2005 Asia-Pacific Conference on Communications, Perth, Western Australia, 3 - 5 October 2005.

A Quality of Service Architecture for IEEE 802.16 Standards

Hamed S. Alavi
Dept. of ECE
University of Tehran
Tehran, Iran
hamed.alavi@ece.ut.ac.ir

Mona Mojdeh
Dept. of ECE
University of Tehran
Tehran, Iran
m.mojdeh@ece.ut.ac.ir

Nasser Yazdani
Dept. of ECE
University of Tehran
Tehran, Iran
yazdani@ut.ac.ir

***Abstract*- A worldwide demand for a high-speed, always-on broadband wireless system across residential and business regions is emerging rapidly due to an increasing reliance on Internet for information, business and entertainment, as well as new bandwidth-intensive applications. The IEEE 802.16 Air Interface Standard is truly a state-of-the-art specification for fixed Broadband Wireless Access (BWA) Systems employing a point-to-multipoint (PMP) architecture.**

Although IEEE 802.16 standards define different mechanisms to provide Quality of Service requirements, it is the responsibility of the developers to obtain efficient designs and thus, providing QoS in BWA systems has become a challenging issue for designers of such systems.

In this paper, we introduce a novel architecture to support Quality of Service in IEEE 802.16 standards. Moreover, we propose a design approach to implement such architecture. Simulation result shows the high performance of our architecture for all types of traffic classes defined by the standard.

Keywords: Quality of Service, Broadband Wireless Access, IEEE 802.16

I. INTRODUCTION

Demands for high-speed Internet access and multimedia service for residential and business customers has increased for last mile broadband access. Broadband wireless access is emerging as a broadband access technology with several advantages; rapid deployment, high scalability, low maintenance and upgrade costs, and granular investment to match market growth [3]. IEEE 802.16 standards have been designed for broadband wireless access systems [1]. These standards utilize different mechanisms to provide Quality of Service requirements. While, these mechanisms are mentioned in standards, the details of the designs are directly left to developers. Therefore, an appropriate architecture to perform those mechanisms that provide QoS requirements seems to be essential. Most of the works to provide QoS in these systems are limited to specific modules such as Scheduling, Admission Control, and Traffic Shaping. The one proposed in [3] includes an upstream scheduler and a traffic shaper module to provide QoS for MAC protocol of IEEE 802.16 BWA systems. A scheduling architecture and the way it deals with each type of service flow are suggested in [4]. The one provided in [6] is an architecture in which downstream generation is described based on a scheduling algorithm. Another architecture defined by [2] proposes a design that considers admission control, schedulers and traffic policers with focus on an uplink packet scheduler. Unfortunately previous architectures do not support conformances regarding up-to-date standards.

In this paper, we propose an inclusive architecture to support QoS mechanisms in IEEE 802.16 standards. Also, we develop some compatible methods for specific modules such as Scheduler, Traffic Shaper, and Request and Grant Manager to optimize Delay, Throughput and Bandwidth Utilization metrics. Our simulation results show that our proposal meets these objectives.

The rest of this paper is organized as follows. Section II provides an overview of IEEE 802.16 standards. We then describe our proposed architecture in details in section III. Afterwards we explain our novel method for request and grant mechanism in section IV. Section V gives the performance evaluation of the proposed architecture through simulations. Finally we conclude and discuss future work in section VI.

II. IEEE 802.16 BROADBAND WIRELESS ACCESS

The IEEE 802.16 standards for fixed BWA systems support metropolitan area network architecture. It assumes a point-to multipoint topology with a Base and several Subscriber Stations. Base Station (BS) controls and manages the entire system and each Subscriber Station (SS) performs as an interface between end users and the Base Station (Figure 1). The IEEE 802.16 standards define a connection-oriented MAC protocol that supports multiple physical layer specifications. The physical layer air interface is optimized for bands from 10 to 66 GHz. The downlink channel on which data flow is directed from BS to SSs uses TDM scheme and the uplink channel in opposite direction applies TDMA scheme [7, 10].

0-7803-9132-2/05/$20.00

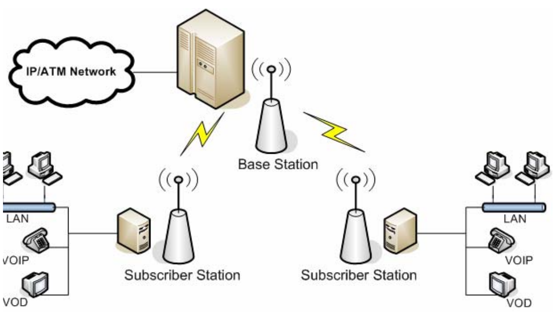


Figure 1. Broadband Wireless Access

IEEE 802.16 defines four types of service flows, each with different QoS requirement [1]:

Unsolicited Grant Service (UGS):

The UGS is designed to support real-time service flows that generate fixed size data packets on a periodic basis, such as T1,E1 and Voice over IP without silence suppression. The service offers fixed size grants on a real-time periodic basis, which assure that grants are available to meet the flow's real-time needs.

Real Time Polling Service (rtPS):

The rtPS is designed to support real-time service flows that generate variable size data packets on a periodic basis, such as moving pictures experts group (MPEG) video.

Non-Real Time Polling Service (nrtPS):

This service is for non-real-time flows which require better than best effort service, e.g. bandwidth intensive file transfer like FTP applications.

Best Effort Service (BE):

This service is for best effort traffic such as HTTP. There is no QoS guarantee.

IEEE 802.16 standards use specific request and grant mechanism in which each SS indicates the amount of uplink bandwidth it needs to the BS. The BS is allowed to allocate bandwidth in two modes; Grant Per Connection (GPC), in which bandwidth is assigned to each connection, and Grant Per Subscriber Station (GPSS), in which an SS requests for transmission opportunities for all of its connections and is allowed to re-distribute the bandwidth among them. The latter is more suitable when there exists many connections per terminal and it is mandatory for systems using the 10–66 GHz PHY specification [1].

III. Proposed Architecture

IEEE 802.16 standards have addressed a couple of mechanisms to provide Quality of Service such as Request and Grant, Polling and Bandwidth Allocation. The proposed architecture figured in 2 is designed based on these mechanisms. This figure shows data and signal transmission path since they enter the sender MAC layer until they leave receiver side.

Considering the fact that IEEE 802.16 standards are connection-oriented, each user first sends a *Connection Establishment Request* to BS. The request is then analyzed in *Call Admission Control* and if accepted, attributes of QoS and also two identifiers for each direction of this connection are registered in *Service Flow Data Bases*. In order to perform QoS process, packets are classified according to their mentioned identifiers in MAC entrance point by *Classifiers*.

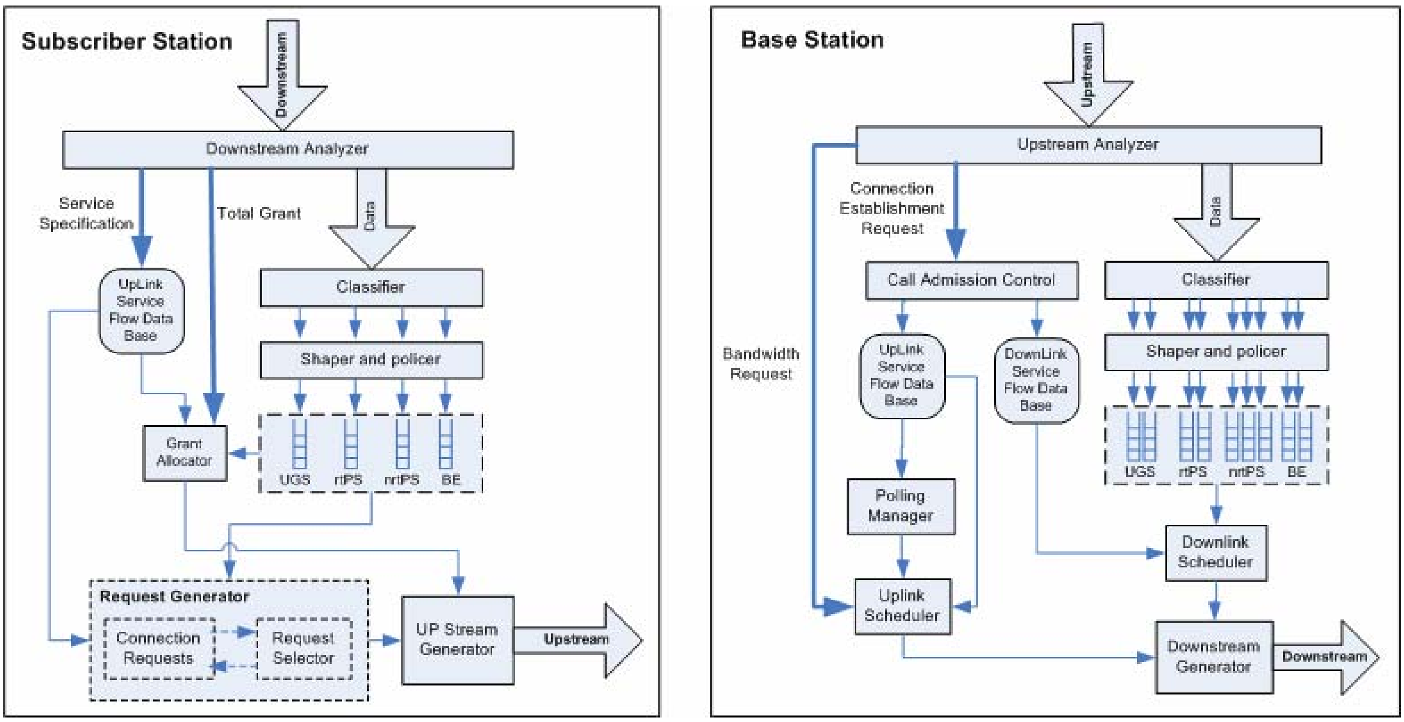


Figure 2. The proposed architecture to provide QoS in IEEE 802.16 standards

The transmission process of data from SS to BS is ordered as below:

- According to the Polling mechanism defined in the standards, BS is responsible to poll each SS in specific intervals, and consequently provide each of the connections with the opportunity to send their bandwidth requests [1]. This is done by *Polling Manager*.
- An SS that has obtained the opportunity of a bandwidth request through polling mechanism is then able to send one bandwidth request according to its QoS attributes and lengths of connections queues.
- *Uplink Scheduler* in BS divides the entire bandwidth of the uplink among active SSs according to the bandwidth requests it has received, QoS attributes of each connection of these SSs, and the outcome of *Polling Manager*. The result is then reported to each SS in the form of specific time slots.
- *Grant Allocator* in each SS divides the total bandwidth it has obtained, among its different connections according to their QoS characteristics.

The transmission process of data from BS to SS is ordered as below:

- *Downlink Scheduler* in BS distributes the entire downlink bandwidth among downlink connections.
- BS *Downstream Generator* sends specific amount of data from each downlink connection according to the output of *Downlink Scheduler*. This module is also responsible to send messages generated in *Uplink Scheduler*.

The traffic entering MAC layer is examined according to its attributes through *Traffic Shaper and Policer* in both SS and BS.

IV. REQUEST AND GRANT MECHANISM

Since the request and grant mechanism in IEEE 802.16 standards is so particular, and there is still no appropriate design for it, an efficient technique is proposed in this section to provide the mechanism. According to this mechanism, BS polls each SS in specific intervals. This can be done by the use of allocating extra bandwidth or sending a Polling message. On the other hand, each SSs is responsible to generate and send bandwidth requests in appropriate situations. Bandwidth requests may be incremental or aggregate. When the BS receives an incremental bandwidth request, it shall add the quantity of bandwidth requested to its current perception of the bandwidth needs of the connection. When the BS receives an aggregate bandwidth request, it shall replace its perception of the bandwidth needs of the connection with the quantity of bandwidth requested [1]. This should be done as follows:

- As shown in Figure 2, the *Connection Request* module in SS is responsible for generating one request for each connection in each frame, this is done according to the amount of data in their queues. The self-correcting nature of the Request and Grant protocol requires that SSs periodically use aggregate bandwidth Requests. The period may be a function of the QoS of a service and of the link quality. Except when there is an aggregate request to send, in all other cases connections would generate incremental request. The request for each connection is generated as below:

$$Inc_i = \sum_{i=LastAccepted}^{current} g_i + (q_i.len)_{current} - (q_i.len)_{LastAccepted}$$

$$Agg_i = (q_i.len)_{current}$$

in which Inc_i and Agg_i are the incremental and aggregate requests for connection i respectively, $(q_i.len)_{current}$ is the current queue length of the ith connection and $(q_i.len)_{LastAccepted}$ is the length of that queue when the last request of that connection was accepted by *Request Selector*. $\sum_{i=LastAccepted}^{current} g_i$ is the amount of grant allocated to connection i from the time the last request of the connection was accepted till the current time.

- *Request Selector* in each SS analyzes the requests generated in *Connection Request*, and considering their types and also their connection QoS attributes, selects and sends at most one of them. The idea used in this module is as follows. First each queue is assigned a weight according to its QoS parameters, and then in each round the selector chooses one request by multiplying the weight and length of the queues and selecting the greatest. The chosen connection is informed as well.

V. SIMULATION AND RESULTS

In order to evaluate and analyze the performance of our architecture, we have developed a simulation tool in a Java based environment. The goal of the experiment was to show that the proposed architecture can provide QoS support in terms of Delay and Bandwidth Utilization.

As depicted in figure 3, the model consists of a BS and four SSs with max uplink and downlink capacity of 80 Mbps. Each SS has four service flows with mentioned average loads.

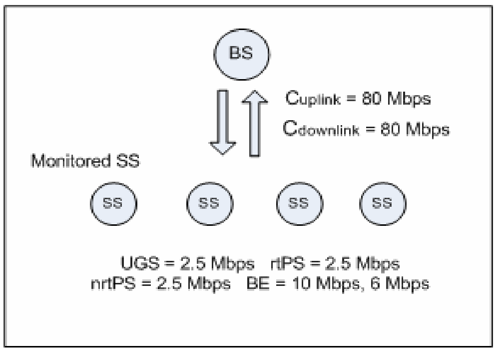
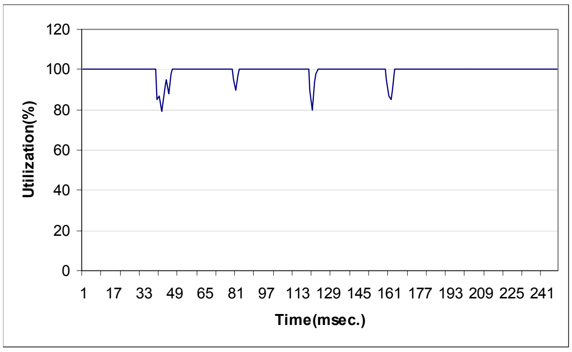


Figure 3. Simulated Model

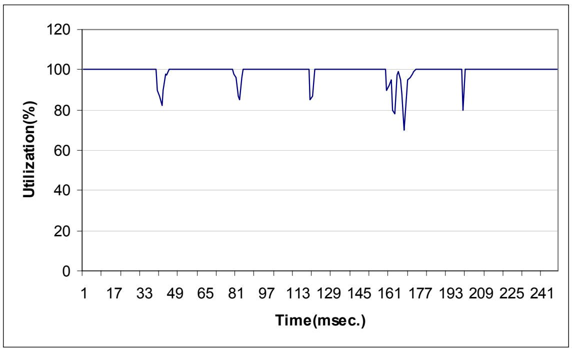
As an appropriate design to implement the proposed architecture, we applied combination of Token Bucket[11] and Leaky Bucket[12] algorithms to *Traffic Shaper and Policer*. For *Uplink* and *Downlink Schedulers* and *Grant Allocation*

Modules, we used Weighted Fair Queuing algorithm [8] and for *Request Generation* the method discussed in section 4 is utilized.

Results were evaluated using Average Delay and Bandwidth Utilization metrics. Average Delay is defined as average latency between the reception of a packet by the BS or SS on its network interface and the forwarding of the packet to its physical interface [1]. Bandwidth Utilization specifies the proportion of the total bandwidth granted to each SS that is utilized for data and message transmission. Figure 4 shows Bandwidth Utilization in two experiments with two different best effort traffic loads, 10Mbps and 6Mbps, in *Monitored SS*. In both, it is assumed that incremental and aggregate requests are generated in 10 and 40 msec periods respectively. In the first test, we achieved 8ms Average Delay, and in the second, the Average Delay was 6ms. While the bursty nature of best effort traffic makes bandwidth utilization more inefficient than other services, the result figured in 4 shows that our method always gains approximately full bandwidth utilization.



a) Load: 6Mbps



b) Load: 10Mbps

Figure 4. Best Effort Bandwidth Utilization

The achieved result is due to the characteristic of the method proposed for request generation. In the case of incremental request, the sum of requested bandwidth never exceeds the sum of arrived data in any interval. The missed bandwidth, as shown in figure 4, is mainly because of the aggregate requests, which happens periodically.

Since the Average Delay obtained is mainly related to BE and nrtPS, real-time services are delayed less than 1ms in our experiments that is adequate for common real-time applications [13].

VI. Conclusion

In this paper, we proposed a novel inclusive architecture to support Quality of Service mechanisms in IEEE 802.16 standards together with a design to implement such architecture. To the best of our knowledge, it seems that no similar work had been presented before. Moreover, the simulation results proved the performance of our architecture. We are currently in the process of providing a model for IEEE 802.16 standards Physical layer in order to analyze overall throughput of the system.

Refrences

[1] IEEE P802.16-REVd/D5-2004: “Air Interface for Fixed Broadband Wireless Access Systems”

[2] Wongthavarawat, K.; Ganz, A.; “IEEE 802.16 based last mile broadband wireless military networks with quality of service support”, *Military Communications Conference (MILCOM 2003)*, IEEE, Volume: 2 , Oct. 2003, PP. 779-784

[3] GuoSong Chu; Deng Wang; Shunliang Mei; ”A QoS architecture for the MAC protocol of IEEE 802.16 BWA system” *Communications, Circuits and Systems and West Sino Expositions, IEEE 2002 International Conference on* , Volume: 1, July 2002, PP. 435-39

[4] Hawa, M.; Petr, D. W.;”Quality of service scheduling in cable and broadband wireless access systems”, *Quality of Service, 2002. Tenth IEEE International Workshop on*, May 2002, PP. 247 – 255

[5] Vaughan-Nichols, S.J.; “Achieving wireless broadband with WiMax”, *IEEE Computer*, Volume: 37 , Issue: 6 , June 2004, PP. 10 – 13

[6] G. Nair; J. Chou; T. Madejski; K. Perycz; D. Putzolu; J. Sydir; “IEEE 802.16 Medium Access Control and Service Provisioning”, *Intel Technology Journa*l, Volume: 08, Issue: 03, August 2004, PP. 213-28

[7] Eklund, C.; Marks, R. B.; Stanwood, K. L.; & Wang, S. (2002), “IEEE Standard 802.16: A Technical Overview of the Wireless MAN Air Interface for Broadband Wireless Access”, *IEEE Communication Magazine*, June 2002, PP 98-107

[8] S. Lu, V. Bhargavan, and R. Srickant, “Fair Scheduling in Wireless Packet Networks”, *IEEE Trans. on Networking*, Vol. 7, No. 4, August 1999

[9] Y. Cao and V.O.K. Li, “Scheduling Algorithms in Broad-band Wireless Networks”, Proc. IEEE, Vol. 89, No.1, January 2001, pp 76-87

[10] Harry R. Anderson ; “Fixed Broadband Wireless Systems”, John Wiley & Sons, Ltd ISBN: 0-470-84438-8, 2000

[11] Puqi Perry Tang; Tai, T.-Y.C.; "Network traffic characterization using token bucket model" INFOCOM '99. *Eighteenth Annual Joint Conference of the IEEE Computer and Communications Societies*. Proceedings. IEEE , Volume: 1 , 21-25 March 1999 PP.51 - 62

[12] Wong, L.-N.; Schwartz, M.; "Access control in metropolitan area networks" *Communications, 1990. ICC 90, Including Supercomm Technical Sessions. SUPERCOMM/ICC '90. Conference Record., IEEE International Conference on*, 16-19 April 1990, PP.1591 - 1595

[13] Tsolakou, E.; Nikolouzou, E.; Venieris, L.;"A study of QoS performance for real time applications over a differentiated services network" *Computers and Communications, 2002. Proceedings. ISCC 2002. Seventh International Symposium on*, 1-4 July 2002, PP.397 - 403