# Hybrid Quality of Service Architecture for Wireless/Mobile Environment

Li Zheng, Arek Dadej, Steven Gordon
*Institute for Telecomunications Research, University of South Australia*

Abstract: Provision of Quality of Service (QoS) guarantees for multimedia applications over IP networks is rapidly becoming a critical research and design issue, especially within the radio access segment of the next generation wireless/mobile environment. While the simple and scalable Differentiated Services (DiffServ) QoS control model is suitable for the core part of the network, more explicit, admission and reservation based QoS mechanisms are required in the wireless access segment of the network where the resources available and the levels of traffic aggregation render the DiffServ principles less effective. We present a suitable hybrid QoS architecture framework, and then study the relevant issues further for a radio access network based on IEEE 802.11 WLAN with 802.11e QoS extensions.

Key words: Quality of Service; QoS consistency; Third Generation Mobile Network; IEEE 802.11 WLAN; Differentiated Services; Integrated Services; Medium Access Control.

## 1. INTRODUCTION

As the deployment of Third Generation (3G) mobile networks and Wireless Local Area Networks (WLANs) rapidly becomes a reality, there is a growing demand for multimedia services over these wireless systems. Such services, especially real-time ones, require at least the same level of Quality of Service as that provided by the wired Internet for which most of the IP-based multimedia applications were originally designed. This leads to a problem of Quality of Service (QoS) consistency across the wireless and wired segments of the network [1, 2, 3].

Radio access network is usually seen as the bottleneck in the end-to-end data path. The QoS limitations of the wireless segment originate from the inherent properties of mobile radio environment.

For delay sensitive or bandwidth intensive applications, resource constraints over the air interface may become a significant technical challenge in the design of new generation wireless multimedia networks. Provided that the total resources available over the air interface are, on average, sufficient to meet the total resource requirements of the user application sessions admitted to the system, the level of QoS desired/expected by users can be provided on an end-to-end basis by means of service differentiation (i.e. sacrificing the system performance for services tolerant of longer delay and higher rates of data loss in order to meet the quality of service specified for the other, less tolerant services).

QoS mechanisms over the wireless segment of the network have to be designed with consideration given to the entire (end-to-end) network QoS. Various end-to-end QoS architectures for networks involving wireline core/transport and wireless access segments have been proposed and discussed in the literature. Some researchers argue for DiffServ principles uniformly applied end-to-end throughout the entire network [e.g. 4, 8, 9], others represent the view that more subtle and explicit QoS control mechanisms are required at the radio access level [10]. A number of proposals have been published for hybrid QoS architectures involving the use of RSVP/IntServ end-to-end QoS control over DiffServ based networks [e.g. 5, 12, 15]. The drawbacks of QoS architectures proposed thus far vary from insufficient level of control implemented with DiffServ model only, especially over the resource-limited air interface, to scalability and complexity problems of solutions involving end-to-end use of RSVP/IntServ model.

In this paper, we present a hybrid QoS architecture framework suitable for new generation wireless IP networks. It is likely to achieve the goals of QoS consistency across the network, as well as scalability and simplicity of QoS control. The proposed hybrid architecture follows the principles of Differentiated Service (DiffServ) model [13] over the core part of the network, and the principles of Integrated Services (IntServ) model [14] locally over the wireless access segment. Unlike the IntServ/RSVP over DiffServ networks models [e.g. 5, 12, 15], our hybrid model does not apply the IntServ principle end-to-end, thus avoids the scalability and complexity problems usually associated with the end-to-end IntServ based QoS control.

We will examine our hybrid QoS control model in the context of IEEE 802.11 wireless access network.

## 2.      HYBRID QOS ARCHITECTURE PRINCIPLES

While the DiffServ model is useful in providing efficient and scalable QoS control within the network segments characterised by high volume of available resources and high aggregation of traffic (i.e. core/transport network), it fails to provide subtle enough tools for controlling QoS where the resources are strongly limited and the levels of traffic aggregation are low, such as in the wireless access network. The last hop (wireless access) radio resource management cannot rely solely on mechanisms providing differentiated treatment of packets that belong to different application sessions. To avoid degradation of QoS as the traffic generated by the users within the same access network increases, a mechanism is needed at the access network level to control the total resource requirements of the sessions admitted to the system (explicit admission control), and to reserve the amount of resources required by each session. Such mechanism, operating on a session-by-session basis, is characteristic of the IntServ QoS model.

The above argument serves as a brief justification for our choice of hybrid model with DiffServ principles applied over the core/transport network domain, and IntServ principles applied locally to the QoS control over the wireless access segment.
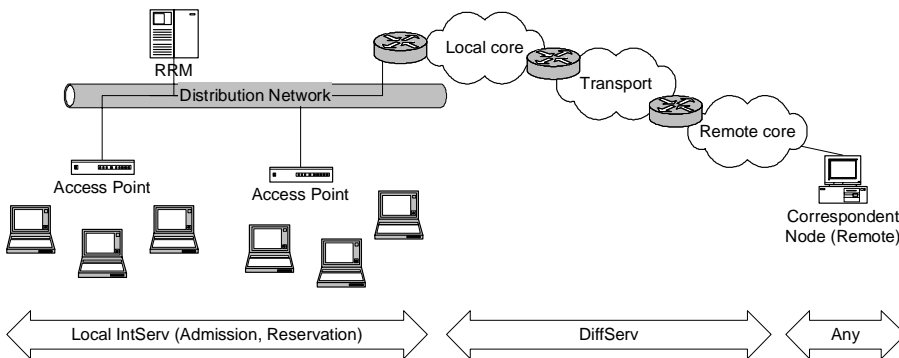


*Figure 1*. Hybrid QoS Architecture

The proposed QoS control architecture (*Figure 1*) comprises of the Differentiated Services part in the core/transport network segments, and explicit resource management (admission control and reservation) part in the radio access network. Our architecture does not presume any specific QoS control model in the remote network where correspondent node (the other party in the application session) is located; it assumes that it is the other

network's responsibility to guarantee at its end a QoS level consistent with that in the remaining parts of the path. The explicit resource management is localised to a single radio access network domain, where stateful and fine-granularity control mechanisms operating at the level of individual flows and application sessions can be applied without causing scalability and complexity concerns.

The resource management at the access network level is based on functional blocks typical of IntServ model, i.e. admission controller and packet (frame) classifier, with multiple queues and service disciplines used to enforce QoS guarantees given to the flows (sessions) upon admission. However, unlike the IETF IntServ architecture, it does not use explicit end-to-end path establishment and resource reservations such as those available with RSVP.

Because of being limited in its scope to the last hop, the signalling required by the reservation based mechanism does not have to be implemented by means of explicit application level signalling protocol; it can be easily implemented at the medium access control level, as part of MAC requests. In case of 3G access networks, the MAC-level admission control and reservation signalling becomes a part of radio resource management necessary to handle admissions of individual terminals and changes of their link states. In case of 802.11 WLANs, the explicit resource management cannot be easily and reliably achieved with the standard MAC data and control frames and random access based Distributed Coordination Function commonly implemented in the current 802.11 products. It will require either application level signalling involving exchange of control packets between the wireless station, access point and WLAN-wide radio resource management entities, and IP-level service differentiation mechanisms, or QoS extensions to the 802.11 MAC layer discussed later in this paper.
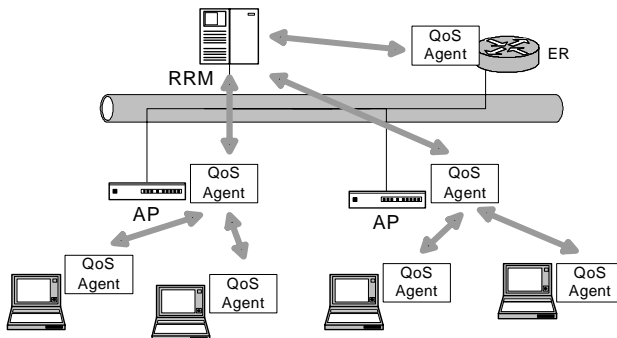


*Figure 2.* QoS agents

To enable exchange of QoS signalling at the local radio access network level, all major entities in the network (wireless stations, access points, edge router) are equipped with QoS agents (*Figure 2*). The role of QoS agents (management plane processes) is to capture QoS requirements known to the application agents and/or application session control processes, and facilitate transfer of the QoS related information to QoS processors (classifiers, schedulers), as well as exchange of QoS signalling with other QoS (Resource Management) entities.

## 3. QOS SUPPORT IN 802.11 WIRELESS LANS

### 3.1 802.11 wireless LANs

The IEEE 802.11 standard, approved in its basic form in 1997, has now become a widely accepted standard for wireless local area communications. It specifies the medium access control (MAC) and physical (PHY) layers for wireless communication between wireless station(s) (STA) and access point(s) (AP). Enhanced data rates have been introduced in the 802.11b (up to 11 Mbps, using Direct Sequence Spread Spectrum in the 2.4GHz ISM band) and later in the 802.11a (up to 54 Mbps, using Orthogonal Frequency-Division Multiplexing in 5 GHz ISM band).

### 3.2 The IEEE 802.11 Medium Access Control

The 802.11 MAC layer specifies two modes of operation: Distributed Coordination Function (DCF) and Point Coordination Function (PCF) [19]. In the DCF mode, the only mode of operation implemented in most of the commercially available 802.11b products, stations contend for transmission opportunities by following the principles of Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA). Before transmission, a station must listen to the channel. If the medium (channel) is sensed free for a time interval greater than the DCF Inter Frame Space (DIFS), then the station can transmit a DATA frame. Otherwise, the transmission is deferred until the channel is sensed free, after which an additional backoff time is applied. The backoff interval is randomly selected from a range of 0 to Contention Window (CW), with the initial CW of 32, doubled for every new attempt to transmit the same frame (up to the max value of 1024). To further reduce the

probability of collisions, especially when hidden terminals may be present in the network, Request To Send (RTS) and Clear To Send (CTS) frames may be exchanged between the source and destination stations prior to the transmission of a DATA frame, causing all stations that hear either the RTS or the CTS to defer their transmissions. The DCF is simple and performs well under light traffic loads, but does not allow for any service differentiation.

In the PCF mode of operation, the AP alternates between Contention Periods (CP) where the DCF access rules are used, and Contention Free Periods (CFP) where the AP explicitly allocates transmission opportunities to wireless clients by polling. The CFP is started by a broadcast of a special beacon frame from the AP; this forces all wireless stations to enter a mode where they can transmit only in response to a poll. The AP sends poll frames, possibly piggybacked on the downlink frames, to stations that have been placed on the polling list; these in turn respond with acknowledgement and uplink frames. At the end of the CFP (signalled by a CF-End frame), stations revert to their normal operation under the DCF rules (i.e. enter the Contention Period), until the next CFP. PCF offers a certain level of service differentiation (the AP can schedule transmissions to and from selected stations), but fails to provide the level of service differentiation control necessary to deliver QoS guaranteed service.

## 3.3      The 802.11e QoS extensions

Several limitations of the basic DCF and PCF in WLAN applications where QoS control is of primary concern, have led to the current work in IEEE (through its 802.11e working group) on QoS extensions to the 802.11 MAC layer. Although this work is still in progress, with the resulting documents enjoying at best a draft status, we will present here a summary of the new features outlined in the 802.11e Draft 3.0 [11].

The first enhancement offered by the 802.11e is the Enhanced DCF (EDCF). In the original 802.11 DCF, all clients enjoy equal access to the radio resource. The 802.11e EDCF offers up to 8 prioritised traffic categories (TCs), with TC7 having the highest priority. Frames to be transmitted can be marked with one of the TCs, and forwarded to a queue specific to that TC. Frames at the head of each of the eight queues contend for a transmission opportunity using the DCF rules, except that the lower the TC priority, the longer the station has to sense the medium idle before gaining its transmission opportunity (TXOP). Therefore, the 8 queues can be seen as 8 virtual stations contending for access to the radio medium. The different waiting times applied to the different queues ensure that the higher

priority frames are given the transmission opportunity more often than the lower priority ones.

The second major enhancement in 802.11e is the new contention-free access mechanism called Hybrid Coordination Function (HCF). Under the HCF the AP can, in addition to polling clients during the CFP, poll clients (allocate them a TXOP) during the CP. The AP can also allocate itself TXOPs. Upon granting the TXOP, the start time and the maximum duration of transmission are specified, i.e. once polled, a client can send multiple data frames within the same TXOP. To allow requests for TXOPs, the AP advertises a Controlled Contention Interval (CCI) in which clients can make resource requests (RRs) i.e. requests for TXOPs. Combined with EDCF, the HCF offers greatly improved QoS control at the wireless MAC protocol level.

In addition to differentiation and short-term medium reservation capabilities offered by the EDCF and HCF, the third 802.11e extension, namely the Traffic Specification (TSPEC) facility, offers flow-based traffic admission and reservation capabilities. When a client wants to reserve resources for a stream of frames (a flow), it sends a TSPEC (Action) frame with specification of the QoS requirements (e.g. data rate, delay) for the flow. The AP now decides to admit the TSPEC, admit a modified TSPEC, or reject the TSPEC, after which it returns the result to the client. The admitted TSPEC is given one of 8 TSPEC IDs (TSIDs) identifying different flows associated with the client in question, and all frames using that TSID are classified for treatment according to the TSPEC registered for this specific flow. The TSPEC facility offers a means for MAC level admission control and reservation signalling between the wireless clients and the Access Points.

# 4. HYBRID QOS ARCHITECTURE FOR AN IEEE 802.11 WLAN

As discussed in Section 2, IntServ principles of explicit admission control and reservation are followed locally in the radio access network. We assume that in the design of our hybrid QoS architecture, we will make use of features defined in 802.11e mentioned in Section 3.3.

We assume that traffic on the WLAN can be divided into two major classes: traffic generated by session-based applications, and traffic generated by non-session-based applications. Session-based applications normally start with session set-up procedure. The QoS parameters need to be guaranteed for the duration of the session and are either negotiated between the

application and network entities at the session set-up time, or implicit in the type of application. Examples include Voice over IP and video streaming sessions. Non-session-based traffic does not need hard QoS guarantees, thus explicit resource reservation is not necessary. The traditional "best effort" service is sufficient for this class of traffic.

At the local WLAN level, the QoS control involves two steps. The first one is stream admission control and resource reservation. The second is service differentiation.

The Admission Control functionality is located at the Radio Resource Manager and is responsible for admission control of session based application streams. The admission decisions are made on the basis of stream QoS requirements and the current RRM's knowledge of the resource usage (reservation) status in the WLAN. The admitted streams are then registered with the edge router for the purpose of mapping between the 802.11e stream QoS descriptors (TSPEC) and stream identifiers (TSID), the user priority levels on the Ethernet distribution network, and the DiffServ DSCPs visible at the edge of core/transport network.

UIIt The QoS signalling between the wireless station and the Access Point is accomplished by means of MAC level TSPEC negotiation defined in 802.11e. The signalling between the Access Point, Radio Resource Manager and the edge router is accomplished at the application level (via exchange of IP packets).
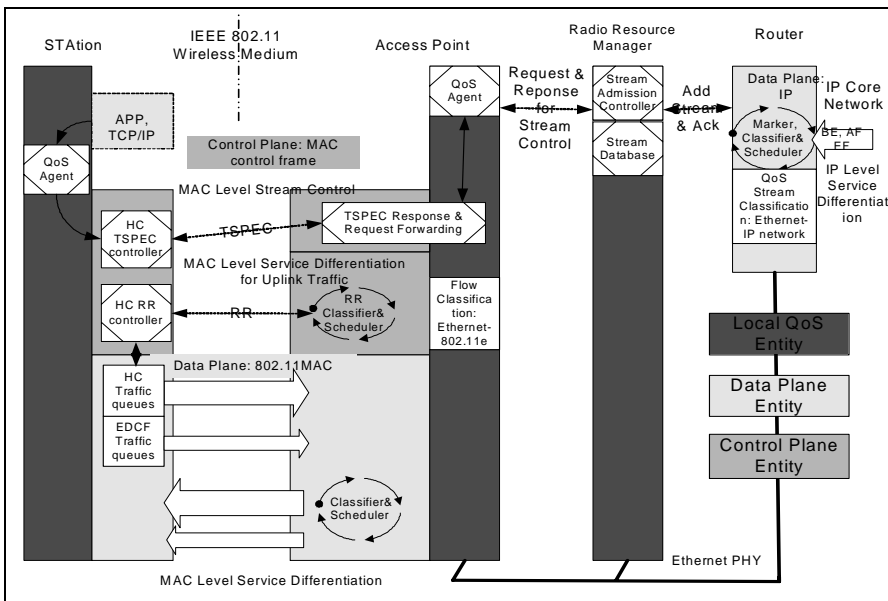


*Figure 3.* The QoS Architecture for 802.11 WLAN

The Service Differentiation performed at the 802.11e MAC layer level ensures that the high priority (session streams) frames have transmission opportunities (TXOP) satisfying their QoS requirements, as promised at the time of stream admission. Lower priority traffic is treated according to "best effort" principles, filling in the bandwidth available after the session-based streams admitted to the system have been satisfied. *Figure 3* explains the details of the QoS architecture described here.

## 4.1 Admission control procedures

When a wireless station (STA) initiates, or is invited to, a session-based application, a session set-up dialog is carried out (we may think of a SIP Invite dialog as an example). The QoS agent in the STA will capture the QoS requirements of media streams involved in the session, and map them to a MAC layer TSPEC description as defined in the 802.11e.
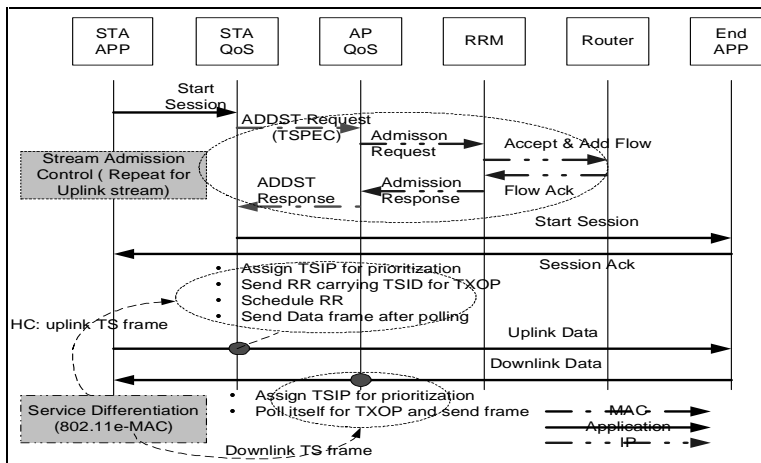


*Figure 4.* A Signalling Diagram for Stream Admission Procedures

In order to request admission and reserve radio resources for the stream, an 802.11e ADDTS-Request (add stream request) frame is sent to the AP. It carries a TSPEC element, which describes the source address (MAC), destination address, TSID, and QoS parameters of the stream. The QoS agent in the Access Point then forwards, in an IP packet, the admission request to the Admission Controller in the Radio Resource Manager. The RRM has "global" knowledge of the WLAN resources and reservation status; it will either admit or reject the stream, taking into account the

resource usage across the WLAN. If the stream is successfully admitted, the RRM registers the stream with the edge router (via IP level communication) and sends a positive reply to the AP's QoS manager. Subsequently, a QoS ADDTS-Response frame is sent back to the wireless station, carrying a TSPEC element for the admitted stream. The admitted TSPEC could be as requested, or altered as a result of resource negotiation at the RRM.

## 4.2      **Flow classification**

Once resources are reserved for a stream, application data frames must be classified in order for the service differentiation mechanism to be applied in the AP at the MAC layer level. The task of the classifier (for downlink traffic, in the edge router) is as follows: given an IP datagram of a particular flow (identified by, for example, the source/destination IP addresses and port numbers; we will refer to this as the flow ID), allocate the stream identifier (TSID) to the corresponding MAC data frame. For user data traffic on the uplink, the classification is straightforward because the QoS agent in the wireless station has knowledge of the streams generated by this station and admitted to the system. As a result, the TSID can be inserted directly into the MAC frame when it is generated at the station. For downlink traffic, classification is more complex because the AP operates only at the MAC layer level, and has no knowledge of the traffic flows at the IP level. The process of classification and mapping of IP flows onto the TSIDs must begin at a layer 3 device, i.e. the edge router.

We propose that, for downlink traffic, the edge router examines IP packets to detect flows and marks the distribution network (e.g. Ethernet) MAC frames with a priority level based on the TSID previously registered for the flow (recall that as part of admission control procedures, the RRM notifies the edge router of new flows). The priority information in the 802.3 MAC frame on the distribution network is carried in the additional 802.1p header (this additional header, which can be processed by most Ethernet products available today, carries a 3-bit user priority field). Therefore, when the router sends a distribution network MAC frame towards the AP, the frame contains the mobile host MAC address and the user priority value equal to the registered TSID. The AP's QoS agent must then interpret the Ethernet user priority field as the TSID for this frame. Together with the identity of the destination station, this determines the service differentiation treatment the frame will receive at the AP.

This method of classifying downlink traffic using the Ethernet user priority field relies on two key assumptions:

1. The Ethernet user priority field is on the distribution network used
   exclusively for signalling of TSID information, i.e. switches on the

distribution network should ignore the Ethernet user priority. We therefore strongly recommend that the WLAN distribution network be separated from other parts of the LAN by an edge router where classification of flows is performed.

2. Fixed hosts attached to the Ethernet (such as servers within the WLAN subnet) must be equipped with a QoS agent that ensures MAC frames sent by them are marked with the appropriate user priority that will be interpreted by the APs as TSID.

## 4.3      Service Differentiation in MAC

In our proposed 802.11 WLAN QoS architecture, service differentiation is performed at the MAC layer level according to the principles defined in 802.11e. It needs to be noted that the principles of the QoS architecture framework described in this paper will still hold for other choices of service differentiation mechanisms at the local radio access network level.

## 5.      CONCLUSION

We presented a hybrid QoS architecture framework for next generation wireless networks. The framework assumes simple and scalable DiffServ QoS control model for the core/transport part of the end-to-end path, but applies IntServ principles of explicit admission control and resource reservation locally in the wireless access network domain. We also presented in detail an example solution consistent with the chosen hybrid QoS architecture principles, for WLAN exploiting the 802.11e QoS extensions being currently standardised within the IEEE.

## 6.      REFERENCES

[1] Maniatis, S.; Grecas, C.; Venieris, I., "End-to-end quality of service issues over next generation mobile Internet", Communications and Vehicular Technology, 2000. Symposium on, P150 –154

[2] Dixit, S.; Yile Guo; Antoniou, Z., "Resource management and quality of service in third generation wireless networks", IEEE Communications Magazine, Volume: 39 - 2, Feb. 2001; P125 –133

[3] Koodli, R.; Puuskari, M., "Supporting packet-data QoS in next generation cellular networks", IEEE Communications Magazine, Volume: 39; Issue: 2, Feb. 2001; P180-188

[4] J. Antonio García-Macías; Franck Rousseau; Gilles Berger-Sabbatel; Leyla Toumi; Andrzej Duda, "Quality of service and mobility for the wireless internet", Proceedings of the first workshop on Wireless mobile internet July 2001, P34 - 42 .

[5] Robles, T.; Kadelka, A.; Velayos, H.; Lappetelainen, A.; Kassler, A.; Hui Li; Mandato, D.; Ojala, J.; Wegmann, B., "QoS support for an all IP system beyond 3G", IEEE Communications Magazine, Volume: 39 Issue: 8, Aug. 2001 P64 -72

[6] S. Mangold; S. Choi; P. May; O. Klein; G. Hiertz; L. Stibor, "IEEE 802.11e Wireless LAN for Quality of Service", Proceedings of the European Wireless, 2002, February, Volume 1, P32-39, Florence, Italy.

[7] G. Anastasi, L. Lenzini, "QoS provided by the IEEE 802.11 wireless LAN to advanced data applications: a simulation analysis", Wireless Networks, Volume 6, Issue 2  1999.

[8] Indu Mahadevan; K. M. Sivalingam, "Architecture and Experimental Framework for Supporting QoS in Wireless Networks Using Differentiated Services", in ACM/Baltzer Mobile Networks and Applications Journal, Vol. 6, No. 4, P385-395, 2001

[9] Suk-Un Yoon; Ji-Hoon Lee; Ki-Sun Lee; Chul-Hee Kang, "QoS support in mobile/wireless IP networks using differentiated services and fast handoff method", Wireless Communications and Networking Conference, 2000. WCNC. 2000 IEEE, Volume: 1, 2000, P266 -270 vol.1

[10] Indu Mahadevan; K. M. Sivalingam, "Architecture and Experimental Results for Quality of Service in Mobile Networks using RSVP and CBQ", in ACM/Baltzer Wireless Networks Journal, Vol.6, No.3, pp.221-234, 2000

[11] IEEE 802.11 Task E WG, Draft Supplement to standard for Telecommunications and Information Exchange Between Systems -LAN/MAN Specific Requirements - Part 11: Wireless Medium Access Control (MAC) and physical layer (PHY) specifications: Medium Access Control (MAC) Enhancements for Quality of Service (QoS), IEEE 802.11e/D3.0, Jul. 2002.

[12] Y. Bernet; P. Ford; R. Yavatkar; F. Baker; L. Zhang; M. Speer; R. Braden; B. Davie; J. Wroclawski; E. Felstaine, "A Framework for Integrated Services Operation over Diffserv Networks", Request for Comments (RFC) 2998, IETF, November 2000

[13] S. Blake; D. Black; M. Carlson; E. Davies; Z. Wang; W. Weiss, "An Architecture for Differentiated Service", Request for Comments (RFC) 2475, IETF, December 1998

[14] B. Braden, D. Clark, and S. Shenker, "Integrated Services in the Internet Architecture: an Overview", Request for Comments (RFC) 1633, IETF, June 1994.

[15] Conforto, P.; Tocci, C.; Losquadro, G.; Sheriff, R.E.; Chan, P.M.L.; Fun Hu, Y., "Ubiquitous Internet in an integrated satellite-terrestrial environment: the SUITED solution", IEEE Communications Magazine, Volume: 40 Issue: 1, Jan. 2002 P98 –107

[16] Kanghee Kim; Seokjoo Shin; Kiseon Kim, "A novel MAC scheme for prioritized services in IEEE 802.11a wireless LAN", ATM (ICATM 2001) and High Speed Intelligent Internet Symposium, 2001, Joint 4th IEEE International Conference on, 2001 P196 –199

[17] Banchs, A.; Perez, X.; Radimirsch, M.; Stuttgen, H.J. "Service differentiation extensions for elastic and real-time traffic in 802.11 wireless LAN", High Performance Switching and Routing, 2001 IEEE Workshop on , 2001 P245 –249

[18] Banchs, A.; Perez, X; "Providing throughput guarantees in IEEE 802.11 wireless LAN", Wireless Communications and Networking Conference, 2002. WCNC2002. Volume: 1, Mar. 17-21, P130 –138

[19] IEEE 802.11 WG, Standard for Telecommunications and Information Exchange Between Systems -LAN/MAN Specific Requirements - Part 11: Wireless Medium Access Control (MAC) and physical layer (PHY) specifications. IEEE 802.11 std, 1999.