Towards QoE-driven Multimedia Service Negotiation and Path Optimization with Software Defined Networking

Andreas Kassler[1], Lea Skorin-Kapov[2], Ognjen Dobrijevic[2], Maja Matijasevic[2], Peter Dely[1]
[1] Karlstad University, Computer Science Dept., Universitetsgatan 2, 65188 Karlstad, Sweden
[2] University of Zagreb, Faculty of Electrical Engineering and Computing, Unska 3, 10000 Zagreb, Croatia
E-mail: {andreas.kassler, peter.dely}@kau.se, {lea.skorin-kapov, ognjen.dobrijevic, maja.matijasevic}@fer.hr

Abstract: **This paper presents motivation and ongoing work towards a system for Quality of Experience (QoE)-driven path assignment for multimedia services. The system goal is to enable negotiation of service and network communication parameters between end-users and to assign the network paths that are used for delivering multimedia flows according to the agreed service configuration. The key concept behind the system is a centralized multi-user optimization of the path assignments, which maximizes QoE by taking into account service utility functions, network topology, link capacities, and delay. Based on the output of the optimization process, the system implementation uses the OpenFlow to set up forwarding paths for the network elements.**

1. INTRODUCTION

When delivering multimedia services over heterogeneous networks, the impact of resource limitations is different for different media types and users' quality expectations. This requires developing mechanisms, which allow negotiation and optimization of multimedia content delivery under resource constraints according to user perceivable quality measures, or Quality of Experience (QoE).

At the service level, a multimedia service involving several media flows may be designed so that it can be adapted and delivered at different quality levels, or "configurations", each with different resource requirements. At the network level, however, in current systems all data flows between a given source and destination follow the same path, although such a path might not be the optimal one for all types of media. For example, an audio flow should be delivered over a path that minimizes delay, while a (non-real time) data flow should typically be delivered over a path that has sufficient capacity available. This requires a system that allows delivering each user media flow over the "best available" path using the "best service configuration" in order to maximize the total QoE for all users.

Implementing a service that jointly optimizes the configuration of the network elements and the end-hosts is difficult, as it requires grained control over the packet forwarding based on application requirements. To this end we propose a new architecture within the framework provided by Software Defined Networks (SDNs), which offer hardware vendor-independent APIs to implement such functionality. The paper reflects work in progress, starting from the overall motivation and high-level general outline of the proposed system, followed by an illustrative example, towards implementation and future work. Following a brief overview of background and related work, Section 3 describes the proposed architecture, its key components and their functionality. Section 4 presents an example and Section 5 concludes the paper by presenting directions for future work.

2. BACKGROUND AND RELATED WORK

2.1 QoE-driven service negotiation and optimization

The recent understanding of user quality perception has resulted in QoE as a concept that goes beyond accounting for the influence of only "traditional", QoS-based technical parameters. While numerous definitions of QoE can be found in literature and standards, a recent (working) definition that has emerged from the COST Action IC1003: European Network on Quality of Experience in Multimedia Systems and Services (QUALINET) community emphasizes that QoE results from meeting an end user's expectations regarding the utility and/or enjoyment of an application or a service, considered in light of the user's personality and current state [1]. Hence, QoE is influenced by a wide range of factors, including those related to the user, service, content, network, device, application, and context of use.

For the purposes of making QoE-driven network resource allocation decisions, it is necessary to understand the correlation between network performance and user perceived quality, while also accounting for the influence of additional previously mentioned factors (e.g., user- and context-related). As a concept adopted from economics, utility functions have often been used as a formal mathematical means for expressing an end user's degree of satisfaction with respect to corresponding multi-criteria service performance [2]. Different types of utility curves, portraying utility as a function of bandwidth, have been shown to correspond to different types of traffic, e.g., elastic and adaptive vs. non-elastic traffic types (as observed by Shenker in [3]). Audio and video traffic usually have adaptive utility functions whereby certain delay/loss may be tolerated (as long as it remains below a certain threshold). The utility curve, in general, depends on application-level configuration parameters, such as used codecs, formats, resolutions, frame rates, etc. In the case of multimodal sessions with multiple media components (e.g., audio and video), different utility functions may correspond to each media component, with overall session utility expressed as some form of a weighted combination thereof.

Following the modeling of QoE, challenges lie in employing such models in the network for the purposes of QoE optimization and control. While QoE optimization may be considered from the perspective of a single user session taking into account current terminal, network, and service constraints, multi-user domain-wide QoE optimization problems involve making domain-wide resource allocation decisions across multiple sessions [4][5].

To provide an input for such optimization problems, feasible session parameters (e.g., in terms of feasible media components, codecs, resolutions, etc.) need to be negotiated between actors involved in the service delivery chain (e.g., end users, service providers, network operators). For example, end users communicating via a multimedia session while located in heterogeneous access networks and using different devices will need to agree on a set of session parameters that are supported by both users and authorized by the underlying network(s). The IETF's Session Initiation Protocol (SIP) [6] and Session Description Protocol (SDP) [7] provide application-level signaling mechanisms which have been adopted by standards organizations including ETSI/TISPAN and 3GPP for the purposes of multimedia session negotiation, establishment, and modification.

In the context of the next generation network (NGN) architecture, previous work has proposed inclusion of different generic application-level network functions involved in negotiation and QoE-optimization decision making [8][9]. In this paper, we apply this concept in the context of software defined networking, as described next.

2.2 Software Defined Networking

Software Defined Networking (SDN) [10] is a new networking paradigm in which the data and control planes are decoupled, and the control (which used to be tied to a particular infrastructure element, and thus be vendor and device specific) is logically centralized and programmable via standardized interfaces. Through implementing their own rules and policies in software and deploying them in an abstracted, virtualized network infrastructure, network operators can achieve flexible control over network services, such as, e.g., routing, traffic engineering, and QoS.

SDN architecture (Figure 1) can logically be viewed as having three layers: 1) an infrastructure layer, 2) a control layer, and 3) an application layer. The OpenFlow protocol, promoted by the Open Networking Foundation (ONF), provides the first standardized interface between the SDN controllers in the control plane and the infrastructure elements. ONF is also working on open APIs between the SDN control and applications layers. What these provide are the means for the applications to use network services and capabilities as needed, without knowing the network specifics, such as network topology. Applications on the application layer can thus issue requests, which are translated by the control layer to device specific configurations.

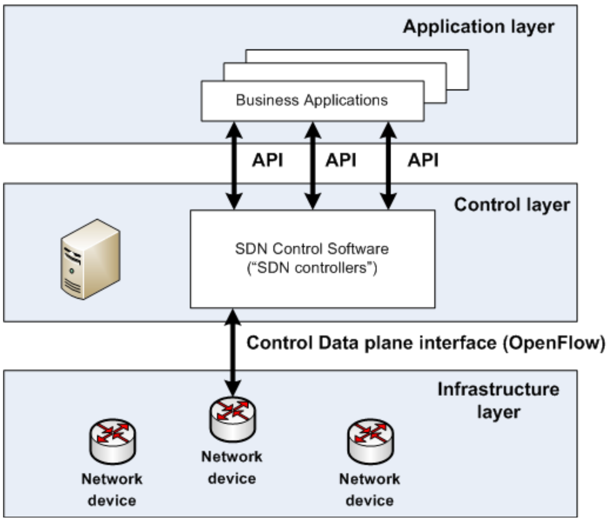

Figure 1 - SDN architecture

As the control layer in SDNs has a global view on the network topology graph, it is possible to implement control applications that use traditional graph optimization algorithms for traffic engineering. For example, Ethane [11] is a centralized control system to optimize traffic flows in data centers. An SDN-based architecture to optimize flows and user associations in wireless mesh networks has been proposed in [12]. Balancing the load between web-servers by using OpenFlow, while considering the available network capacity, has been presented in [13].

While the works mentioned optimize flows with the goal to maximize the network throughput, increased throughput does not necessarily correspond with a better user perceived quality, or QoE. Hence, to maximize QoE, an understanding of the correlation between QoS, i.e., the network-related factors, and QoE is necessary. The joint consideration of end-user QoE and path assignment at the infrastructure layer may open the possibility of optimizing paths for multimedia flows so as to maximize QoE. In the following section we present a novel architecture, which combines service negotiation at the application level, with the concept of SDN, leading to optimal path assignment.

3. THE PROPOSED ARCHITECTURE

The proposed architecture is illustrated in Figure 2. The key enablers for the proposed functionality are the *QoS Matching and Optimization Function* (QMOF) and the *Path Assignment Function* (PAF).

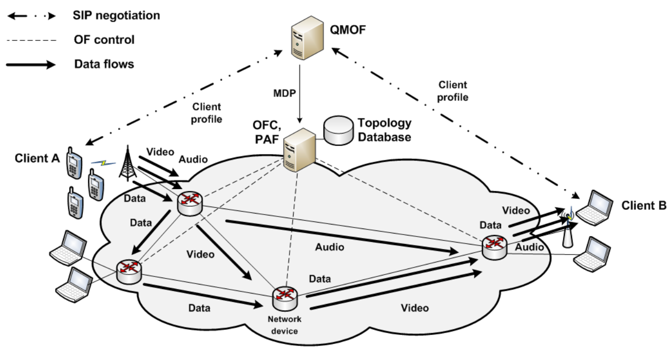

Figure 2 – The proposed architecture for QoE-driven service optimization and path assignment

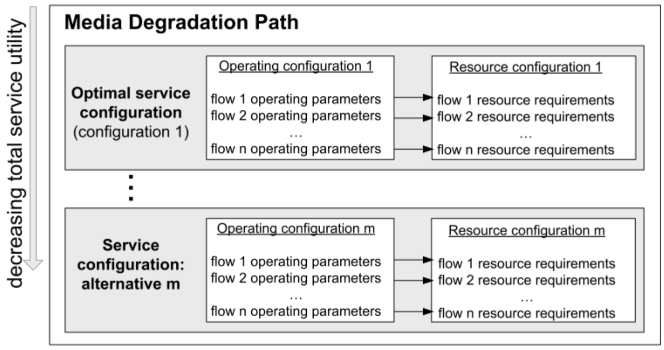

Figure 3 – An example of a session Media Degradation Path

The QMOF resides in the SDN application layer, while the PAF is located in the SDN control layer. We note that there are two complementary optimization procedures that we refer to: the first involves calculation of a Media Degradation Path (MDP) and it is conducted by the QMOF, while the second involves the calculation of optimal path assignment and it is conducted by the PAF.

3. 1 QoS Matching and Optimization Function

The QMOF has been previously proposed as a generic and reusable service capability, supporting optimized session delivery and controlled session adaptation by being included along the end-to-end (E2E) signaling path [14]. The role of the QMOF in session negotiation is as follows. Upon initiating a service request, user capabilities, preferences and service utility mappings (specifying the resource requirements for individual media flows) are signaled by the end users and collected by the QMOF. The QMOF conducts an initial parameter matching process to determine feasible service configurations, followed by a utility-based optimization process used to determine the optimal service configuration and corresponding resource requirements for the given service session, in light of user, service, and network constraints. The overall session utility (QoE) is expressed as a weighted combination of individual media flow utility functions based on parameters such as individual user preferences, usage context, and service content.

The optimal service configuration specifies the configuration of flow operating parameters (e.g., frame rate, codec), resource requirements (e.g., bit rate), and a utility value that represents a numerical estimation of the particular configuration's QoE. Besides the optimal configuration, several suboptimal configurations are calculated and ordered by their decreasing utility value, thus forming a *Media Degradation Path* (MDP).

The goal of the MDP is to serve as a "recipe" for controlled service adaptation, thereby achieving maximum possible utility in light of dynamic conditions. For example, in the case of a user indicating that he/she prefers video over audio for a given audiovisual service, an MDP may be constructed so as to first degrade video quality in case of a decrease in network resource availability, while maintaining high audio quality. Hence, in case of decreased resource availability, a suboptimal configuration can be activated (thus preventing unpredictable degradation of a service). An example of an MDP is given in Figure 3.

Following its calculation, the MDP is then forwarded to the Path Assignment Function (PAF), which subsequently optimizes the network paths in order to meet the resource requirements of a currently active service configuration.

3.2 Path Assignment Function

The PAF is an application executed on the OpenFlow controller (OFC). The PAF maintains a network topology data base, which contains the network connectivity graph along with link capacities. The topology database is built using the topology discovery mechanisms provided by OpenFlow. Furthermore, the PAF, as part of the OFC, has an up-to-date view on the currently active flows in the network.

The PAF exposes an API, which the application layer uses to inform the PAF about the MDP referring to a set of flows in a session. In particular, the QMOF reports the MDP of a new session upon its establishment. The PAF then optimizes the paths routes. The aim of the optimization is to maximize the session QoE, while taking into account the quality degradations described in the MDP. Using the configurations in the MDP, the PAF will try to assign such paths to each of the flows in the given configuration, so as to meet the specified service requirements, starting from trying to meet the requirements of the optimal configuration and then moving down along the MDP. If no solutions are found, an option would be to conduct a global optimization process taking into account the MDPs of multiple sessions active in a given network domain, with the goal of reassigning paths so as to be able to admit incoming sessions.

The path assignment problem can be formulated as a variant of the min-cost multi-commodity flow problem, where the costs are the negative QoE values. For this class of problem a large universe of solution methods can be found in literature [15]. Solving the optimization yields the network paths and rates for each flow.

To implement the solution in the network, the PAF configures the forwarding tables of the network devices using the OpenFlow protocol, similar to the approach presented in [12]. With the OpenFlow protocol the PAF can specify at each network device, which is the next-hop of a flow and what the minimum rate is.

Except for OpenFlow support, our architecture does not require the network devices to have any other special capabilities. The PAF is therefore independent of the used link layer technology. A similar functionality – link-layer independent QoE-optimized flow routing – could be achieved by using overlay networks [16]. However, our architecture is more lightweight and flexible, since no overlays need to be maintained and the routing does not depend on any overlay structures.

4. EXAMPLE: PATH ASSIGNMENT FOR FLOWS IN A MULTIMEDIA SESSION

This section presents an example scenario to illustrate the workings of the proposed architecture. The scenario consists of establishing a multimedia session between two users, Alice and Bob. The session involves multiple media flows, which are routed over different network paths to satisfy their requirements regarding network resources and optimize overall session QoE. The feasibility of a specific media flow quality settings are determined at the QMOF, based on user preferences, terminal capabilities (in terms of hardware and software configuration), service requirements and capabilities, access network characteristics, and current availability of network resources.

Alice uses a smart-phone (*Client A*) to connect to the network over a High-Speed Packet Access (HSPA) link, while Bob connects his laptop computer (*Client B*) to the network via a wireless local area network (WLAN). The clients A and B both run a conferencing tool that allows them to establish audio/video conferences, to exchange files, and to send instant messages. The conferencing software employs SIP/SDP signaling to negotiate a set of feasible session flows and their parameters among the users. The signaling diagram that describes high-level control interaction among the proposed architecture elements is depicted in Figure 4.

Alice requests an audio/video conference to be set up with Bob, and her client issues an SIP INVITE request for session establishment, which is used for conveying Alice's user profile to the QMOF and then forwarded to client B (steps 1–2 in Figure 4). Client B responds with a SIP response to Client A, which delivers Bob's user profile to the QMOF (step 3). The session establishment sequence triggers the QMOF to perform the matching process (step 4), and offer the set of potential session parameters (from feasible service configurations) to the users A (steps 5–6) and B (steps 7–8).

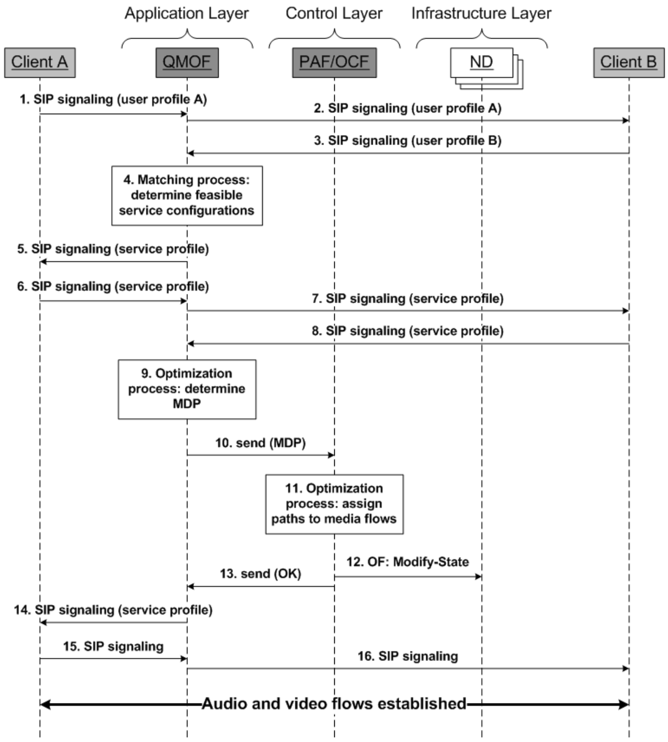

Figure 4 – Simplified signaling diagram for the example

(The user may accept or refuse the (subset of) offered parameters.) From the given user, service, and network constraints, the QMOF determines the MDP for the given session (step 9), which includes several service configurations, ordered by their utility values (corresponding to user-perceived quality). The configurations are produced in such a way to consider all media flows that may be used in the conferencing session, i.e., audio, video, and data (file transfer) flow, even if some of them are not requested in the establishment phase. This relieves the system from the need to invoke the (re)negotiation process for every addition/removal of a flow that may occur later during the ongoing session. In the current example, the users have requested an audio-visual session and a service configuration with active audio and video flows is to be applied. Once it calculates the MDP, the QMOF uses an application-to-control plane API to forward the MDP to the PAF (step 10).

Based on service configurations from the MDP, the network topology database, and the currently active flows established by other users, the PAF assigns paths for the audio and the video flows so as to maximize QoE for the audio-visual session (step 11). In this scenario, audio flow is delivered over a network path with the minimal delay, while video flow is assigned the path that combines minimal delay and sufficient capacity (as illustrated in Figure 2). The PAF then configures forwarding tables of the network devices (NDs) via the OF protocol (step 12) and confirms the configuration (step 13). Finally, SIP signaling continues (steps 14–15–16) in order to initiate the service delivery to

Clients A and B, according to the agreed service profile, and allowing media flows to be established over their respective paths.

If at a later point in time Alice decides to share a file with Bob, the Client A will send a SIP re-INVITE request to add data (file transfer) flow to the ongoing session. This SIP request is delivered to the QMOF, which invokes the PAF to recalculate path assignment by using the existing MDP. For the purposes of this scenario, the PAF assigns a network path with sufficient capacity available to the data flow and, if needed, modifies the rules at the NDs accordingly to include the new flow.

5. CONCLUSIONS AND FUTURE WORK

We have described the motivation and the idea how software defined networking could be used to perform optimized path assignment, while achieving the desired level of QoE from the users' perspective. Since the SDN controller has a global view of the network domain, both optimization goals may be addressed simultaneously. Our current work is focused on formulating and solving the multi-user domain-wide QoE optimization problem. We also plan to implement the proposed system in an emulated network environment to create a proof-of-concept prototype and to be able to run performance experiments with diverse multimedia sessions.

ACKNOWLEDGMENT

This work has been supported by the European Community's 7th Framework Programme under grant agreement no. 285939 (ACROSS), and the research projects 036-0362027-1639 and 071-362027-2329 funded by the Ministry of Science, Education, and Sports of the Republic of Croatia.

REFERENCES

[1] –, Qualinet White Paper on Definitions of Quality of Experience, Output version of the Dagstuhl seminar 12181 (Version 1.1, Dagstuhl, June 2012), available at http://www.qualinet.eu/.

[2] Peter Reichl, Bruno Tuffin, Raimund Schatz: "Logarithmic laws in service quality perception: where microeconomics meets psychophysics and quality of experience", Telecommunication Systems, Springer, June 2011, pp. 1-14.

[3] Scott Shenker: "Fundamental design issues for the future Internet," IEEE Journal on Selected Areas in Communications, Vol. 13, No. 7, pp. 1176-1188, September 1995.

[4] Krunoslav Ivesic, Maja Matijasevic, Lea Skorin-Kapov: "Simulation Based Evaluation of Dynamic Resource Allocation for Adaptive Multimedia Services", Proc. of the 7th International Conference on Network and Service Management (CNSM), October 2011, pp. 1-8.

[5] Srisakul Thakolsri, Shoaib Khan, Eckehard G. Steinbach, Wolfgang Kellerer: "QoE-Driven Cross-Layer Optimization for High Speed Downlink Packet Access", Journal of Communications, Vol. 4, No. 9 (2009), pp. 669-680, Oct 2009, doi:10.4304/jcm.4.9.669-680.

[6] Jonathan Rosenberg et al.: "SIP: Session Initiation Protocol," IETF RFC 3261, June 2002.

[7] Flemming Andreasen: "Session Description Protocol (SDP) Capability Negotiation", IETF RFC 5939, Sept. 2010.

[8] Lea Skorin-Kapov, Miran Mosmondor, Ognjen Dobrijevic, Maja Matijasevic: "Application-level QoS Negotiation and Signaling for Advanced Multimedia Services in the IMS", IEEE Comm. Mag, Vol. 45 No. 7, July 2007, pp. 108-116.

[9] Janez Sterle, Mojca Volk, Urban Sedlar, Janez Bester, and Andrej Kos: "Application-Based NGN QoE Controller", IEEE Comm. Mag., Vol. 49, No. 1, Jan. 2011.

[10] –, Open Networking Foundation, "Software-Defined Networking: The New Norm for Networks", ONF White Paper, 2012.

[11] Martin Casado, Michael J. Freedman, Justin Pettit, Jianying Luo, Nick McKeown, and Scott Shenker: "Ethane: taking control of the enterprise", Proc. of SIGCOMM 2007, ACM, pp. 1-12.

[12] Peter Dely, Andreas Kassler, Nico Bayer: "OpenFlow for Wireless Mesh Networks", Proc. of the 20th Int. Conf. on Computer Communications and Networks, 2011, pp. 1-6.

[13] Richard Wang, Dana Butnariu, and Jennifer Rexford: "OpenFlow-based server load balancing gone wild", Proc. of Workshop on Hot Topics in Management of Internet, Cloud, and Enterprise Networks and Services (Hot-ICE), Boston, MA, March 2011, 6 pp.

[14] Lea Skorin-Kapov, Maja Matijasevic: "Modeling of a QoS matching and optimization function for multimedia services in the NGN", Proc. of the 12th IFIP/IEEE Int. Conf. on Management of Multimedia and Mobile Networks and Services, 2009, pp. 55-68.

[15] Ravindra K. Ahuja, Thomas L. Magnanti, and James B. Orlin: "Network flows: theory, algorithms, and applications", Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1993.

[16] Bart De Vleeschauwer, Filip De Turck, Bart Dhoedt, Piet Demeester, Maarten Wijnants, and Wim Lamotte: "End-to-end QoE Optimization Through Overlay Network Deployment", Proc. of the 22nd IEEE International Conference on Information Networking (ICOIN 2008), Busan, Korea, January 2008, 5 pp.