

QoS Signalling in 4G Scenarios

Susana Sargento, Diogo Gomes, Alfredo Matos, Rui L. Aguiar
Universidade de Aveiro, Instituto de Telecomunicações, Portugal

Abstract— This paper presents an evaluation of possible a QoS architecture and the corresponding QoS signalling strategies for next generation networks, being developed inside the IST project Daidalos. We address the main results achieved in terms of their performance in an experimental testbed developed inside the project. The result provide indications on the expected performance of different signalling strategies for next generation heterogeneous networks.

Index Terms—QoS, signalling, multimedia, experimental evaluation

I. INTRODUCTION

Next generation wireless communication systems will handle diverse types of services, across different types of access technologies. This trend, already present in 3G networks and in the current explosion of hotspots, is expected to become a universal characteristic in communications by the end of this decade. Providing mobility across domains using different access technologies in a seamless way, with no perceived service degradation for the user, is a major requisite for the next generation networks. Scalability concerns make this requirement still harder. Current wireless operators have dozens of millions of customers, and, as cell sizes decrease, handovers will become more and more frequent, potentially reaching the hundred thousands per second in a large telecom operator.

The DAIDALOS project IV aims at seamlessly integrating these heterogeneous network technologies, thus allowing network operators and service providers to offer new and profitable services (voice, data, multimedia). The architecture integrates both wired and wireless technologies, with quality of service capabilities under a common authentication, authorization, accounting, auditing and charging (A4C) framework and in a secure communication environment.

Providing mobility across domains using different access technologies in a seamless way is a major requirement for the next generation networks. The provision of seamless end-to-end QoS in such a demanding and heterogeneous scenario, requiring no perceived service degradation for the user when moving across different access technologies, is one of the main challenges in Daidalos. This technical problem is further compounded by the complex telecom business, with multiple types of operators foreseen in the market, covering a wide range, from basic transport to intelligent service and multimedia provision.

Daidalos has already proposed a next-generation 4G architecture IV and signalling strategies taking in consideration both QoS and mobility aspects. The proposed architecture acknowledges the existence of legacy applications in the network, and provides seamless support for these applications. This paper presents results of the experimental evaluation of these signalling strategies, and comments on their adequacy to next generation networks. The paper will focus on evaluating the session setup and negotiation provided by each signalling solution on an experimental testbed. The different concepts of signalling

strategies are analyzed according to three major scenarios: the mobile terminal itself performs the QoS requests to a QoS Broker (responsible for resource management at the access network) either through implicit mechanisms (i) or explicit signalling (ii); a multimedia service proxy is responsible for requesting network resources (iii). Even though the multimedia service proxy represents a similar approach to 3G, for SIP based services, two more signalling scenario provide full support for legacy and unaware applications.

This paper is organized as follows. Section II presents the network architecture, its elements and generic implementation considerations. Section III presents a experimental evaluation of the different signalling scenarios for QoS support and illustrates their main characteristics using message signalling charts. Section IV performs a comparison between these different strategies, discussing our key conclusions.

II. NETWORK ARCHITECTURE

Next generation communication systems will aim at providing seamless mobility of users through networks with different access technologies and services. In this sense, the network needs to be capable of supporting heterogeneous access technologies. These communication systems, usually referred to as 4G networks IV, may support network technologies such as WiFi, UMTS, and new emerging technologies, such as WiMax and DVB-T. These technologies are quite different from each other, ranging from Local Area Networks (LAN) to Broadcast Diffusion Networks with quite distinct network architectures.

The main focus of 4G systems is the support of all these technologies under an unified network architecture capable of supporting the different access technologies. The IPv6 protocol is thus an adequate convergence layer to provide such a unified platform, creating an abstraction layer at transport level that hides technology specific parameters from advanced services. Moreover, its intrinsic support of mobility is quite important for 4G communication systems; since it provides “almost” seamless mobility between different technologies (support for fast mobility IV must be usually used IV).

Figure 1 depicts the proposed QoS architecture that supports several access (wireless) networks, each of them capable of handling several access technologies. Differentiated Services (DiffServ) IV is used to support QoS in the core network, achieving scalability and performance. In broad terms, this architecture is very flexible, and presents a more comprehensive set of characteristics, such as: a fully integrated approach to IP-based communication with different types of applications and protocols (e.g. both legacy and SIP-based applications are supported), including adaptive applications, multicast and broadcast; the customization/optimization of the architecture according with the expected service mix to support; and the integrated support of multiple QoS service models, according to the overall network configuration (defined by operator policies).

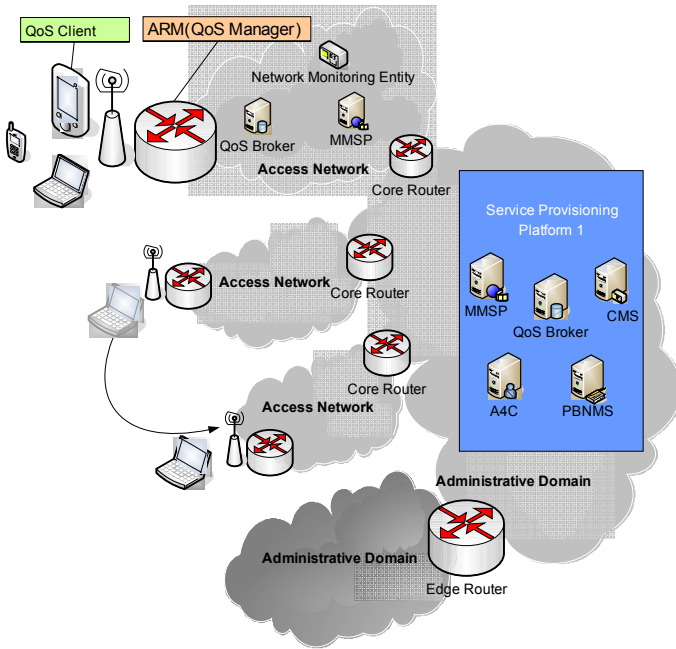


Figure 1: Daidalos QoS network architecture

The architecture distinguishes three networks: the Core Network, where the Service Provisioning Platform (SPP) is located; the Access Network (AN), which further extends the SPP by distributing some of its functionalities; and the Access Link Network where the Terminals lie.

A. Network elements

The architecture is constituted by three main functions: QoS, A4C and Multimedia Control. These functions are distributed by several elements in order to improve scalability.

The most important QoS element of the architecture is the QoS Broker, which performs admission control and manages network resources; it controls the network routers according to the active sessions and their requirements. It also performs load balancing of users and sessions among the available networks (possibly with different access technologies) by setting off network-initiated handovers. This is a quite important feature, since it provides the means to optimize the usage of operator resources.

While basic QoS services are provided intrinsically by the AN, more advanced services are supported by the SPP, in the core network. The MultiMedia Service Proxy (MMSP), aware of the requirements of user services, and the QoS Broker in the Access Network (AN QoS Broker) can have a very close relation. Merging the high level knowledge of running services with the available network resources might, for instance, enable the network to move a video stream from a Wi-Fi network to a DVB-T link, and provide the adequate network-level QoS to a multimedia stream.

For the provisioning of multimedia streaming services, MultiMedia Servers (MMServer) may also be present, located in the application servers' garden (a special Access Network for Application Servers). The QoS definitions at the domain level are provided by a Policy Based Network Management System (PBNMS), and then proxied by the AN QoS Brokers to the Access Routers (AR) in the different access networks. For authentication and accounting purposes, an A4C (Authentication, Authorization, Accounting, Auditing and Charging) server is also present in each domain. The Core Network (CN) also has a QoS Broker, to deal with aggregates of flows traversing the core network and communication with other administrative

domains. To aid in the admission control procedure performed by the QoS Brokers, a real time network monitoring system is included, which comprises Network Monitoring Entities (NME) located in several points of the network, and a Central Monitoring System (CMS).

On the terminals, and in order to perform explicit signalling of QoS requirements, the QoS Client entity provides a lightweight RSVP daemon functionality (in the future this might be converted into NSIS, e.g.), and other QoS signalling capabilities.

4G networks must support all types of services. Although the support of multimedia services can be provided by means of interaction between the MMSP and the QoS Broker, this solution is not suitable for IP legacy applications, which may equally have service quality requirements IV. In order to provide QoS to those legacy, QoS unaware applications, some advanced functions must be added to the access routers. The required functionality comprises connection tracking, similar to what may be found in a Network Address Translation (NAT) router with port translation, per-application flow DiffServ Code Point (DSCP) marking, and the means to translate other QoS reservation mechanisms, such as Integrated Services (IntServ) IV resource ReSerVation Protocol (RSVP) reservations, into DiffServ DSCP marking and QoS Broker requests. We refer to the entity supporting all these functions as the Advanced Router Mechanisms (ARM) IV.

The ARM provides functionality equivalent to a basic proxy without the need to change any of the legacy applications, and can be considered as a dedicated intelligent transparent proxy. Note that the ARM can also perform application to network level QoS mapping for multimedia services, e.g. for Session Initiation Protocol (SIP) IV services, issuing the resource reservation requests to the QoS Broker and filtering the QoS configurations in the application signalling messages IV. The ARM may, therefore, perform the QoS related functions that are typically required also from a MMSP, if the operator so desires, since the application/service logic allows these operations to be delegated to the ARM.

B. QoS support

In order to provide end-to-end QoS to the application flows, enough resources must be available along the entire flow path. In the most demanding scenario, where the mobile terminals communicating are attached to different access domains, this path comprises (1) the access networks of both terminals, (2) the core networks of the access domains where the access networks belong and (3) the inter-domain path, consisting of all the transit domains traversed by the flows. Daidalos handles QoS control in all these segments with a scalable approach based on DiffServ with resource control: resource management at the core is performed on a per-aggregate basis and based on information from a monitoring platform, whereas in the wireless link, where (radio) resources are scarce, per-flow QoS control is applied.

This 4G network provides overall control and management mechanisms, relying on the overlay setup of a distributed set of QoS Brokers, and achieving QoS at the access link by means of explicit reservation in an IntServ like manner. In order to coordinate all these mechanisms, several signalling strategies arise. Implicit signalling in the DiffServ environment is very simple, but requires applications to produce marked packets with the right DSCP code, and reduces the control flexibility achieved by IntServ-like reservations in the access network, where (radio)

resources are scarce. Explicit signalling can be done by application signalling protocols (for example, RSVP messages or using SIP and its companion Session Description Protocol – SDP to describe the required session and QoS parameters) and interaction between the MMSP and the QoS Broker.

These solutions are more complex and involve a larger set of signalling messages, but increase the flexibility of the characteristics of the services to be offered to the user. These different solutions are very closely tied to the QoS service model that will be used, which is directly associated to the business model chosen by the operators: (1) application oriented, IntServ like, (2) user oriented, where the user asks for the service characteristics he wants to, and (3) service oriented, where the user has some well known contracted services with the network and/or service operator. Note that these QoS models are not disjoint and independent from each other; they are closely related to the QoS signalling strategies in place, which will be discussed in the next sections.

Since the ANQoS only has information on the network services to be delivered, it is required to map the application QoS parameters to network QoS parameters. This mapping can be made in the QoS client, ARM or in the MMSP, depending on the signalling strategy used. The network service is described by two parameters: the network service QoS class (where the class is specified by a set of QoS parameters), and the bandwidth to be reserved. The definition of the network service classes is conformant with DiffServ network architecture.

To enable fast QoS and session setup for specific services and users, beyond the services defined and negotiated by the user, we also introduce the concept of the “well-known services”. These are the set of network services characterized by pre-defined parameters that are offered by the specified network operator. These services do not accept negotiation procedures, since the granularity is the one already pre-defined by the operator.

III. PROTOCOL STRATEGY EVALUATION FOR QoS CONTROL

All the signalling strategies that will be presented below are based on this architecture, and therefore, are based on the QoS Broker concept for resource reservation. The protocol used for the communication between QoS Brokers (PDP) and its “clients” (ARM and MMSPP) is the Common Open Policy Service (COPS). These strategies were tested in a basic network with two Access Routers, one for the MT and the other for the correspondent node, and a core network, where all the control entities lay. The signalling message charts presented below were obtained with a ping6 test.

Several signalling methodologies are currently being thought for the support of QoS in mobility environments, (e.g. the ones being defined by the NSIS WG). These signalling approaches can be later incorporated in this architecture for the signalling in the access network. Currently, we used RSVP for making explicit QoS reservations. For Multimedia Applications support we used the SIP (Session Initiation Protocol) IV protocol between the terminal and the MMSPP.

A. Implicit QoS Reservation

This signaling flow assumes the presence of a DiffServ Network, and as such, makes use of the DiffServ Code Point (DSCP) of each and every packet to implicitly signal the QoS System of the desired QoS for the given packet. For this

scheme to work, three entities must be involved: QoS Client, ARM and QoS Broker.

The QoS Client, using its DSCP marking module, is able to install per flow filters capable of marking the packets leaving the terminal. The DSCP values used are in this case provided by the A4C and correspond to a subset of the NVUP (Network View of the User Profile) which is nonetheless than a set of QoS Policies. The NVUP is transferred during registration phase of the terminal to both the QoS Broker of the Access Network where the terminal logged in, and to the QoS Client (in the form of a subset). This procedure gives an assurance of information synchronization between the different QoS elements.

Figure 2a depicts a message sequence chart (MSC) of an implicit QoS Reservation. The reservation process is triggered by the first packet of a flow; the flow corresponds to a “ping6”, so we can see the process being triggered by the “echo request” ICMPv6 packet. Instead of immediately forwarding the packet, it outsources to the QoS Broker the decision to accept or drop the flow based on the availability of resources in the network. The Decision also includes QoS Policer Rules, and Queue Configurations thus accounting for the relative long delay between the COPS Unsolicited Decision to Router 2 / Request Handle message and COPS Decision to Router 2 / Exit of the ICMPv6 packet from Router 1. Also of relevance is the time delay of the QoS Broker to issue the first decision, as this action involves the processing of the NVUP, calculation of the network trunks occupied by the session, and resource reservation of the same.

The reservation performed here is solely based on the NVUP information available at the QoS Broker. This signaling mechanism, although highly efficient, is rather limited to previously agreed QoS Profiles.

B. Explicit QoS Reservation

Besides Implicit QoS Reservation the QoS Client is also involved in Explicit QoS Reservation, through a RSVP module. This signaling scenario tries to fill the gap of the previous one in terms of flexibility, at the expense of extra complexity. Our use case for explicit signaling is nonetheless not related to the protocol used (RSVP). RSVP is regarded as an IntServ protocol, but as stated in previous sections, our architecture is DiffServ based. The use of RSVP arises from the need to have a protocol capable of extensively describing QoS parameters, between the MT and the AR. Thus, its use is limited to the Access Link Network, for signaling between Terminals and Routers. In the router a RSVP agent translates the RSVP messages into COPS objects in order to perform the same outsourcing mechanisms as in the previous signaling flow.

That said, we can now understand the absence of RSVP messages PATH and RESV between Routers (opposed to the normal RSVP behavior), a further example that RSVP could be substituted by NSIS or any other protocol capable of describing QoS parameters.

Figure 2b depicts the mentioned signaling flow. The RSVP PATH message is sent toward the Correspondent Node, thus enabling Router 1 to retrieve information on the end to end flow. The R-Spec and T-Spec information from RSVP is translated into Client Specific messages sent to the QoS Broker, the Admission process of the QoS Broker in this case is more flexible, thus taking slightly longer to process all the parameters. In order for the RSVP PATH message to be generated in Router 2 the QoS Broker sends new R-Spec and T-Spec objects to the Router. The longest delay in this signaling flow is associated to the response to

the RSVP PATH which in our case took over 2 seconds, due to propagation and rsvpd processing delays. It is our understanding that the rsvpd delay is caused by the usage of less efficient algorithms with regard to RSVP options processing and internal communication overhead. This delay could be overcome by using a different RSVP daemon (more efficient).

Overall this signaling flow takes longer than the implicit reservation, (even disregarding the 2 seconds required by the RSVP daemons). Nonetheless it provides the terminal with the capability to reserve any given QoS under limitative parameters defined in the NVUP.

Multimedia Applications are a potential target of this signaling flow, as they require precise QoS constraints but can depend of the CODECs used for encoding of audio and video. Through such flexible interface the application can request the network a broad set of QoS Profiles in addition to the ones already defined in the NVUP, and thus maintaining the scalability of the NVUP by restricting it to the most common scenarios.

C. Multimedia Proxy based Reservation

Although the previous signaling scenario seems appropriate for multimedia applications, the requisites of the terminal in terms of processing and storage capabilities might limit the usage of an explicit QoS Signaling daemon, further Operator business cases might want to limit the QoS profiles used in the network. These factors associated with the availability of the Multimedia Service Provisioning Proxy (MMSPP) in the SPP, makes the emergence of a signaling flow for this situation quite logical. The SIP signaling carries CODEC information that can be used to map in to QoS Parameters. Considering the MMSPP as yet another Policy Enforcement Point, it triggers a COPS Request to the QoS Broker, that like in previous scenarios checks the NVUP and issues Unsolicited Decisions to the Routers serving the end points containing the necessary QoS configurations.

The first messages in this signaling scenario are SIP OPTION and not SIP INVITE. This is a requisite for the correct profiling of the flow in terms of IP/protocol/ports.

The OPTIONS and OK messages enable the User Agents to agree on communication ports. Only after this does the SIP INVITE trigger the signaling of QoS reservation, Based on the capabilities of the terminal, the MMSPP sorts the CODECs of the INVITE and request the QoS Broker QoS reservations for two flows (in our case an audio stream, composed of audio RTP and corresponding RTCP).

Analyzing Figure 2c, we acknowledge that the bottleneck of this scenario is the processing of SIP messages. This conclusion is not a novelty, but the reason of this delay is. It was actually expected, as SIP is a text based protocol, while COPS is a binary object protocol much easier parsed by machines. In practice, SIP processing took several dozens milliseconds, but that includes the end-to-end communication delay. The major problem is the synchronization messages required to set the network QoS according to the effective requirements of the specific multimedia codec being used. If this optimum resource management is not required, then several of these messages can be omitted, and the process will be faster.

Nonetheless this scenario proved to be more efficient than the previous RSVP-based scenario. Because of the

reservation setup problems of the RSVP daemons, this evaluation shows that terminal driven explicit signaling is less efficient than a central server collocated in the SPP.

IV. CONCLUSIONS

This paper presented an evaluation of the signalling behaviour of a QoS architecture being implemented inside the Daidalos project. This architecture is able to provide end-to-end QoS (in a heterogeneous mobile environment) for many types of services and applications, legacy and multimedia, unicast and multicast, with optimized network resource usage and network configuration. We addressed the signalling scenarios for the integration of the QoS signalling with the application signalling, both multimedia and legacy traffic, and both explicit and implicit.

The results showed that implicit signalling was quite fast, on the order of 25ms. Explicit signalling, using RSVP, is also fast, around 35ms, but then it takes the processing time of the RSVP daemons, which was very large (almost 2 sec). Finally multimedia-proxy based signalling, of particular interest for future evolutions of 3G systems, showed reasonable performance, 150 ms. Note however that most of this delay is due to the optimum resource management, in function of the codec, that is achieved through close cooperation between the network and service entities (the QoS broker and the MMSPP).

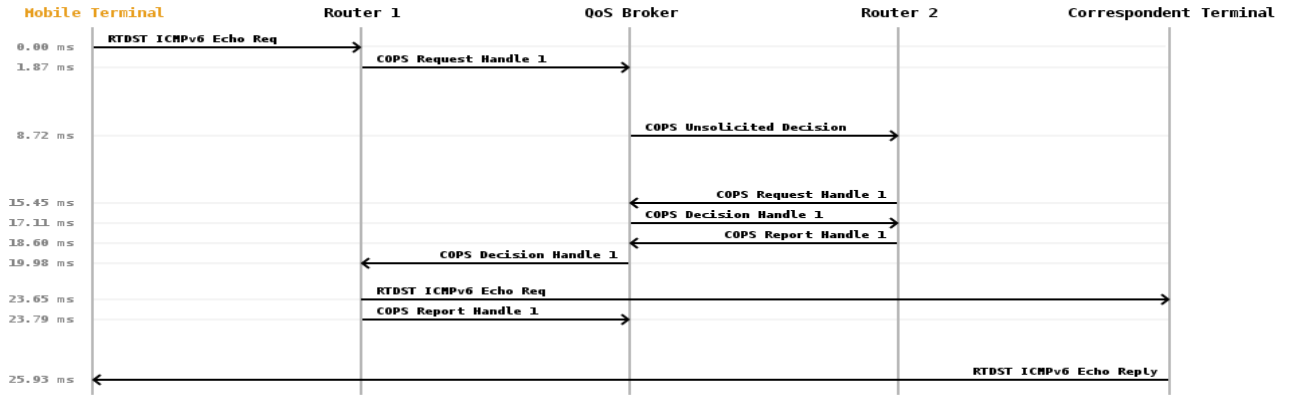
ACKNOWLEDGMENT

The work presented in this paper was funded by the EU project IST-2002-506997 "Daidalos" IV, and developed with the cooperation of multiple partners.

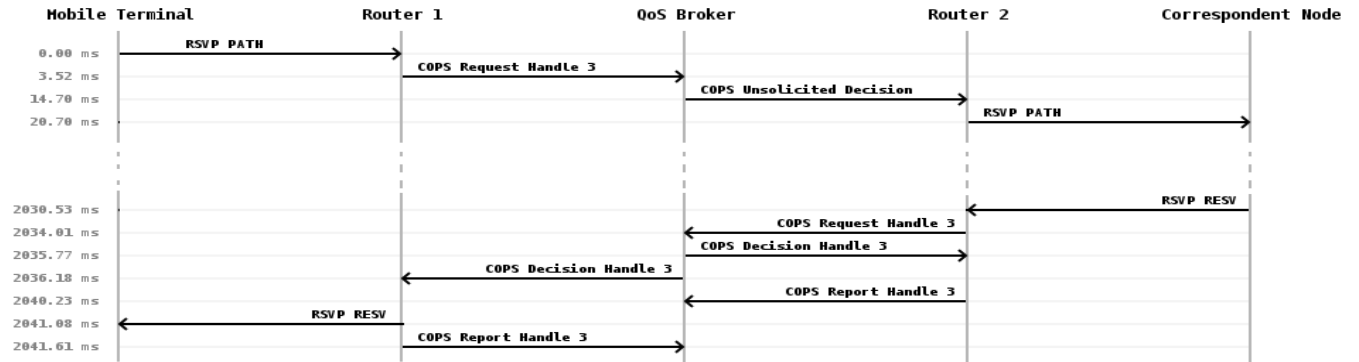
REFERENCES

- [1] Rui L. Aguiar et al, *Daidalos: An Operator and Scenario driven Integrated Project*, IST Mobile Communications Summit 2004, Lyon, Jun 2004.
- [2] R Prior et al, *Heterogeneous Signalling Framework for End-to-end QoS support in Next Generation networks*, HICCS-38, Hawaii International Conference of System Sciences, Hawaii, January 2005
- [3] V. Marques, et al., *An IP-Based QoS Architecture for 4G Operator Scenarios*, IEEE Wireless Communications, June 2003.
- [4] R. Koodli (ed.), *Fast Handovers for Mobile IPv6*. Internet Draft, October 2003.
- [5] S. Blake (ed) et al., *An Architecture for Differentiated Services*, IETF RFC 2475, Dec. 1998.
- [6] ITU-T Recommendation Y.1541, *Network performance objectives for IP-based services*, May 2002.
- [7] R. Braden et al., *Integrated Services in the Internet Architecture: an Overview*, RFC 1633, June 1994.
- [8] D. Gomes, et al, *A transsignaling strategy for QoS support in heterogeneous networks*, ICT'2004, 11th International Conference on Telecommunications, Fortaleza, Brazil, eds: José Neuman de Souza, Petre Dini, Pascal Lorenz, ISBN: 3-540-22571-4, pp 1114-1121, Springer Verlag..
- [9] J. Rosenberg et al., *SIP: Session Initiation Protocol*, IETF RFC 3261, June 2002.

a)



b)



c)

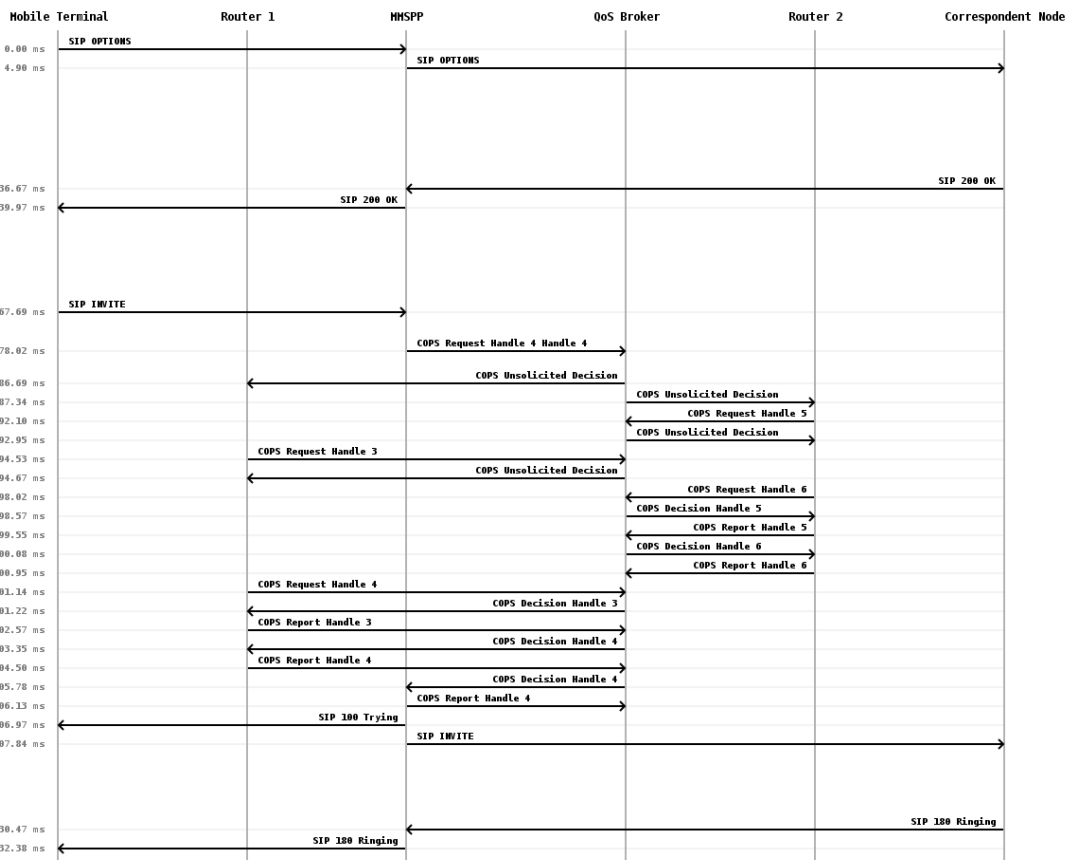


Figure 2: Experimental measures on the signalling flows. From top to bottom: implicit QoS reservation, explicit QoS reservation, multimedia-proxy based reservation

