases during the experiment time period. It is evident that the proposed SDN and wireless access optimization scheme achieves higher perceived quality measured in PSNR, than the baseline scenario. The results are consistent for the two video sequences. On average, the proposed framework results in higher PSNR by $1.78dB$, than the non-optimised video delivery. The performance increases in the case of "Park Crowd", where the resulted PSNR is $2.34dB$ higher than the baseline scenario.

The results are consistent and in the case of the NTIA

Video Quality Metric (VQM) [18]. Since VQM considers the impact of the blurring, jerkiness global noise, block and color distortion to the overall video quality, it correlates better with the human visual system than PSNR. Hence the use of VQM results would enhance the validity of the proposed framework's performance evaluation. In fact, the results illustrated in Fig. 9 and Fig. 10, indicate clearly the advantages of the proposed scheme. For both video sequences VQM results are significantly better (i.e., lower values than in the case of non-optimised transmission).

V. CONCLUSIONS

Evidently, a new wave of IP convergence is being fueled by real-time multimedia applications. As new network architecture paradigms are constantly evolving, there is an ongoing effort by content providers and operators to provide and maintain premium content access with optimized QoS
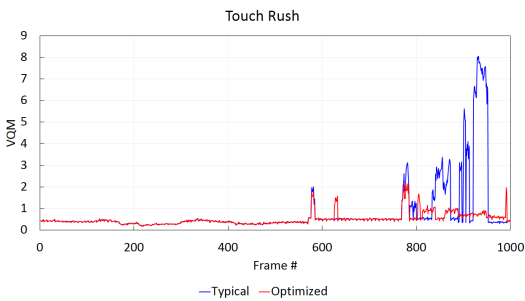
Fig. 10. VQM Score for "Touch Rush" Video Sequence

and QoE. This paper presented a novel quality-aware video optimisation scheme designed to couple SDN's dynamic video flow adaptation and path switching with a media-aware optimisation mechanism located inside the LTE EPC. The proposed synergy prioritised the base layer video packets and increased its throughput resulting in a higher than the baseline video delivery scenario, perceptual video quality. The performance of the proposed framework was evaluated using the full-referenced objective video quality metrics PSNR and VQM, which indicated the overall better video quality achieved with the proposed optimisation scheme.

ACKNOWLEDGEMENT

REFERENCES

[1] I. Politis, L. Dounis, C. Tselios, A. Kordelas, T. Dagiuklas and A. Papadakis "A model of network related QoE for 3D video," IEEE Globecom Workshops, Anaheim CA, December 2012.

[2] C. Tselios, I. Politis, K. Birkos, T. Dagiuklas and S. Kotsopoulos "Cloud for Multimedia applications and services over heterogeneous networks ensuring QoE," IEEE CAMAD, Berlin, September 2013.

[3] L. Vaquero, L. Rodero-Merino, J. Caceres, & M. Lindner, "A break in the clouds: towards a cloud definition," SIGCOMM Comput. Commun. Rev., Vol. 39, pp. 50-55, January 2009.

[4] IEEE Internet Computing, Special Issue on Cloud Computing, October 2009.

[5] https://www.opennetworking.org/images/stories/downloads/white-papers/wp-sdn-newnorm.pdf

[6] S. Shenker, M. Casado, T. Koponen, and N. McKeown, "The future of networking, and the past of protocols," Keynote at Open Networking Summit, October 2011.

[7] https://www.opennetworking.org/sdn-resources/openflow [Accessed 14/03/2015]

[8] https://www.openstack.org/ [Accessed 12/03/2015]

[9] http://opennaas.org/ [Accessed 14/03/2015]

[10] Kordelas A., Politis I., Lykourgiotis A., Dagiuklas T., Kotsopoulos S., "An media aware platform for real-time stereoscopic video streaming adaptation," in Communications Workshops (ICC), 2013 IEEE International Conference on, pp.687-691, 9-13 June 2013.

[11] IETF Draft RFC 3984, "RTP Payload Format for H.264 Video," draft-ietf-avt-rtp-rfc3984bis-12, Oct., 2010

[12] IETF Draft Audio/Video Transport WG, "RTP Payload Format for SVC Video", draft-ietf-avt-rtp-svc-27, Feb. 2011.

[13] Xiph.Org Foundation https://media.xiph.org/video/derf/

[14] Vanguard Software Solutions H.264/SVC SDK, http://www.vanguardsw.com/

[15] Floodlight http://www.projectfloodlight.org/floodlight/

[16] Open vSwitch (OvS) http://openvswitch.org/

[17] Specification, OpenFlow Switch. "Version 1.3.0", Open Networking Foundation (2012).

[18] ITU-R Rec. BT.1683, "Objective perceptual video quality measurement techniques for standard definition digital broadcast television in the presence of a full reference," 2004.