Telecommun Syst (2007) 35: 55–65
DOI 10.1007/s11235-007-9039-z

QoS abstraction layer in 4G access networks

Gustavo Carneiro · Manuel Ricardo

Published online: 25 August 2007

Abstract Emerging access networks will use heterogeneous wireless technologies such as 802.11, 802.16 or UMTS, to offer users the best access to the Internet. Layer 2 access networks will consist of wireless bridges (access points) that isolate, concatenated, or in mesh provide access to mobile nodes. The transport of real time traffic over these networks may demand new QoS signalling, used to reserve resources. Besides the reservation, the new signalling needs to address the dynamics of the wireless links, the mobility of the terminals, and the multicast traffic.

In this paper a new protocol is proposed aimed at solving this problem—the QoS Abstraction Layer (QoSAL). Existing only at the control plane, the QoSAL is located above the layer 2 and hides from layer 3 the details of each technology with respect to the QoS and to the network topology. The QoSAL has been designed, simulated, and tested. The results obtained demonstrate its usefulness in 4G networks.

Keywords QoS · L2 · Wireless · 4G · Cross-layer

The work described in this article is based on results of IST FP6 Integrated Project DAIDALOS. DAIDALOS receives research funding from the European Community's Sixth Framework Program. Apart from this, the European Commission has no responsibility for the content of this article.

G. Carneiro (✉)
INESC Porto—Instituto de Engenharia de Sistemas e Computadores do Porto, Campus da FEUP,
Rua Dr. Roberto Frias, 378, 4200-465 Porto, Portugal
e-mail: gjcarneiro@gmail.com

G. Carneiro (✉)
e-mail: gjc@inescporto.pt

M. Ricardo
Faculdade de Engenharia, Universidade do Porto, Porto, Portugal
e-mail: mricardo@inescporto.pt

1 Introduction

The "4G Wireless Network" [1] is currently under research following a number of considerations. On one hand mobile nodes will be offered a set of services such as voice and video calls, streaming, and web browsing, which may run on top of IPv6. On the other hand the wireless technology used at any moment will be abstracted; mobile nodes will be allowed to use the best, or multiple, wireless technologies to access the same network, and will be allowed to perform a "vertical handover." For example, a user entering a 802.11 hotspot may switch from the more expensive UMTS access to 802.11, without loss of connectivity. In addition, Quality of Service (QoS) will be built into the architecture from the ground up.

These goals will drive the creation of new functionality, with advantages for both users and operators. Users are offered the flexibility to choose the most convenient access technology available at any given time and place; operators see the technology-specific part of the network decoupled from the rest of the network, and can offer the same services over multiple access methods, with little duplication of network components.

There have been end-to-end QoS solutions for many years. The Integrated Services framework offers a QoS reservation service for individual application flows, also called micro-flows. Scalability concerns, in particular in core networks, led to a new framework some time later, DiffServ, which groups micro-flows with similar QoS into aggregate flows, thus reducing the number of flows to process to a reasonably low value. Traffic engineering techniques using MPLS followed, complementing the previous architectures. At the IP layer, traffic shaping and scheduling mechanisms are necessary in order to provide QoS. In the wireless part of the network, in particular, the IP QoS mechanisms need to be complemented by the QoS mechanisms of the layer 2 technologies. As a consequence, Wireless LAN (IEEE 802.11e), UMTS, and IEEE 802.16 (broadband metropolitan area network), among others, have developed support for L2 QoS.

IEEE 802.11e [2] has been adopted in 2005, adding layer 2 QoS capabilities to the IEEE 802.11 standard (Wireless LAN). In 802.11e the time between beacons is divided into two distinct periods: Contention Period (CP) and Contention Free Period (CFP). The former is analogous to the traditional 802.11 Distributed Coordination Function (DCF), where stations have to compete for access to the medium. But 802.11e defines a new access algorithm called Enhanced Distributed Channel Access (EDCA) that allows four access categories, which are basically different priorities given to stations when competing for access to the radio channel.

However, EDCA still does not provide enough guarantees for the most demanding delay sensitive applications, and an additional access mechanism was defined by 802.11e—the Hybrid Coordination Function (HCF) Controlled Channel Access (HCCA), which operates during the CFP. This mechanism is coordinated by the Access Point and is based on polling. In HCCA the scheduling of polling frames is aimed at satisfying the requirements of Traffic Streams (TS) of stations. A Traffic Stream represents a packet flow associated with a station, and is characterized by a set of QoS parameters. Stations explicitly reserve L2 QoS for a TS using the `MLME-ADDTS` service primitive, which includes not only QoS parameters such as mean data rate and maximum delay, but also the identifiers used to classify packets as belonging to that stream. Several classification mechanisms are supported, including the source or destination MAC, or the IPv6 Flow Label. In addition to the reservation primitive, there is also the primitive `MLME-SCAN.confirm`, which reports the resource status of an AP, such as the number of stations, or the available capacity. This information is carried in 802.11e beacons so that a station can use it to select the least loaded AP among those in range.

In UMTS networks the packet transfer service has been designed with QoS in consideration from the very beginning [3]. UTMS provides primitives to signal the creation of L2 "tunnels" for packet transfer called Packet Data Protocol (PDP) Contexts. These tunnels have an associated set of QoS attributes, and the UMTS network ensures that the packets entering PDP Contexts are transferred according to the negotiated QoS parameters. The QoS parameters include traffic type, transfer delay, and guaranteed bitrate. Among others, UTMS also provides the primitive `SMREG-PDP-MODIFY-IND`, which is used by the network to notify the terminal when a QoS reservation can no longer be maintained with the same parameters, usually due to fading problems in the radio channel. The affected applications can thus adapt to the degradation in the access network in a timely fashion.

In IEEE 802.16 [4] QoS is provided around the concept of *service flow*, a transport service offered by the MAC layer that supports the unidirectional transfer of packets with QoS guarantees. The QoS attributes supported by 802.16 include not only QoS parameters, such as reserved traffic rate and maximum latency, but also an 802.16 specific *scheduling type* parameter that is associated with different MAC scheduling types and different application requirements. With the *best effort service* contention slots are used to transmit data and, as the name suggests, is suitable for best effort traffic. With the *non-realtime polling service* the Subscriber Station (SS) is polled periodically, on the order of one second, for variable sized data transmission, but contention slots may also be used; this service is suitable for high bandwidth non-realtime services, such as FTP. If the *realtime polling service* is selected the SS is polled periodically for transmission of variable sized data, with a polling period tightly controlled to meet the QoS requirements of the service flow. This service is suitable for real-time, variable sized traffic, such as MPEG video. Finally, the *unsolicited grant service* grants SS periodic time slots with no polling, and no latency and overhead associated with polling. A downside of this service is that the station is forbidden from using contention slots for this service flow, which means that only isochronous traffic is allowed, such as Voice over IP. Service flows are dynamically created, usually by the SS, through the *Dynamic Service Addition* (*DSA*) *Request* primitive.

Table 1 summarizes some aspects of the QoS service interfaces for IEEE 802.11e, UMTS, and IEEE 802.16. While leaving out details such as multicast primitives, we can notice the complexity of using multiple technologies in the context of 4G heterogeneous networks. In order to overcome this problem we have developed the "QoS Abstraction Layer" (QoSAL). It is a generic protocol aimed to reserve QoS in L2 wireless networks, which hides from the hosts and the Access Routers the topology and heterogeneity of the underlying L2 network.

The rest of this paper is organized as follows. The following section formalizes the problem that is addressed by the QoSAL. Section 3 presents some similar work in the field. Section 4 presents a high level overview of the QoSAL architecture, which is complemented by a description of the service interface in Sect. 5, and of the protocol in Sect. 6. Simulation results and prototype evaluation are included in Sect. 7. Section 8 concludes with final remarks on the obtained results and future work.

Table 1 Comparison of different L2 QoS interfaces

	802.11e	UMTS (Release 5)	802.16-2004
QoS parameters	Nominal MSDU size Min/mean/max data rate Mean/max service interval Traffic type (isochronous, asynchronous) Burst size	Traffic class (conversational, streaming, interactive, or background) Guaranteed, maximum bitrate Maximum SDU size SDU/bit error ratio Transfer delay	Traffic priority Maximum sustained traffic rate Maximum traffic burst Minimum reserved traffic rate Scheduling type (best-effort, non-realtime polling, real-time polling, unsolicited grant) Tolerated jitter, maximum latency
Traffic classification	Source/destination MAC + ethertype IPv4 DSCP/Protocol IPv6 Flow Label 802.1D user_priority 802.1Q VLAN ID	Source address, mask Protocol number (IPv4)/next header (IPv6) Destination port range Source port range IPSec security parameter index Type of service (ToS) (IPv4)/Traffic class (IPv6) and mask IPv6 Flow Label	IP ToS/DSCP IPv6 Flow Label IP protocol IP source/destination address Source/destination port range Source/destination MAC Ethertype 802.1D user priority 802.1Q VLAN ID
Resource information	Station count, channel utilization, available admission capacity	N/A	N/A
QoS reservations	`MLME-ADDTS.request`, `MLME-ADDTS.confirm`	`SMREG-PDP-ACTIVATE-SEC-REQ`, `SMREG-PDP-ACTIVATE-SEC-CNF`	`DSA-REQ`, `DSA-RSP`
Initiator	Terminal	Terminal	Terminal or network
Flow direction	Unidirectional or bidirectional	Unidirectional or bidirectional	Unidirectional
QoS notifications	N/A	`SMREG-PDP-MODIFY-IND`	`DSC-REQ`, `DSC-RSP`

2 Problem statement

The basic scenario that we addressed consists of a Mobile Node that uses an Access Router to obtain connectivity to an operator's network. However, the Access Router does not offer direct wireless connectivity to the Mobile Node; instead, it is connected to one or more Access Points, which provide the layer 2 connectivity to the Mobile Node.

A more complex scenario is characterized by concatenated wireless links, as shown in Fig. 1. Here an IEEE 802.16 link (Base Station + Subscriber Station) is used to connect two Access Points in parallel: one 802.11 (Wireless LAN) and one IEEE 802.15.1 (Bluetooth). These scenarios are only examples; the ultimate topology addressed by the QoSAL is a generic L2 network composed of an arbitrary number of equipments interconnected to form a tree rooted at the Access Router.

The following goals were considered for the design of the QoSAL:

1. The main goal, to allow reservation of L2 QoS resources in wireless networks.
2. A flexible QoS model which adopts the common QoS models in the Internet: IntServ and DiffServ.
3. The parameters of the QoS model should be simple and generic, so that each Access Point can map them into its own technology.
4. Multiple concatenated wireless links should be supported, and QoS should be reserved in all of them.
5. No a priori knowledge of the L2 network topology should be required by either Access Router or Mobile Node.
6. Dynamic changes in the L2 topology should be supported.

3 Related work

Although bearing many similarities with the work described in this paper, the Subnet Bandwidth Manager (SBM) protocol/framework developed by the IETF Integrated Services over Specific Link Layers (ISSLL) group is tied to the RSVP protocol and not directly reusable outside the IntServ framework. SBM uses IP multicast packets for signalling, and because of that it is very much IP-oriented and does not fit the conceptual model of MAC bridges. In addition, there can be only one Designated SBM (DSBM) in each LAN segment, and the router uses a simple predefined multicast address to communicate with the DSBM. This means that supporting complex L2 topologies in the same LAN segment (as

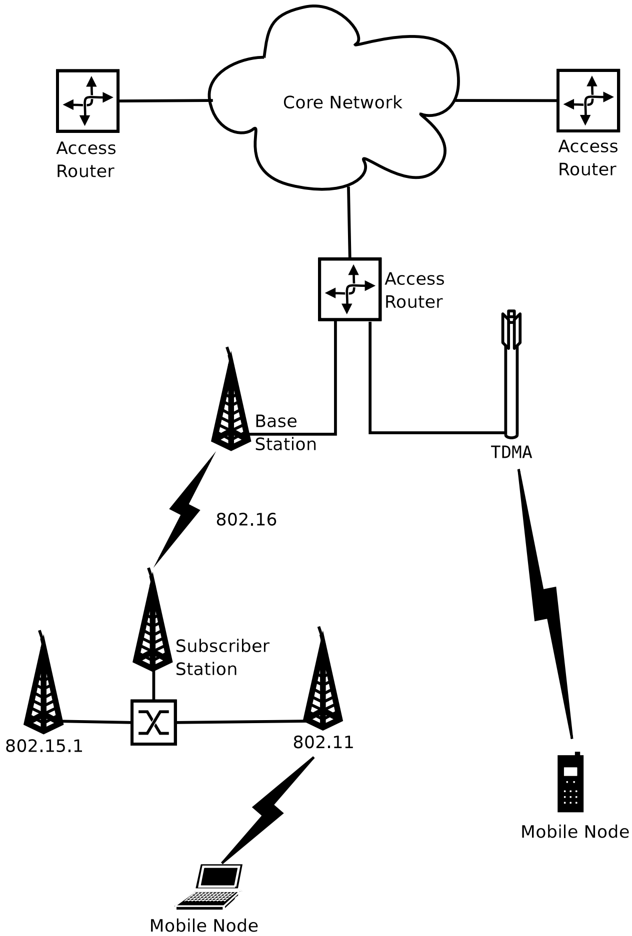


Fig. 1 Example of multi-hop wireless scenario

perceived by the Access Router), such as the one in Fig. 1, is not possible without gross simplification and loss of efficiency. Finally, SBM only performs basic admission control procedures and, at most, 802.1Q user_priority marking, for priority-based L2 QoS, while certain applications demand a more detailed QoS reservation model.

The IEEE 802.21 Working Group is currently working on handover and interoperability between heterogeneous networks. The only overlap in functionality with the work presented here is the definition of abstract link layer indications. In 802.21, QoS issues are hardly addressed at all; it does not contain primitives to reserve L2 QoS resources. Supporting concatenated networks (e.g. IEEE 802.16 + 802.11) with the current 802.21 service model would be challenging. It also does not contain information on “available bitrate”, which is a fundamental parameter in “network initiated handover”.

Unlike IEEE 802.21, QoSAL does not target mobility as main focus; while mobility is supported by QoSAL, it assumes the presence and cooperation of existing mobility management modules, such as Fast Handovers for Mobile IPv6.

The Unified Link Layer API (ULLA) [5] features an abstract API for configuring radio link parameters and receiving generic handover triggers. It is similar to 802.21 in purpose, but is more limited in the sense that it is a host local interface; no protocol is defined for delivering commands or events across L2 network segments. Like 802.21, it does not cover L2 QoS.

4 Architecture

The architecture of the QoSAL is shown in Fig. 2. It represents an access network, with an Access Router, an Access Point, and a Mobile Node. The QoS Abstraction Layer (QoSAL) runs on these network elements, and its instances communicate using a protocol that is transported directly over L2 bearers. At the Access Router, the QoSAL accepts service requests from a L3 entity called *QoS Manager*. The QoS Manager is responsible for end-to-end QoS management, but it only addresses IP subnetworks; for instance, it is unaware of the Access Points in Fig. 1, so the reservation of resources in the L2 network is delegated to the QoSAL. A more in depth overview of the L3 architecture can be found in [6].

The QoSAL exists only in the control plane of the communications stack. In response to abstract QoS requests from the QoS Manager, the QoSAL modules running on Access Point and Mobile Node ask the technology-specific *QoSAL Driver* modules to implement the request at L2, which configure the data-plane modules (e.g. shaping and scheduling), and prepare them for the new flow. For instance, in an UMTS network interface, at the Mobile Node side, the primitive `SMREG-PDP-ACTIVATE-REQ` is used to reserve L2 QoS resources. A more detailed description of QoSAL drivers for IEEE 802.11e, 802.16, and Bluetooth can also be found in [6].

At the IP layer individual flows may have separate queues where packets are shaped and their IPv6 *Flow Label* field is marked with the *Connection Identifier*, which is a number used by the QoSAL to uniquely identify a flow. Then packets are multiplexed into a single queue leading to the network interface driver. From then on the QoSAL-aware network interfaces, not only at Access Router and Mobile Node but also at Access Points, use the Flow Label to demultiplex packets into different L2 queues, and apply different QoS treatment to different flows. Currently only the IPv6 Flow Label is used to associate packets with QoS connections, but a mapping based on L2 or even new L2.5 headers is being researched.

5 The service interface

The services offered by the QoSAL to layer 3, such as the QoS Manager in Fig. 2, can be classified into five groups:

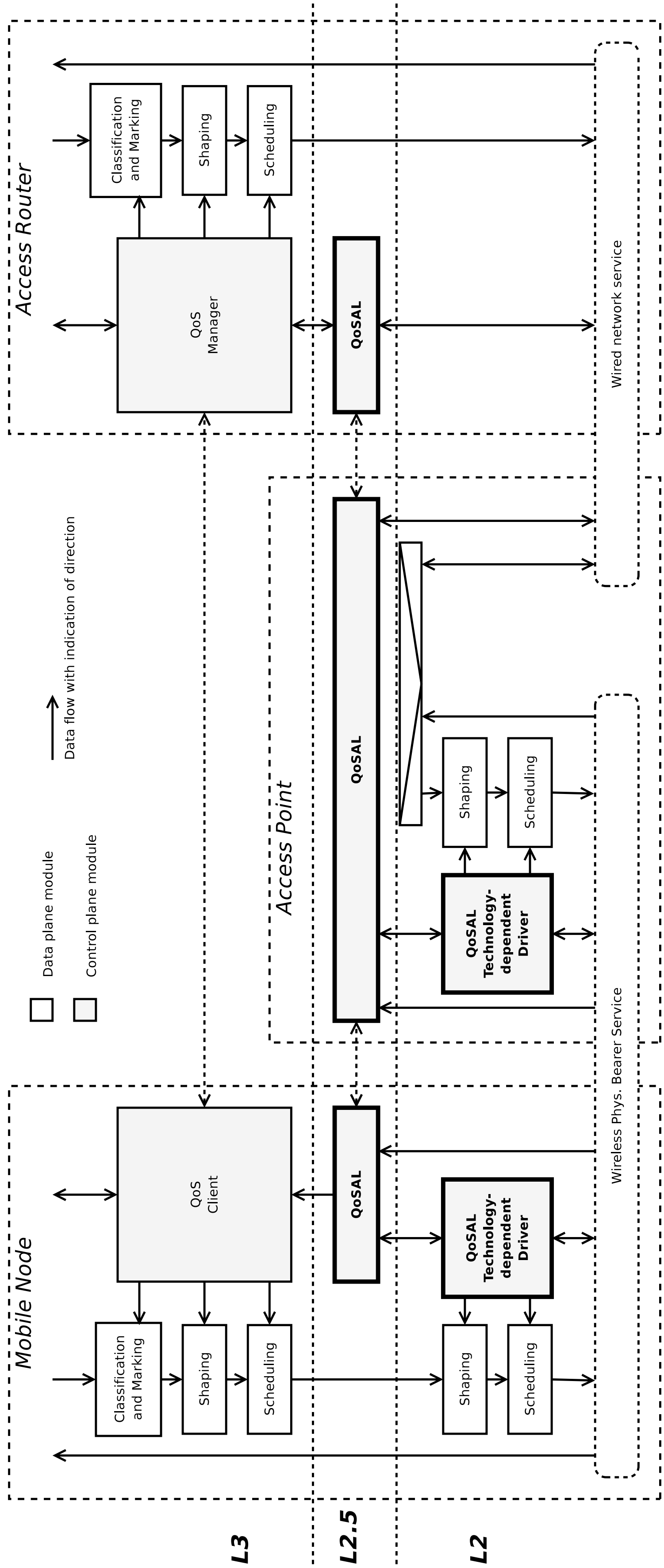


Fig. 2 QoS Abstraction Layer architecture

(1) QoS reservation; (2) resource querying; (3) QoS degradation notification; (4) mobility; (5) multicast.

5.1 QoS reservation

The main service offered by the QoS Abstraction Layer consists in the creation of "QoS connections" between the Access Router and the Mobile Node. These "connections" are virtual channels between the two elements offering QoS guarantees, such as bitrate and delay.

The service of QoS reservation is offered by the primitive `AL-CNX-ACTIVATE-REQ`. The parameters required for this primitive are (1) the MAC address of the destination Mobile Node that terminates the QoS reservation, and (2) the Quality of Service parameters.

The QoS parameters include a Tspec and an Rspec. The Tspec describes the characteristics of the flow to be transferred, and includes the parameters *mean data rate*, *burst size*, *peak data rate*, *minimum policed unit*, and *maximum transmission unit*. These are standard *token bucket filter* parameters that have the meaning defined in [7]. The Rspec describes the actual QoS resources being requested to the link layer. It consists solely of two values: *Reserved Bitrate* and *Class Identifier*. The Class Identifier implicitly determines the values of many other QoS attributes, such as transmission delay and error rate, considering the application requirements. The values considered for Class Identifier are:

- *Conversational.* Interactive voice and video (e.g. audio and video conferencing)
- *Transactional.* Transaction data, interactive (e.g. WEB browsing, telnet, e-commerce)
- *Streaming.* Short transactions, bulk data, video streaming (e.g. video on demand, FTP)
- *Best effort.* Legacy applications/low cost services

The `AL-CNX-ACTIVATE-REQ` primitive triggers the signalling described in Sect. 6, and returns a *connection identifier* in an `AL-CNX-ACTIVATE-RESP` primitive if the reservation succeeds. The entity that requested QoS reservation is then responsible to setup the flow marking module that will mark the Flow Label of the IPv6 packets associated to the connection. The connection identifier is also used for modification and deactivation of QoS connections, and it is included in QoS degradation notifications.

QoSAL offers a reservation mechanism that is adequate for both DiffServ and IntServ. To support DiffServ the end-to-end QoS management entity (e.g. the QoS Manager) should create, for each terminal, one QoSAL reservation per class of service, and adjust the amount of reserved bitrate of these reservations dynamically to accommodate evolving application requirements.

5.2 Resource querying

There is a primitive used to request a report of the available (free) resources in the access network: `AL-RESOURCE-QUERY`. Its single parameter, `DestAddr`, identifies the "path" for which resources are to be queried. If `DestAddr` is the L2 address of a Mobile Node a single report is generated for all Access Points in the path towards the given Mobile Node. If, on the other hand, the address is that of an Access Point then the reply will report resources along the path to the give Access Point. Finally, a broadcast address as `DestAddr` means to request reports for all Access Points in the Access Network, one report for each unique path towards an edge Access Point.

The resource reports will arrive some time later via `AL-RESOURCE-INDICATION` up-calls. This primitive includes an estimation of free bandwidth in the Access Point; if the path includes multiple Access Points, the returned bandwidth is the minimum of all bandwidths of individual Access Points. `AL-RESOURCE-INDICATION` can also be sent spontaneously by the QoSAL, either periodically or whenever significant changes in the available resources are detected.

5.3 QoS degradation notifications

The primitive `AL-CNX-INDICATION` is issued at the Access Router and the Mobile Node by the QoSAL to indicate that it was forced to modify the QoS for a specific connection due to changing conditions in the wireless medium. The primitive includes the connection identifier and Rspec as parameters. This primitive can be used for cross-layer adaptation [8], and may serve as trigger for higher-level handover decisions.

During the lifetime of a connection, the QoSAL is allowed to spontaneously modify the reserved bitrate within the bounds of the initial reservation. For instance, the QoSAL may sense a degradation in the signal strength, and temporarily switch to a transmission mode with lower bitrate but lower error probability; since the actual bitrate obtained by the connection has decreased, an `AL-CNX-INDICATION` primitive is issued at Mobile Node and Access Router to notify the QoS Manager and/or applications of this. When the effect that caused the degradation ceases, the link is again gradually switched to higher bitrate modes, new QoS levels become possible in the connections, and `AL-CNX-INDICATION` primitives are sent again, until the initial QoS level is attained.

5.4 Mobility

In order to support node mobility scenarios a couple of additional primitives were defined. `AL-HANDOVER-PREPARE` is used when the Mobile Node makes a decision to handover to a different Access Point. When the QoSAL receives this primitive a message is sent to the new Access Point notifying it to prepare the QoS resources in advance. Then, as soon as the Mobile Node switches to the new Access Point, the primitive `AL-HANDOVER-EXECUTE` is issued by the mobility control module, and the QoSAL at Mobile Node communicates with the new Access Point in order to activate the resources previously prepared. A smooth handover is obtained, and QoS is preserved under mobile conditions.

5.5 Multicast

The support of multicast QoS will be important in future networks, as multimedia streaming services are likely to gain prominence. Therefore, multicast support had also to be designed into the QoSAL.

There are two stages required to obtain multicast QoS. First, a QoS reservation is performed using the `AL-CNX-ACTIVATE-REQ` primitive, but specifying a L2 multicast address instead of the L2 address of a Mobile Node. As usual, a connection identifier is returned, which identifies the *multicast session*.

After a multicast session is created, membership management takes place. The primitive `AL-MULTICAST-JOIN` takes a connection identifier and a L2 address as parameters, and it is used to notify the QoSAL that a new Mobile Node is joining the indicated multicast session. Conversely `AL-MULTICAST-LEAVE` is used to indicate that a Mobile Node is leaving a multicast session.

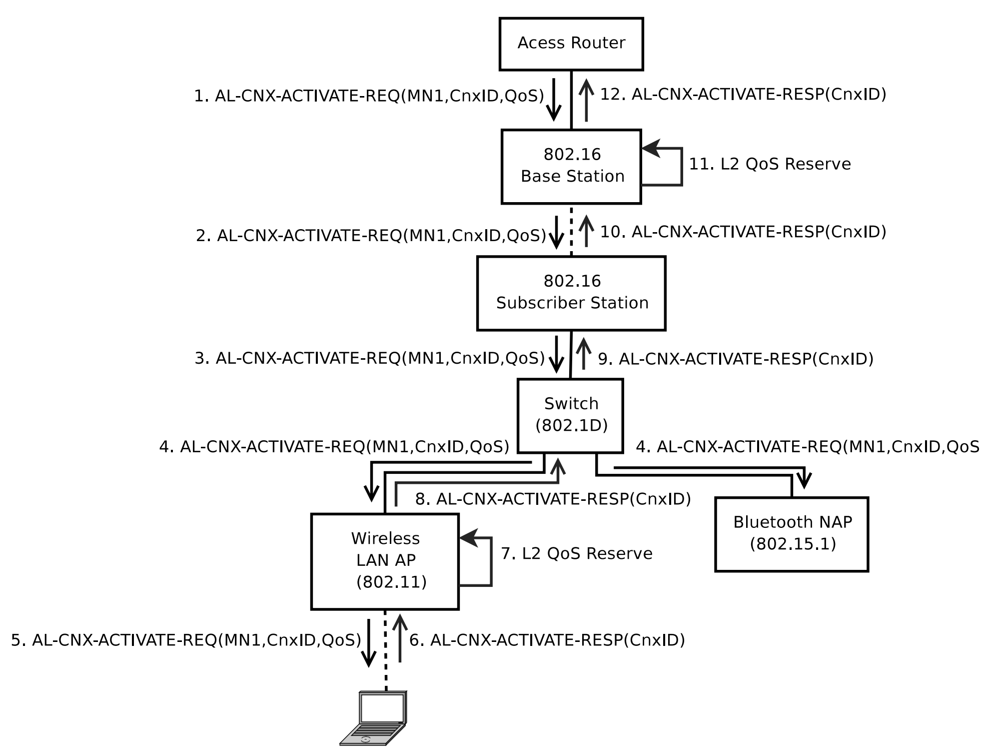
6 The protocol

The QoSAL Protocol Data Units (PDUs) are transmitted as IEEE 802 frames[1] with a dedicated protocol type. It is inspired by the RSVP protocol in the sense that it uses in-band signalling and soft-state. It should be noted, however, that RSVP and QoSAL have different and complementary scopes; the former addresses end-to-end L3 QoS, while the latter covers only L2 access networks.

Figure 3 illustrates how the `AL-CNX-ACTIVATE-REQ` primitive is handled by the protocol. First, a unique identifier for the QoS reservation, CnxID, is derived. Then an Ethernet frame is sent, with MN1 as destination MAC address. The frame contains an `AL-CNX-ACTIVATE-REQ` PDU, with CnxID and QoS as parameters. It is important to emphasize that by putting the Mobile Node address as destination address of the signalling frames, the design goals 4–6 in Sect. 4 are met. This comes "for free" with the IEEE 802.1D Learning Bridges, implemented by switches and wireless Access Points.

[1] Other L2 networks can be supported with little additional effort.

Fig. 3 QoS connection activation example



The request PDU follows the path towards the Mobile Node traversing a number of Access Points. However QoSAL-enabled Access Points automatically recognize the AL PDU by its Ethernet protocol number, and pass it to the QoSAL code for special processing, before allowing it to be forwarded. No reservation is performed at this point, though; the Access Point has to wait for an `AL-CNX-ACTIVATE-RESP` primitive, at which point the reservation is committed, in case of a successful response from downstream.

The reason why the confirmation is required before committing the reservation is related to the Learning Bridge algorithm. A Learning Bridge is not programmed with any routes; it learns them from the received traffic, and maintains a cached table of MAC address/output port. This property is both a strength, since it allows it to work with no previous configuration of routes, and a weakness because it is not guaranteed to always know the route for a particular destination. When a Learning Bridge does not know the route for a particular MAC address, it simply forwards the frame through all other output ports, as in our example in Fig. 3. Thus, forcing the Access Point to wait for a response from the Mobile Node before acting on the request prevents this (infrequent) situation.

There is an analogy between `AL-CNX-ACTIVATE-REQ` and RSVP's *Path* messages, and between `AL-CNX-ACTIVATE-RESP` and RSVP's *Resv*. Like RSVP, the QoSAL protocol is soft-state, which means that QoSAL connections must be periodically refreshed, or else they expire. In addition, the QoSAL protocol works even in the presence of L2 nodes that do not support QoSAL. In these nodes QoSAL PDUs are treated as simple Ethernet frames; no special processing is done, but they eventually reach the other QoSAL-aware nodes. This allows QoSAL to be used even in environments where QoS support has not been fully deployed.

Regarding the resource query primitive described in Sect. 5.2, the three reservation styles are simply mapped into Mobile Node, Access Point, and broadcast addresses. Whatever the addressing scheme used an `AL-RESOURCE-QUERY` PDU is sent from the Access Router towards the requested destination. When this PDU reaches either the destination or an edge Access Point its forwarding stops and a response is sent back to original requester. The response consists of an `AL-RESOURCE-INDICATION` containing a bandwidth parameter. As the message passes through the Access Points, each Access Point combines its own available bandwidth with the value found in the message. The overall available bandwidth, reported to the Access Router, is equal to the minimum of all the bandwidths in each Access Point.

7 Validation and results

7.1 Simulations

In order to assess the scalability of the QoSAL protocol a simulation study has been conducted. The scenario that was simulated consisted on an Access Router connected through a FastEthernet link to an 802.11e Access Point with two Mobile Nodes. The SimPy[2] simulator was used. The Ethernet link was simulated using the fast-easy method described in [9] with a maximum delay of 3.36 ms, while the 802.11 link was approximated by an exponential delay distribution [10].

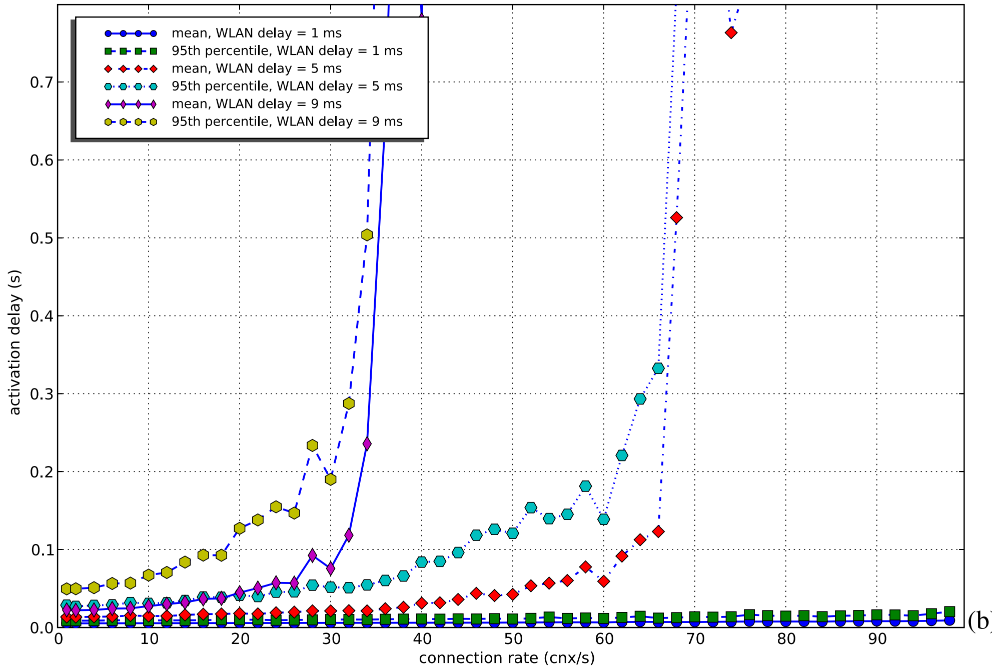
Taking as reference the 3GPP approach to characterize signalling performance [11] we have characterized the delay of activating a connection, and how that delay evolves with the arrival rate of connections (the parameter λ of a Poisson process) and with the WLAN average delay. Fig. 4(a) shows the results for a simple DCF MAC. Looking at 3GPP performance metrics, in particular the "through connection delay" for the "internal and terminating traffic" case we consider that the mean delay must be below 250 ms, and its 95th percentile must be below 300 ms. In the simulation those limits correspond to a connection rate of about 30 cnx/s when the WLAN delay is 9 ms, or 64 cnx/s in case the WLAN delay is 5 ms.

It could be argued that the 802.11 average delay can become much larger than the simulated 9 ms. However, as shown in [12], in 802.11 networks when the load increases above 70%, corresponding to about 10 active stations and a delay below 4 ms, delays increase drastically. We may thus assume that operators will attempt to limit, through load balancing and network initiated handover, the number of active users in any given Access Point.

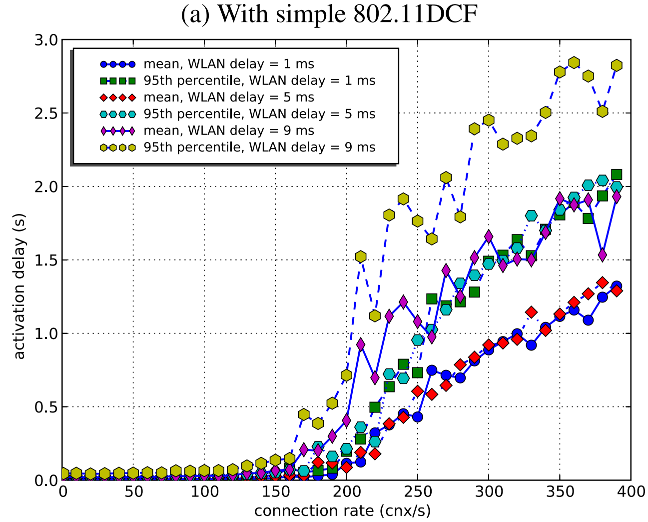
Considering that QoSAL PDUs are relatively short (about 60 octets), and knowing that the medium access delay represents the biggest portion of the delay in WLAN transmission for short packets, we decided to investigate the impact of the Contention Free Burst (CFB) option of 802.11e. When CFB is available, a parameter called $TXOP_{\text{limit}}$ is defined; when a station gains access to the medium, it is allowed to transmit an arbitrary number of consecutive frames in a single go, without having to compete for the medium between each frame, for a maximum period of $TXOP_{\text{limit}}$. The simulation results in Fig. 4(b) were obtained with an 802.11e $TXOP_{\text{limit}} = 1$ ms (a rather conservative value) and $PHY_{\text{rate}} = 2$ Mbit/s. As expected, scalability improves considerably with this option, with over 150 cnx/s possible, even with a 9 ms WLAN delay, while still meeting the 3 GPP performance goals.

[2] http://simpy.sourceforge.net/.

Fig. 4 Simulation of connection activation delays



(a) With simple 802.11DCF



(b) With 802.11e CFB

7.2 Prototype

A prototype of the QoSAL has been developed, with the goal of further validating the protocol and its implementation. The testbed consisted of three PCs running the Linux operating system. One plays the role of Access Router, and is connected to the second one, the Access Point, with an Ethernet cable. The Access Point has a second Ethernet interface which connects to the third PC acting as Mobile Node, and Ethernet bridging is performed between these two interfaces. Figure 5 shows the QoSAL messages, which were captured during a test using `tcpdump` at the Access Router. These messages are represented as an Ethereal view, for which a custom protocol dissector was developed. We can see QoSAL PDUs for activation, modification, and deactivation, as well as resource indications.

8 Conclusions

In this paper the key functions behind the QoS Abstraction Layer were presented. It is a Layer 2.5 signalling protocol for reserving QoS resources across a wireless access network. It leverages the Learning Bridge property of IEEE 802 wireless Access Points and IEEE 802.1D switches to pro-

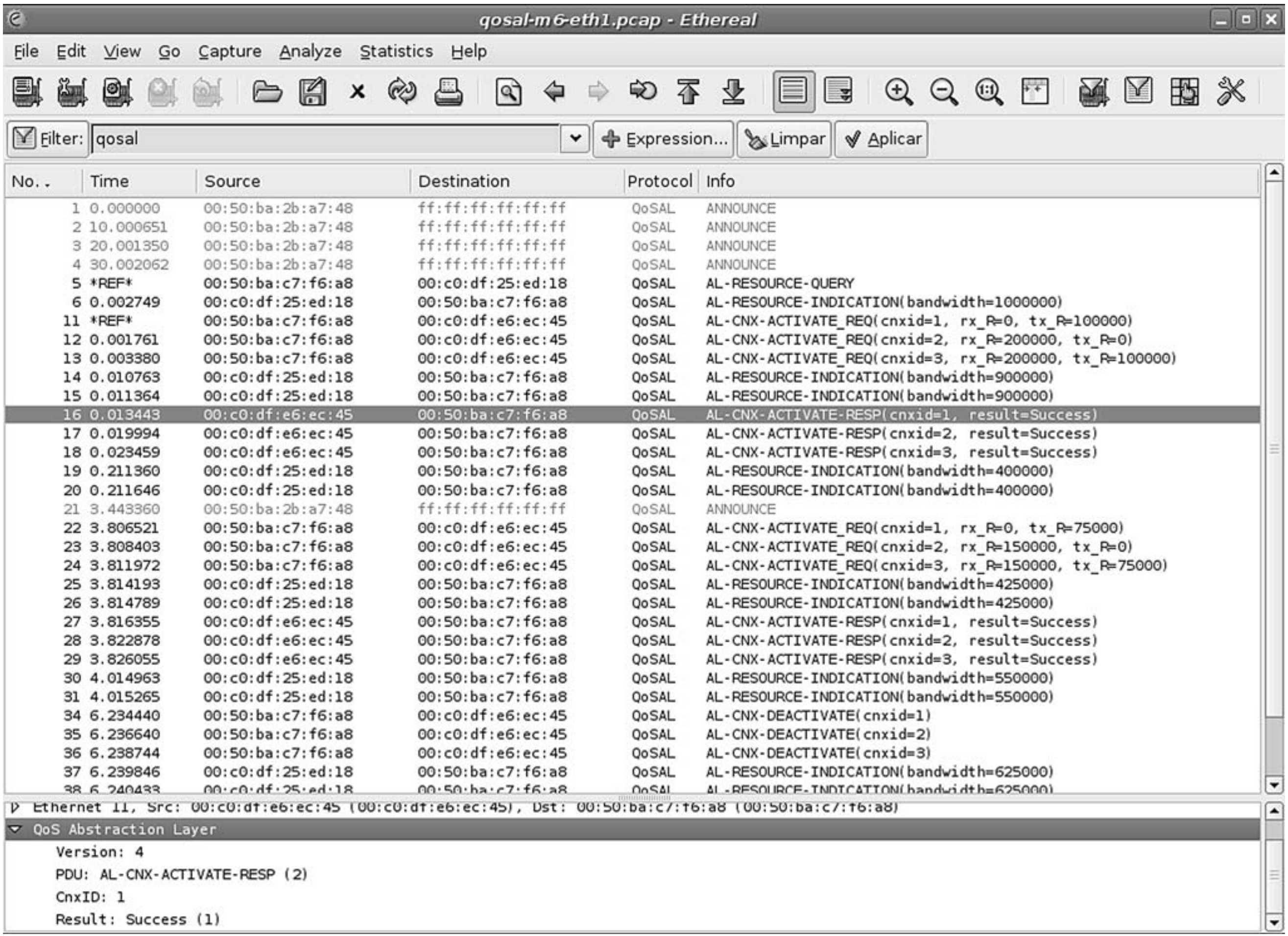


Fig. 5 Ethereal packet analysis of QoSAL signalling

vide automatic discovery of the path for the reservations towards a given terminal. It works with no prior knowledge of the topology of the access network, which may even include multiple concatenated Access Points. Also provided is the ability to query available resources in the network, receive QoS degradation notifications, and make multicast reservations. The concept has been validated through simulations and prototyping, which also provided valuable insights into possible optimizations.

Future work in this protocol includes the design and implementation of an optimization consisting on coalescing multiple similar requests in a single message, which would provide benefits particularly in 802.11 networks without CFB. In addition, QoSAL drivers for 802.11e, 802.16, and 3GPP TDMA, already under development, will continue to be improved. Also some research will be devoted into alternatives to replace the use of IPv6 Flow Label for L3/L2 QoS mapping, in data plane, to avoid the dependency of a L2.5 protocol (QoSAL) on a L3 protocol (IPv6). Finally, integration of security and header compression functionality into QoSAL is being researched.

References

1. Frattasi, S., Fathi, H., Fitzek, F. H. P., Prasad, R., & Katz, M. D. (2006). Defining 4G technology from the user's perspective. *IEEE Network*, *20*(1), 35–41.
2. IEEE Computer Society (1999). IEEE standard for information technology: Telecommunications and information exchange between systems. Local and metropolitan area networks. Specific requirements part 11: wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications Amendment 8: Medium Access Control (MAC), Quality of service enhancements. IEEE Std 802.11e-2005 (Amendment to IEEE Std 802.11) (Reaff 2003, 2005).
3. 3GPP (2002). QoS concept and architecture. TS 23.107, Release 5.
4. IEEE (2004). IEEE standard for local and metropolitan area networks part 16: Air interface for fixed broadband wireless access systems.
5. Farnham, T., Gefflaut, A., Ibing, A., Mähönen, P., Melpignano, D., Riihijärvi, J., & Sooriyabandara, M. (2005). Toward open and unified link-layer API. In *Proceedings of the 14th IST mobile and wireless summit*, Dresden, June 2005.
6. Carneiro, G., García, C., Neves, P., Chen, Z., Wetterwald, M., Ricardo, M., Serrano, P., Sargento, S., & Banchs, A. (2005). The DAIDALOS architecture for QoS over heterogeneous wireless

networks. In *Proceedings of the 14th IST mobile and wireless summit*, Dresden, June 2005.
7. Shenker, S., & Wroclawski, J. (1997). General characterization parameters for integrated service network elements. RFC 2215, IETF.
8. Carneiro, G., Ruela, J., & Ricardo, M. (2004). Cross-layer design in 4G wireless terminals. *IEEE Wireless Communications*, *11*(2), 7–13.
9. Wang, J., & Keshav, S. (1999). Efficient and accurate ethernet simulation. In *Conference on local computer networks, LCN '99*, Lowell, MA, October 1999 (pp. 182–191). Washington, DC: IEEE Computer Society.
10. Zhai, H., Kwon, Y., & Fang, Y. (2004). Performance analysis of IEEE 802.11 MAC protocols in wireless LANs. *Wireless Communications and Mobile Computing*, *4*(8), 917–931.
11. 3GPP (2004). Technical performance objectives. TS 43.005, Release 5.
12. Ni, Q. (2005). Performance analysis and enhancements for IEEE 802.11e wireless networks. *IEEE Network*, *19*(4), 21–27.

Gustavo Carneiro received a Diploma degree in Electrical and Computer Engineering from Porto University, Portugal, in 2001, and since then he is a researcher at INESC Porto, and completed a Master in Network and Communications Services in 2006. He has actively participated the European IST ARROWS project and, currently, he participates in the European IST DAIDALOS project. In addition, he has enrolled in a Ph.D. program in Electrical Engineering in University of Porto.

Manuel Ricardo received a Diploma degree in 1988, a M.S. in 1992, and a Ph.D. in 2000, all in Electrical and Computer Engineering from Porto University, in Portugal. He is an Associate Professor at the Faculty of Engineering of Porto University, where he gives courses in Mobile Communications, and Computer Networks. He also leads the Communication Networks and Services Area at INESC Porto.