Extending QoS support in the IP multimedia subsystem: Mobility-aware session reconfiguration

Ognjen Dobrijevic, *Member, IEEE,* and Maja Matijasevic, *Senior Member, IEEE*

***Abstract*—This paper presents an approach to controlling service continuity and quality of service (QoS) in the 3rd Generation Partnership Project (3GPP) IP multimedia subsystem (IMS). The approach extends existing support to manage QoS by dynamically producing multimedia session configurations with respect to different mobility types, which is referred to as *mobility-aware session reconfiguration*. We first give a state-of-the-art overview regarding IMS support for service continuity and QoS management. Then, we propose to enhance QoS support in IMS by introducing a Session Initiation Protocol (SIP) application server that is responsible for (1) generating QoS specifications which conform to the mobility-induced constraints and (2) reconfiguration decision-making. Moreover, we describe several SIP signaling procedures which enable the session reconfiguration, and demonstrate the application of our approach in an IMS laboratory prototype, which also serves for a performance evaluation of the solution.**

***Index Terms*—Multimedia communication, Quality of service, Mobile communication.**

I. Introduction

WITH the widespread use of multimedia services, deployment of network architectures for their provisioning has gained momentum. One of such architectures is the Internet Protocol (IP) multimedia subsystem (IMS) [1]. A critical aspect of service provisioning is quality of service (QoS) support, which enables service-level negotiation of QoS parameters and transport-level control of network resources based on the agreed service parameters [2]. By deploying IP networks that adhere to the Next Generation Network (NGN) concept [3], including IMS, the providers are committed to enable users to communicate and access services independently of changes that may stem from different mobility types, e.g. *terminal mobility* and *session mobility*. The former provides uninterrupted communication when a user terminal switches between access networks of different technologies (called *vertical handover*), while the latter allows to seamlessly move multimedia sessions between user terminals. To facilitate QoS continuity that meets such a prerequisite, management approaches also include means to adapt service to mobility.

The 3rd Generation Partnership Project (3GPP) IMS [4] has been specified as a network subsystem that offers session control for multimedia service provisioning, with the control being based on the Session Initiation Protocol (SIP) [5]. However, current 3GPP specifications reveal some limitations regarding QoS support. The basic IMS specification [4] describes procedures for the session end-points to negotiate session parameters, which is based on the simple offer-answer model [6]. On the other hand, the 3GPP-defined QoS framework [7], which also applies to IMS, focuses on resource allocation and QoS provisioning at transport level in access networks. We argue for an approach to managing QoS that would enable:

- service delivery that is tailored to match user preferences, user terminal capabilities, service requirements, and access network characteristics;
- reservation of network resources that is optimal in, e.g., sharing bandwidth across media components; and
- service adaptation to different mobility types in order to regulate QoS or prevent its degradations.

The goal of QoS management is to apply “the best” session configuration and network resources allocation that maximize QoS parameters in terms of, e.g., bandwidth and delay.

Our approach envisages the following use case scenarios. Users access an IMS network to establish multimedia services, e.g., video-on-demand, or audio-video conferencing, which are hosted by application servers. Each user may utilize different terminals and may move the ongoing sessions between them.

1) During the establishment phase, session parameters are negotiated between user’s terminal and the related application server (AS), and include feasible media components, their characteristics, and the needed QoS specification. This leads to the resources reservation.
2) When a user decides to replace her/his terminal, session parameters are reconfigured to meet capabilities of the targeted terminal. The reconfiguration may, e.g., modify media encoding parameters and produce a new QoS specification, thus leading to the adjustment of resources allocation before moving media components.
3) When a user’s terminal changes its location, session may be reconfigured by moving its media components to another AS instance in order to maintain QoS. AS instances are deployed for the load balancing purposes and are associated with distinct locations.
4) When a user’s terminal changes access network, session parameters are reconfigured to match capabilities of the new access technology. The reconfiguration is completed after allocating network resources, possibly based on a newly negotiated QoS specification.

To the best of our knowledge, there is no a comprehensive IMS-based solution that addresses all these scenarios (e.g., the IMS service continuity specification [8] defines procedures that maintain a service with regards to scenarios 2) and 4), but it does not regulate how to facilitate QoS continuity).

O. Dobrijevic and M. Matijasevic are with the University of Zagreb, Faculty of Electrical Engineering and Computing, Unska 3, 10000 Zagreb, Croatia e-mail: {ognjen.dobrijevic, maja.matijasevic}@fer.hr.

Manuscript received *May 14*, 2012; revised *August 1*, 2012.

This paper presents our approach to enhancing QoS support in the 3GPP IMS. By dynamically producing session configurations with respect to session and terminal mobility, we are able to maintain or adapt QoS across different user terminals and access networks. We outline requirements for implementing the session reconfiguration in IMS, which are derived from the use case scenarios and our generic session reconfiguration model (presented in a previous work [9]). We propose to introduce an SIP AS that (1) generates QoS specifications conforming to the mobility-induced constraints and (2) decides on the reconfiguration. This SIP AS, named the Session Configuration Management (SCM) AS, also controls IMS service continuity. The SCM AS can be considered a reusable service offered by the IMS network, which would relieve equipment manufacturers and service providers of implementing its specific functionalities. We design several SIP procedures, which are based on the 3GPP specifications, to enable the session reconfiguration. We also demonstrate the application of our approach in an IMS laboratory prototype, which also serves for a performance evaluation of the solution.

The remainder of this paper is organized as follows. An overview of service continuity and QoS support in IMS is given in Section 2. In Section 3, we outline requirements for implementing the session reconfiguration in IMS. Section 4 presents the SCM AS and the IMS prototype. We demonstrate the approach and analyze its performance evaluation in Section 5, followed by the conclusion section.

II. Service continuity and QoS support in IMS

A. 3GPP specifications

The 3GPP IMS is an NGN-compliant architecture, which was designed to be independent of the access network technologies. A simplified view of the 3GPP IMS architecture is given in Figure 1. To control multimedia sessions, IMS defines a number of functional entities. The main entities are:

- IMS User Equipment (UE), which issues requests for session establishment and modification;
- SIP AS, which enables introduction of new services to the IMS network by hosting and executing them;
- Proxy-Call Session Control Function (P-CSCF), which represents the first contact point for a UE on the signaling path towards the rest of IMS core entities;
- Serving-CSCF (S-CSCF), which offers a coordinated SIP interaction among IMS entities and selects an SIP AS depending on the service to be invoked; and
- Home Subscriber Server (HSS), which stores subscription information to authenticate and authorize users, and information related to the user's location and IP address.

The 3GPP specifications consider different aspects of the overall QoS support in IMS. Technical specification (TS) 23.228 "IP Multimedia Subsystem (IMS)" [4] provides a high-level description of SIP procedures for the session end-points to negotiate multimedia session parameters. The specification elaborates on how to determine media characteristics in the session establishment phase or when a session is modified in the context of, e.g., adding a media component or changing bandwidth requirements. The negotiation procedure is based on the simple offer-answer model [6], which proposes a mechanism for the end-points to reach a common view of the session. In the model, one end-point offers a set of desired session parameters to the other, while the other end-point answers with the session parameters that are wanted from its perspective. Such a procedure may take multiple negotiation steps until the media characteristics are agreed upon.

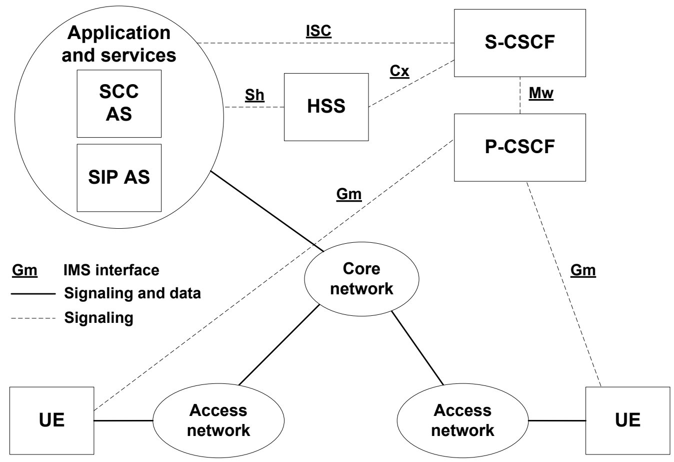

Fig. 1. A simplified view of the 3GPP IMS architecture

TS 23.207 "End-to-end QoS concept and architecture" [7], on the other hand, defines the 3GPP QoS framework, which also applies to IMS. The framework, however, focuses on QoS management and resources allocation at the IP-bearer level and the transport-bearer level of different access networks.

TS 23.237 "IMS Service Continuity" [8] provides a high-level description of service-level (SIP) procedures that transfer a session between different access networks (the specification refers to this process as the *access transfer*, AT) or different UEs (the *inter-UE transfer*, IUT). This specification employs two functional entities to execute the service continuity mechanisms: (1) the SCC AS, which is an SIP AS (Figure 1), and (2) a UE with the associated support. Multimedia sessions started by UEs are anchored at the SCC AS, which uses the 3rd party call control (3pcc) mechanism to facilitate session transfer. The SCC AS is inserted in the SIP signaling path and selected by an S-CSCF to control the AT or the IUT, while the UE initiates the transfer procedures. The AT is triggered based on the criteria such as operator policy, user preferences, and access network conditions, while user input starts the IUT.

B. Other related work

The related QoS research efforts, for which a summary and a comparison are given in Table I, mainly differ in two aspects. The first one regards different degrees of QoS support offered:

1) service continuity is provided, but without any QoS guarantees;
2) resource allocation is performed, but no form of QoS negotiation is included;
3) QoS negotiation is employed solely in a service set-up phase; and
4) QoS adaptation is supported during the course of a service lifetime.

TABLE I
A SUMMARY AND A COMPARISON OF RELATED WORK

Approach	Considered deployment levels	Degree of QoS support	Regarded mobility types	Supported service customization
Y. C. Yee et al. [10]	Service level	Service continuity	Terminal mobility	Move sessions among user terminals, change network attachment point for the terminals, switch media codecs
E. Cerqueira et al. [11]	Transport level	QoS adaptation	Terminal mobility	Remove/add media components of different priority from/to sessions, assign different QoS classes to sessions
K. S. Munasinghe and A. Jamalipour [12]	Service level	Service continuity	Terminal mobility	N/A
M. Rawashdeh and A. Karmouch [13]	Service level	Service continuity	Session mobility	Change video framerate
P. Bellavista et al. [14]	Service level	Service continuity	Terminal mobility	N/A
W.-K. Chiang and P.-C. Kuo [15]	Service level	Service continuity	Terminal mobility	N/A
M. Navarro and Y. Donoso [16]	Transport level	Resource allocation	N/A	Assign different QoS classes to sessions
S.-R. Yang and W.-T. Chen [17]	Service level + transport level	QoS negotiation	Terminal mobility	N/A
T. Renier et al. [18]	Service level + transport level	QoS negotiation	Terminal mobility	N/A
J. Liao et al. [19]	Service level + transport level	QoS adaptation	Terminal mobility	Change point of access network attachment for user terminals
L. Skorin-Kapov et al. [20]	Service level + transport level	QoS adaptation	None	Add/remove media components to/from sessions, switch media codecs, adapt allocation of network resources

While most of the referred solutions do not employ QoS adaptation, those that facilitate it focus on control procedures at transport level and consider a limited set of parameters to be adjusted. The other research efforts' aspect relates to handling distinct mobility types – a majority of the approaches is centered on either terminal mobility or session mobility.

The *Proactive and Adaptive Handover* (PAHO) system [10] is an SIP-based approach that customizes service configuration in the event of network performance degradation caused by terminal mobility. To maintain service continuity, the PAHO, e.g., moves media components among user terminals or switches media codecs. The PAHO, however, does not include QoS negotiation and resources reservation. The *Multi-User Session Control* (MUSC) approach [11] regulates session QoS parameters in response to terminal mobility and network performance variations. The MUSC employs transport-level procedures to coordinate QoS adaptation, which leads to removing or adding "lower-priority" media components, or, assigning different QoS classes to a session. It enables to match session QoS requirements with the available QoS classes, but does not support service delivery that adjusts to user preferences and terminal capabilities, as well as handling of session mobility.

An IMS platform that controls mobility between a Wireless Local Area Network (WLAN) and a Universal Mobile Telecommunications System (UMTS) network is presented in [12]. It offers service continuity with a make-before-break type of handover between the two network types, but does not allow to adjust QoS parameters to the targeted network characteristics nor to allocate necessary resources. A seamless video transfer for session mobility in IMS is described in [13]. It focuses on minimizing the disruption time when moving a video session between UEs, but does not tackle the issue of providing QoS. The IMS-compliant Handover Management AS (IHMAS) [14] is introduced in the IMS network to achieve session continuity upon a vertical handover. This solution proactively triggers SIP signaling with the targeted access network by having UEs predict the handover. A similar SIP AS approach, referred to as the Centralized Service Continuity (CSC), is presented in [15]. The continuity is managed by the CSC AS, which, similarly to the SCC AS and the IHMAS, acts as the session anchor point and performs the 3pcc for session re-establishment. However, the last two approaches do not provide any QoS guarantees.

An IMS-centered enhancement of the 3GPP QoS framework is presented in [16]. Its transport-level approach reassigns QoS classes to the sessions based on the network state and resource availability. Another QoS framework for IMS is proposed in [17]. It alleviates the influence of terminal mobility on QoS by having UEs trigger resources reservation at "neighboring" IMS networks which they may visit during service execution. This way, QoS agreements from the service set-up phase tend to be preserved, but with possibly a large waste of the resources that will not be used. An approach for IMS that offers session continuity in response to vertical handover is described in [18]. It enables delivery of agreed QoS parameters between P-CSCFs that control different access networks, but assumes that network conditions remain the same after the handover.

An improvement of the framework presented in [17] is an enhanced IMS handover mechanism (EHM) [19]. The EHM employs a mobility prediction algorithm to detect a UE's movement between network attachment points. Before the UE moves to a new IMS domain, the EHM chooses the most appropriate access network, which leads to reserving resources in advance. Still, this approach lacks both the means to react to session mobility and parameters such as UE capabilities, service requirements, and budget constraints when customizing service delivery. Our previous work [20] proposes an SIP AS for IMS that hosts a function for matching communication requirements of the parties involved in service establishment and for calculating the reservation of network resources that is optimal in distributing them among the media components. The matching function was designed to assist in the QoS negotiation by producing an offer of feasible media components for a session and their characteristics. However, this work does not consider a service adaptation to mobility.

III. Proposed enhancements to QoS support in IMS

After presenting an overview of IMS support for service continuity and QoS management, we now focus on requirements regarding the session reconfiguration and meeting them in IMS. The session negotiation procedure in IMS resides on a simple matching among capabilities and requirements of multimedia session's end-points, which must agree on session parameters such as type and encoding of media components. In order to provide QoS guarantees by applying controllable values of network performance indicators, the associated QoS support must somehow map session requirements to, e.g., expected network bandwidth and delay.

As mentioned in the previous section, we have proposed a common function, named the QoS Matching and Optimization Function (QMOF), that produces session configurations based on user preferences, user terminal capabilities, multimedia service requirements, and access network constraints [20]. A session configuration is feasible when it meets these criteria:

1) user terminal capabilities comply with the processing requirements of desired media components;
2) access network constraints (e.g., available bandwidth and delay) support the minimum requirements on network performance for desired media components; and
3) user preferences, such as respective relevance of media components, are fulfilled.

The purpose of the matching functionality is to enhance the negotiation procedure by offering a number of potential configurations for a particular session, with all of them meeting the mentioned criteria, but differing in calculated parameter values. Another aspect of the QMOF relates to determining the optimal QoS reservation of network resources across media components of the session (the optimization objective may be formulated in various ways and specified by, e.g., the network operator). The interested reader is referred to our previous work [20] for details on the optimization and the QMOF.

Furthermore, we argue for additional mechanisms that would improve IMS support beyond service continuity and facilitate QoS continuity in the event of session mobility and terminal mobility. Besides offering procedures to control session transfer between different user terminals and different access networks, we believe that IMS support should include the means to decide on how a session is reconfigured in response to mobility-induced changes that emerge during service execution. For instance, if the available access bandwidth decreases after a vertical handover, the reconfiguration could lead to reducing bandwidth demands for all the session's media components or to removing a media component from the session in accordance to user preferences. Moreover, if capabilities of the user terminal improve as a result of session mobility (e.g., replacing a mobile phone with a personal computer), a session could be reconfigured to apply codecs with higher bit rates.

Our generic session reconfiguration model [9] presents requirements and functional entities in the context of NGN [3] that are needed to control session parameters and manage QoS. The model is based on two main concepts, the one of a *mobility event*, and the other of a *reconfiguration primitive*. Mobility events represent changes stemming from mobility:

1) *Change of terminal* – a change of the user terminal due to session mobility;
2) *Change of location* – a change in user terminal's location due to terminal mobility; and
3) *Change of access network* – a change of the terminal's access network due to the vertical handover.

When a mobility event occurs, the associated *mobility event notification* carries context information regarding the change, which is used to direct the reconfiguration process. One of the key functions in the reconfiguration model is the Session Reconfiguration Function (SRF), which analyzes delivered event notifications and decides on the reconfiguration primitive(s) to be executed. A result of this decision may, as well, involve invoking the QMOF, e.g., if new session configurations need to be produced. The reconfiguration primitives are management operations which are executed to regulate session configurations: (1) *start media flow*, (2) *modify media flow*, and (3) *stop media flow* (the reconfiguration model distinguishes among media components, which it refers to as media flows).

3GPP defines the IMS architecture in a way to enable multimedia services to take advantage of common IMS functions and different service enablers via standardized interfaces. To meet the outlined requirements and implement the session reconfiguration in the 3GPP IMS, we exploit the IMS mechanisms currently specified, and propose to combine the QMOF and the SRF on an SIP AS that we name the Session Configuration Management (SCM) AS. The SCM AS would be responsible for producing session configurations during the service establishment phase and for customizing the configurations in order to adapt them to mobility-induced constraints (i.e., changes in UE capabilities and access network characteristics that follow after the mobility). One of its most important functional aspects is calculation of QoS specifications that define required network performance in terms of bandwidth, delay, jitter, and packet loss ratio. Another aspect relates to deciding on how to perform the reconfiguration.

As IMS specifies functional entities rather than network nodes, we also propose to integrate functionalities of the SCC AS [8] into the SCM AS, which would allow it to control the service continuity as defined by IMS. Being accessible over the IMS Service Control (ISC) interface, the SCM AS can be regarded to as a reusable common service offered by the IMS network. This would relieve UE manufacturers and service providers of implementing specific and rather complex QMOF and SRF functionalities.

IV. The IMS laboratory prototype

To demonstrate the application of our approach, we implement an IMS laboratory prototype with the SCM AS (Figure 2). For the demonstration purposes, we design several SIP signaling procedures, which build on the IMS session control procedures (described in [4], [21], [22], [8]), to enable the session reconfiguration. The central part of the prototype

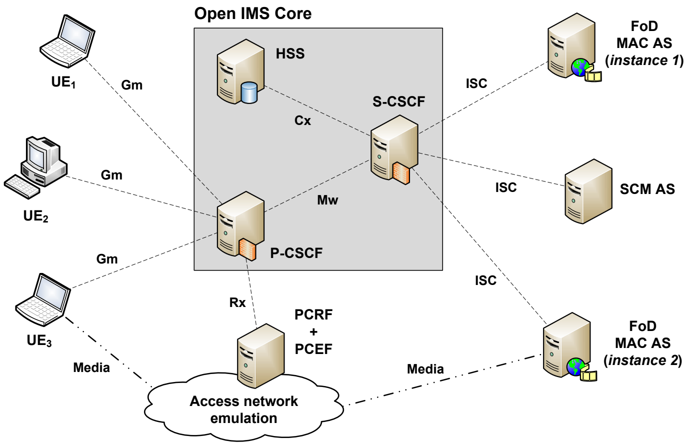

Fig. 2. Architecture of the IMS-based laboratory prototype

is the Open IMS Core (OIC) [23], which is a reference, open source implementation of IMS CSCFs and HSS. We utilize OIC to implement P-CSCF, S-CSCF, and HSS, where the latter is only used for user authentication and authorization.

A. Prototype components

The prototype uses two multimedia applications we developed: Audio-Video Call (AVC) and Football on Demand (FoD). Each of these multimedia services is described with a *service profile*, which specifies its media components and supported encodings, demands on network performance, available transport protocols, etc. AVC offers a conversational service that enables two users to establish an audio-video call. This application incorporates the VLC media player [24] to support live media streaming by the Real-time Transport Protocol (RTP). FoD is a simple Video-on-Demand service for users to watch prerecorded football matches. It is hosted by an SIP AS and also uses VLC for RTP streaming. An FoD feature includes the existence of multiple SIP AS instances, which are organized in a way to serve UEs at different locations. For the demonstration purposes, the location is determined by IP address a UE is assigned to while on the move. These instances are deployed for the load balancing purposes, while moving media flows between them may aid in maintaining QoS.

Each user holds a *user profile*, which specifies her/his preferences (e.g., favored access technology) and capabilities of the associated UEs. For the demonstration purposes, user profile also contains types and predefined characteristics (e.g., available bandwidth, delay, and jitter) of different access networks, which are then signaled in the given use case scenario. A UE enables the user to establish multimedia sessions and to access the offered services. SIP signaling for UE is provided by a signaling application programming interface (SAPI), which relieves application developers from the need to know signaling protocol details. UE also implements modules for VLC streaming control and media reproduction, which are executed for AVC and FoD. FoD is hosted and executed by the SIP AS implementation we refer to as FoD MAC AS (this SIP AS is also built upon the SAPI functionality). As such, it may be the responsibility of either an IMS operator or a 3rd party service provider. FoD MAC AS holds source files of the football matches, adapts their content according to the negotiated session parameters, and streams them by using VLC. In addition, it stores the accompanying service profile.

By implementing the QMOF logic and the SRF logic independently of a specific multimedia service, the SCM AS can be used for different users and multimedia applications. Its inclusion in the SIP signaling path is decided based on service control rules specified at the chosen S-CSCF. The QMOF matches parameters from user and service profiles to recommend feasible combinations of session media flows and their operating parameters, and produces an *optimized service profile*, which contains QoS specification(s) for the resources reservation. One of the most significant aspects of the QMOF is calculation of a Media Degradation Path (MDP). The MDP is a list of an optimal and a number of suboptimal combinations of resources allocations across media flows, which are referred to as MDP configurations and conveyed in SIP messages. They may be used for adapting the resources allocation. The SRF, on the other hand, analyzes delivered mobility context information and controls the reconfiguration.

The Policy and Charging Control (PCC) architecture [25] is responsible for policy control in NGNs and can, thus, be applied in IMS. PCC involves the Policy Control Resource Function (PCRF) and the Policy Control Enforcement Function (PCEF), which are implemented in our prototype [26]. The PCRF decides of the resources authorization and of a suitable MDP configuration to apply in accordance to resource availability, while the PCEF imposes chosen QoS rules. The PCRF interacts with a P-CSCF and the PCEF via Diameter [27]. The P-CSCF selects MDP from a related SIP message and delivers it to the PCRF. The PCRF then invokes the PCEF to identify resource availability and to carry out the authorization/allocation. The PCRF sends a final decision about the allocation back to the P-CSCF. Network resources are emulated by using the Wide Area Network Emulator (WANem) [28]. This tool can be used for emulating multiple network characteristics, including bandwidth, delay, and jitter, which then affect network performance for the media flows.

User and service profiles are organized in an eXtensible Markup Language (XML) format and conveyed among the prototype components in SIP messages. XML was chosen to support modularity and extensibility, but our goal is to introduce the profiles which employ the Session Description Protocol [29], a standardized format for describing multimedia sessions. By then, we are working on the introduction of the XML Document Management (XDM) [30]. XDM enables to manage data stored in XML format on a central file repository. Such an approach would allow to retrieve the profiles past SIP signaling and, thus, to reduce signaling overhead.

B. SIP control procedures

In order to facilitate session reconfiguration and QoS management, we have designed five SIP signaling procedures in IMS for the use case scenarios:

1) *Session establishment* and *Session termination*,
2) *Session reconfiguration upon a change of terminal*,

3) *Session reconfiguration upon a change of terminal's location*, and
4) *Session reconfiguration upon a change of terminal's access network*.

Session establishment negotiates QoS and other session parameters between two end-points, while *Session termination* stops media flows and releases the allocated resources. The procedure applied in response to *Change of terminal* negotiates QoS parameters that conform to the targeted terminal capabilities and adjusts the resources allocation. When a terminal changes location, the associated procedure may result in transferring media flows to maintain QoS. The procedure invoked due to *Change of access network* tunes QoS parameters to the new access characteristics and reserves resources in the network.

1) Session establishment: Figure 3 shows SIP message sequence for establishing media flows between a UE and an MAC AS (e.g., for FoD), which focuses on agreeing upon session parameters and reserving necessary resources.

The sequence assumes that a user and her/his UE are registered to the IMS network. When the user requests a service via its SIP address (i.e. SIP Uniform Resource Identifier, SIP URI), the UE sends an SIP *INVITE* request (step 1, Figure 3) that conveys the user profile to the SCM AS and the MAC AS. The corresponding service profile is delivered to the SCM AS in an SIP *183 (Session Progress)* response (steps 7-8), which triggers the QMOF to generate a *feasible service profile* that comprises an offer of media flows and their parameters (step 9). This profile is then sent to the UE, from which the user chooses among the offered session parameters (step 15).

The resulting service profile is delivered in an SIP *PRACK* request to the MAC AS. When an SIP *OK (to PRACK)* response traverses the SCM AS, the QMOF invokes the *optimization process* to generate an *optimized service profile* (step 23). This profile includes a determined MDP, which is employed for the resources allocation (steps 26-27). If the allocation is successful, the optimized service profile and the applied MDP configuration are forwarded to the UE and the MAC AS (steps 29-38) to start media transmission and establish the agreed flows.

2) Session reconfiguration upon a change of terminal: Figure 4 shows SIP message sequence that negotiates QoS parameters while transferring media flows from, e.g., UE1 to UE2. This signaling sequence assumes that UE1 and UE2 are controlled by the same P-CSCF and the same S-CSCF, but this does not affect its generality. Definition of the procedure is based on the IMS service continuity specification and SIP specification for managing session transfer [31].

In the first part of the procedure, the targeted UE (UE2) is required to establish the current media flows with the MAC AS (thus applying the *start media flow* primitive). This part is identical to the *Session establishment* procedure. To complete the transfer, the flows then need to be terminated between the originating UE (UE1) and the MAC AS. An SIP *REFER* request is employed for delivering the established service information to UE2 (Figure 4, steps 1-6), including address of the used MAC AS, which leads UE2 to invite the MAC AS to establish the flows. The SRF at the SCM AS is invoked (step 20) to examine whether a feasible service profile has already been produced for UE2 and the associated user profile. The latter, together with the delivered session parameters, represent mobility context information for this scenario. If there is no profile produced, the QMOF determines a new service profile offer that takes capabilities of UE2 into account.

After the flows are established between UE2 and the MAC AS (steps 13-59), UE2 sends an SIP *NOTIFY* request to inform UE1 of the transfer. This SIP request triggers UE1 to terminate its participation in the communication (thus applying *stop media flow*), which is initiated by sending an SIP *BYE* request to the MAC AS (steps 72-76). When the MAC AS receives the termination request, it sends an SIP *OK (to BYE)* response to UE1, which is also used for invoking the release of allocated resources (steps 81-82).

3) Session reconfiguration upon a change of terminal's location: If, e.g., by replacing UE the location is also changed, the SRF at the SCM AS may decide upon moving media flows between different MAC AS instances. That way, media flows can be established with the instance that is "closer" to the used UE, which could help in maintaining QoS. In that case,

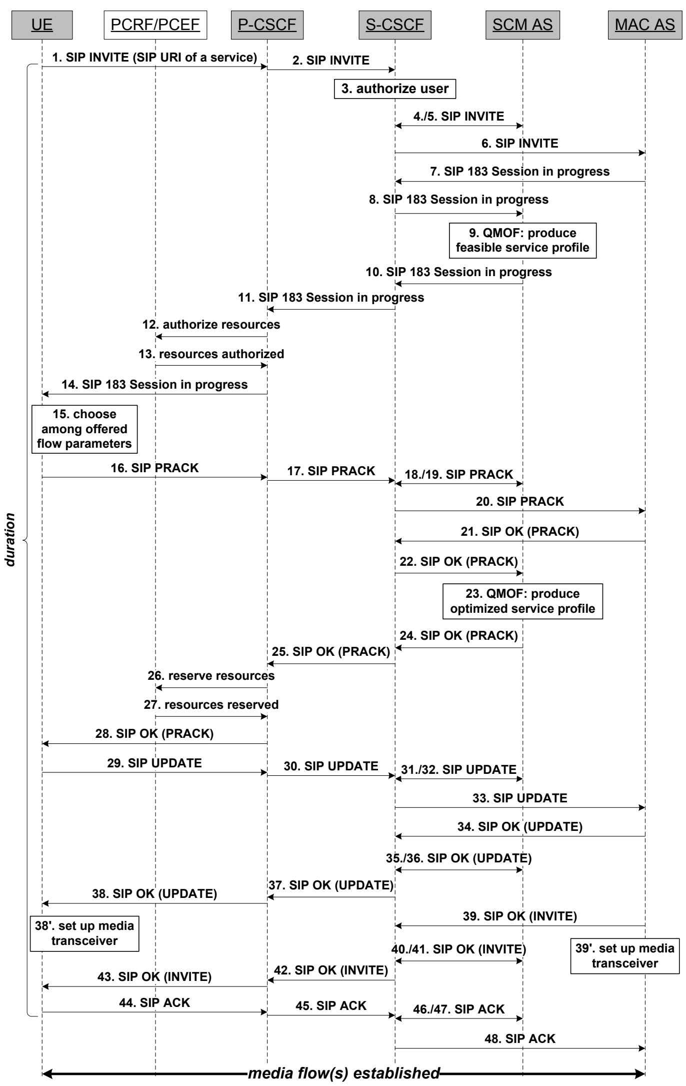

Fig. 3. SIP signaling for *Session establishment*

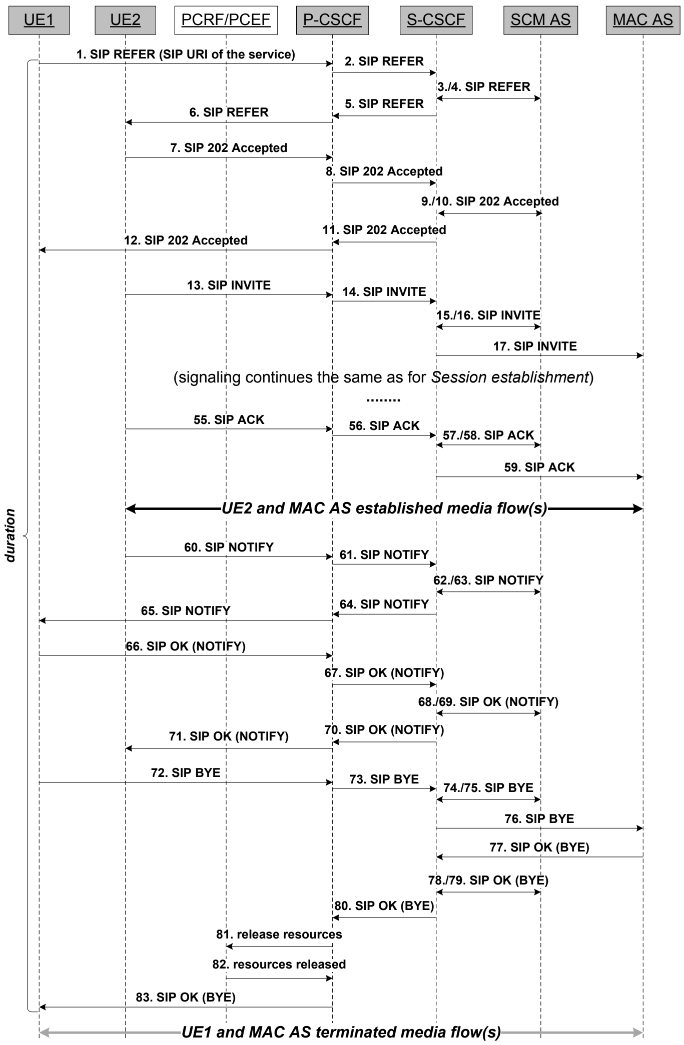

Fig. 4. SIP signaling for *Session reconfiguration upon a change of terminal*

the SCM AS will modify the SIP *REFER* request (step 3 in Figure 4) to target another MAC AS instance (e.g., MAC ASi2 instead of MAC ASi1) by providing its address. This would instruct UE2 to establish the flows with MAC ASi2.

V. Case Study and Performance Evaluation

A. Experimental testbed

Case study demonstration and performance evaluation measurements are conducted in an experimental network shown in Figure 5. Configuration of the nodes which host the components of the IMS laboratory prototype is depicted in Table II. It must be emphasized that the network does not involve any traffic besides the one pertaining to the applied SIP procedures and to media delivery within the prototype services.

B. Case study scenario

The purpose of this case study is to demonstrate application of our approach when QoS for media flows is negotiated during the establishment phase and adapted when a user decides to change the terminal for communication, which also includes a change of location and of access network.

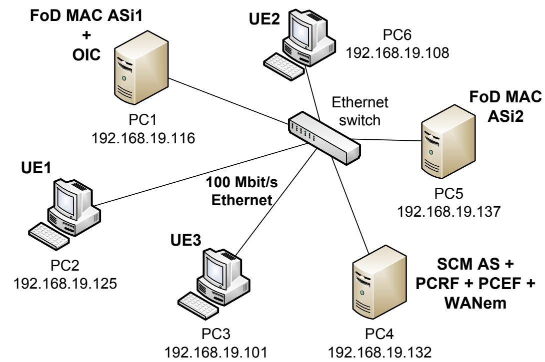

Fig. 5. Topology of the experimental testbed

TABLE II
Configuration of the testbed nodes

Node	Oper. system	Configuration
PC1	Linux Ubuntu	Pentium IV, CPU 3.0 GHz, RAM 1 GB
PC2	Linux Ubuntu	Pentium IV, CPU 2.4 GHz, RAM 512 MB
PC3	Linux Ubuntu	Pentium IV, CPU 1.7 GHz, RAM 1 GB
PC4	Linux Ubuntu	Pentium IV, CPU 1.7 GHz, RAM 1 GB
PC5	Linux Ubuntu	Pentium IV, CPU 1.6 GHz, RAM 512 MB
PC6	Linux Ubuntu	Pentium IV, CPU 1.7 GHz, RAM 1 GB

1) Session establishment: In the first part of the scenario, two friends, Alice and Bob, decide to watch a football match together over the IMS network. Their IMS operator offers a FoD service via a 3rd party service provider, which deploys multiple MAC AS instances for the service (e.g., FoD MAC ASi1 and FoD MAC ASi2). Alice is at home. She uses her laptop computer (represented by UE1) over an Asymmetric Digital Subscriber Line (ADSL) connection, which supports a downlink of 10 Mbps and an uplink of 512 kbps, to establish a session with an FoD MAC AS. This session will be referred to as *session1*. Based on the Alice's location, UE1 establishes the session with FoD MAC ASi1, which comprises one audio and one video flow (Figure 6).

At the same time, Bob is traveling home by train. He uses his smarthpone (UE2) over a High-Speed Packet Access (HSPA) connection, which supports a downlink of 3.6 Mbps and an uplink of 384 kbps, to watch the game. Based on his location, UE2 establishes the session with FoD MAC ASi2.

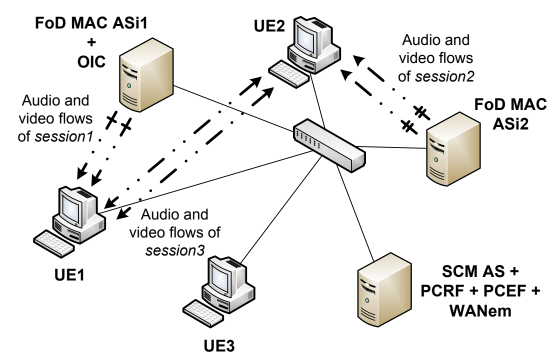

Fig. 6. Flow map after establishing sessions

This session will be referred to as *session2*. Just after the match started showing, Alice invites Bob to an audio-video call, so they can comment on the match together. The latter session, *session3*, comprises two audio and two video flows (Figure 6). An example of the QoS parameters that result from the *Session establishment* procedures is depicted in Table III, with eight media flows established in total. These parameters are used at WANem to reserve necessary resources and provide QoS.

TABLE III
RESULTING QOS PARAMETERS AFTER ESTABLISHING SESSIONS

Media flow	**Bandwidth** (kbps)	**Delay** (ms)	**Jitter** (ms)	**Drop** (%)
session1: video	1024	200	100	0.4
session1: audio	64	200	100	0.4
session2: video	512	200	100	0.7
session2: audio	48	200	100	0.7
session3: video1	128	100	50	1.0
session3: audio1	32	100	50	1.0
session3: video2	128	100	50	1.0
session3: audio2	32	100	50	1.0

2) Session reconfiguration: After coming home, Bob decides to transfer the communication to his laptop computer (UE3), which is connected to the network via an ADSL connection. For the demonstration purposes, this request also includes a change in location, which is represented by different IP addresses. While processing the SIP *REFER* request (Figure 4, step 3), the SRF at the SCM AS processes the change in location and instructs UE3 to establish media flows of *session2* with FoD MAC ASi1, instead of FoD MAC ASi2. After *session2* and *session3* are transferred (Figure 7), Bob continues watching the match and chatting with Alice on UE3.

To achieve service continuity, SIP messages of the transfer procedure carry information about the elapsed time for the match, which enables Bob to resume watching the game from the right moment. An example of the resulting QoS parameters, which are enforced after the flow transfer, is shown in Table IV. The parameters from the first part of the scenario are improved regarding enhancements in hardware configuration of UE3 and its access network. Transport parameters of *session2* and *session3* are updated to reflect the new UE.

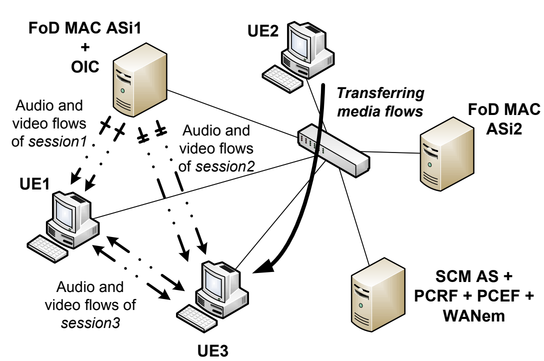

Fig. 7. Flow map after transferring media flows

TABLE IV
RESULTING QOS PARAMETERS AFTER TRANSFERRING MEDIA FLOWS

Media flow	**Bandwidth** (kbps)	**Delay** (ms)	**Jitter** (ms)	**Drop** (%)
session1: video	1024	200	100	0.4
session1: audio	64	200	100	0.4
session2: video	1024	200	100	0.4
session2: audio	64	200	100	0.4
session3: video1	256	100	50	0.7
session3: audio1	64	100	50	0.7
session3: video2	256	100	50	0.7
session3: audio2	64	100	50	0.7

C. Performance evaluation

A preliminary performance evaluation of our approach in an IMS setting is conducted to analyze delay induced by the reconfiguration procedures, with the focus on examining a scalability of the solution. For the purposes of this work, we define the *duration* performance metric that refers to the time interval required to complete a specific SIP procedure from the user perspective (Figures 3 and 4). This metric is similar to the SIP performance metrics specified in [32]. For *Session reconfiguration upon a change of terminal*, *duration* is the interval between sending *1. SIP REFER* and receiving *83. SIP OK (BYE)* at UE1, while its "reference value" implies procedure completion for a single UE. *Duration* is measured in relation to the number of UEs simultaneously executing a particular procedure with, e.g., an MAC AS. The measurement results for the analyzed SIP procedures are given in Table V. Average *duration* was obtained over 30 test runs.

TABLE V
AVERAGE *duration* FOR THE ANALYZED SIP PROCEDURES

The procedure / **Number of UEs**	**1**	**4**	**7**	**10**
Session establishment [s]	6.13	6.61	7.37	8.32
S. recon. u. a chan. of terminal [s]	7.62	8.13	8.85	9.89
S. recon. u. a chan. of term. locat. [s]	7.66	8.18	8.93	10.00

Signaling load of multiple UEs exchanging SIP messages with an MAC AS instance is achieved by employing the SIP traffic generator called SIPp [33]. SIPp is able to create SIP messages as per user-defined scenarios, and we customize it to send the messages of the applied reconfiguration procedures. The performance, besides the SIP message exchange, is influenced by time duration of the matching and optimization processes at the SCM AS (comparing to the SRF, for which it is negligible). Table VI shows average QMOF processing time.

TABLE VI
AVERAGE QMOF PROCESSING TIME

The QMOF process / **Number of UEs**	**1**	**4**	**7**	**10**
Matching process [s]	0.74	1.10	1.46	1.88
Optimization process [s]	0.19	0.31	0.45	0.61

As user and service profiles constitute a signaling overhead, which affects the overall performance, we alter the SIP procedures in a way that SIP messages only reference the profiles, instead of carrying them along the signaling path. This required all the profiles to be produced in advance and stored at each prototype component that uses them. The measurement results are given in Table VII.

TABLE VII
AVERAGE *duration* FOR THE PROCEDURES WITH PROFILE REFERENCING

The procedure / **Number of UEs**	**1**	**4**	**7**	**10**
Session establishment [s]	5.08	5.55	6.28	7.33
S. recon. u. a chan. of terminal [s]	6.59	7.11	7.84	8.79
S. recon. u. a chan. of term. locat. [s]	6.64	7.15	7.92	8.89

The results show that *duration* increases "slightly faster" than the increase in the number of UEs, and in a non-linear fashion, which does not promise a good scalability. In addition, overall *duration* of the procedures poses a QoS violation itself, by leading to the signaling delays that are, e.g., around a hundred times longer than requested QoS delays. But, the results are encouraging when we compare them to results from [34], where IMS session establishment delay is reported as 3.37 seconds, or to results from [35], where session reestablishment delay due to vertical handover is reported as around 2.5 seconds. Moreover, it can be noticed that *duration* improves by over a second when profile conveyance is removed from SIP messages, which could justify the decision to introduce the XDM management. Different mechanisms will be investigated to mitigate the mentioned effects, with a focus on a notable delay introduced by the SCM AS processes. As the SCM AS represents a bottleneck in the current prototype deployment, several of its instances could be employed to serve different UEs and share the load. In addition, the SCM AS should be realized as the session anchor point, regarding the SCC AS, which would suppress the need for end-to-end signaling and may lead to faster reconfiguration procedures.

VI. CONCLUSION

This paper presents an approach to enhancing QoS support in the 3GPP IMS by dynamically producing multimedia session configurations with respect to session and terminal mobility. We propose to introduce an SIP AS that, based on received mobility context information, steers the reconfiguration in the IMS network and produces QoS specifications conforming to the mobility-induced constraints. We design several SIP signaling procedures, which are built upon the 3GPP specifications, to enable the session reconfiguration. We also implement an IMS laboratory prototype and describe a case study, in which QoS is negotiated during session establishment, and successfully adapted when a user replaces her/his UE and changes its access network. An initial performance evaluation of our solution indicates the signaling delay of a few seconds, which is generally unacceptable, but comparison to the similar research results encourages us to investigate different mechanisms in order to reduce this delay. Future work will include additional performance evaluation to address scalability of the solution in the context of various background traffic conditions.

ACKNOWLEDGMENT

This work was carried out within the research project 036-0362027-1639 "Content delivery and mobility of users and services in new generation networks", which is supported by the Ministry of Science, Education and Sports of the Republic of Croatia. The authors would also like to thank anonymous reviewers for their constructive comments and suggestions.

REFERENCES

[1] Camarillo, G. and Garcia-Martin, M.-A., The 3G IP Multimedia Subsystem (IMS): Merging the Internet and the Cellular Worlds, 2nd ed. Chichester, West Sussex, England: Wiley, 2006.

[2] Yun, C. and Perros, H., "QoS Control for NGN: A Survey of Techniques," *Journal of Network and Systems Management*, vol. 18, no. 4, pp. 447–461, 2010.

[3] ITU Telecommunication Standardization Sector (ITU-T), Recommendation Y.2001 "General Overview of NGN," 2004.

[4] 3rd Generation Partnership Project (3GPP), Technical Specification 23.228 "IP Multimedia Subsystem (IMS); Stage 2 (Release 11)," 2011.

[5] Rosenberg, J., Schulzrinne, H., Camarillo, G., Johnston, A., Peterson, J., Sparks, R., Handley, M., and Schooler, E., "SIP: Session Initiation Protocol," RFC 3261 (Proposed Standard), 2002.

[6] Rosenberg, J., Schulzrinne, H., "An Offer/Answer Model with the Session Description Protocol (SDP)," RFC 3264 (Proposed Standard), 2002.

[7] 3GPP, Technical Specification 23.207 "End-to-end Quality of Service (QoS) concept and architecture (Release 10)," 2011.

[8] 3GPP, Technical Specification 23.237 "IP Multimedia Subsystem (IMS) Service Continuity; Stage 2 (Release 11)," 2012.

[9] Dobrijevic, O. and Matijasevic, M., "Multimedia Session Reconfiguration for Mobility-aware QoS Management: Use Cases and the Functional Model," *Smart Innovation, Systems and Technologies*, vol. 11, pp. 209–218, 2011.

[10] Yee, Y. C., Choong, K. N., Low, A. L. Y., and Tan, S. W., "SIP-based Proactive and Adaptive Mobility Management Framework for Heterogeneous Networks," *Journal of Network and Computer Applications*, vol. 31, no. 4, pp. 771–792, 2008.

[11] Cerqueira, E., Veloso, L., Neto, A., Curado, M., Monteiro, E., and Mendes, P., "Mobility Management for Multi-user Sessions in Next Generation Wireless Systems," *Computer Communications*, vol. 31, no. 5, pp. 915–934, 2008.

[12] Munasinghe, K. S. and Jamalipour, A., "Interworking of WLAN-UMTS Networks: An IMS-based Platform for Session Mobility," *IEEE Communications Magazine*, vol. 46, no. 9, pp. 184–191, 2008.

[13] Rawashdeh, M. and Karmouch, A., "Seamless Video Handoff in Session Mobility over the IMS network," in: 10th IEEE International Symposium on a World of Wireless Mobile and Multimedia Networks, pp. 1–6. IEEE, New York, New York, 2009.

[14] Bellavista, P., Corradi, A., and Foschini, L., "IMS-Compliant Management of Vertical Handoffs for Mobile Multimedia Session Continuity," *IEEE Communications Magazine*, vol. 48, no. 4, pp. 114–121, 2010.

[15] Chiang, W.-K. and Kuo, P.-C., "IMS-Based Centralized Service Continuity" (2012) [Online], *Wireless Personal Communications*, available: DOI: 10.1007/s11277-012-0503-z

[16] Navarro, M. and Donoso, Y., "An IMS Architecture and Algorithm Proposal with QoS Parameters for Flexible Convergent Services with Dynamic Requirements," *International Journal of Computers, Communications & Control*, vol. 7, no. 1, pp. 123–134, 2012.

[17] Yang, S.-R. and Chen, W.-T., "SIP Multicast-based Mobile Quality-of-Service Support over Heterogeneous IP Multimedia Subsystems," *IEEE Transactions on Mobile Computing*, vol. 7, no. 11, pp. 1297–1310, 2008.

[18] Renier, T., Larsen, K. L., Castro, G., and Schwefel, H.-P., "Mid-session Macro-mobility in IMS-based Networks," *IEEE Vehicular Technology Magazine*, vol. 2, no. 1, pp. 20–27, 2007.

[19] Liao, J., Qi, Q., Zhu, X., Cao, Y., and Li, T., "Enhanced IMS Handoff Mechanism for QoS Support over Heterogeneous Network," *The Computer Journal*, vol. 53, no. 10, pp. 1719–1737, 2010.

[20] Skorin-Kapov, L., Mosmondor, M., Dobrijevic, O., and Matijasevic, M., "Application-level QoS Negotiation and Signaling for Advanced Multimedia Services in the IMS," *IEEE Communications Magazine*, vol. 45, no. 7, pp. 108–116, 2007.

[21] 3GPP, Technical Specification 23.218 "IP Multimedia (IM) session handling; IM call model; Stage 2 (Release 11)," 2011.

[22] 3GPP, Technical Specification 24.228 “Signalling flows for the IP multimedia call control based on Session Initiation Protocol (SIP) and Session Description Protocol (SDP); Stage 3 (Release 5),” 2006.
[23] OpenIMSCore.org | The Open Source IMS Core Project (2012) [Online]. Available: http://www.openimscore.org/
[24] VideoLAN – Official page for VLC media player, the Open Source video framework! (2012) [Online]. Available: http://www.videolan.org/vlc/
[25] 3GPP, Technical Specification 23.203 “Policy and charging control architecture (Release 11),” 2011.
[26] Grgic, T. and Matijasevic, M., “Online Charging in IMS for Multimedia Services with Negotiable QoS Requirements Based on Service Agreements,” *International Journal of Intelligent Information and Database Systems*, vol. 4, no. 6, pp. 656–672, 2010.
[27] Calhoun, P., Loughney, J., Guttman, E., Zorn, G., and Arkko, J., “Diameter Base Protocol,” RFC 3588 (Proposed Standard), 2003.
[28] WANEM : The Wide Area Network Emulator (2012) [Online]. Available: http://wanem.sourceforge.net/
[29] Handley, M., Jacobson, V., and Perkins, C., “SDP: Session Description Protocol,” RFC 4566 (Proposed Standard), 2006.
[30] Releases – XML Document Management v2.0 (2012) [Online]. Available: http://www.openmobilealliance.org/technical/release_program/xdm_v2_0.aspx
[31] Shacham, R., Schulzrinne, H., Thakolsri, S., and Kellerer, W., “Session Initiation Protocol (SIP) Session Mobility,” RFC 5631 (Informat.), 2009.
[32] Malas, D. and Morton, A., “Basic Telephony SIP End-to-End Performance Metrics,” RFC 6076 (Proposed Standard), 2011.
[33] Welcome to SIPp (2012) [Online]. Available: http://sipp.sourceforge.net/
[34] Munir, A. and Gordon-Ross, A., “SIP-based IMS Signaling Analysis for WiMax-3G Interworking Architectures,” *IEEE Transactions on Mobile Computing*, vol. 9, no. 5, pp. 733–750, 2010.
[35] Chiba, T., Yokota, H., Dutta, A., Chee, D., and Schulzrinne, H., “Performance Analysis of Next Generation Mobility Protocols for IMS/MMD Networks,” in: 4th International Wireless Communications and Mobile Computing Conference, pp. 68–73. IEEE, Piscataway, New Jersey, 2008.

Ognjen Dobrijevic received his Dipl.-Ing., M.Sc., and Ph.D. degrees in Electrical Engineering from the University of Zagreb, Croatia, in 2004, 2007 and 2011, respectively. He is currently employed as a senior research and teaching assistant at the University of Zagreb, Faculty of Electrical Engineering and Computing. His research interests include next generation networks and Future Internet, quality of service (QoS) and mobility management for multimedia services, and networked virtual environments. Ognjen is a member of IEEE.

Maja Matijasevic is a Full Professor at the University of Zagreb, Faculty of Electrical Engineering and Computing, Croatia. She received her Dipl.-Ing. (1990), M.Sc. (1994), and Ph.D. (1998) degrees in Electrical Engineering from the University of Zagreb and her M.Sc. in Computer Engineering (1997) from the University of Louisiana at Lafayette, LA, USA. Her research interests include quality of service (QoS) and quality of experience (QoE) for multimedia services in IP-based next generation networks, with the particular focus on session negotiation, adaptation, and mobility. Dr. Matijasevic has over 70 journal and conference publications, and she has co-authored several books and book chapters. She is a member of IEEE and ACM.