Performance Analysis of QoS Scheduling in Broadband IEEE 802.16 Based Networks

V. Rangel[1], J. Gómez, J. Ortiz

Wireless Networks Research Lab
National Autonomous University of Mexico
Ed. Valdés Vallejo 3er piso,
Telecommunications Department
CU, 04510, México, D.F.
E-mail: victor@fi-b.unam.mx

Abstract

This paper presents the design and performance analysis of a scheduling technique for the provision of QoS over Broadband Wireless Access Networks (BWA). The proposed scheduling algorithm is based on the MAC protocol of the IEEE 802.16 standard and focuses on the uplink channel, which is the limiting factor of BWA networks and is critical in the delivery of services to individual users. Although the IEEE 802.16 standard had proposed several QoS service classes for various types of applications, they do not suggest how to schedule traffic to fulfill timing critical services such as compresed/un-compresed voice, audio and video streams. We have derived a mechanism called EBSA that combines several scheduling algorithms to closely match VBR-like and CBR-like traffic over the IEEE 802.16 air interface. Simulation results show that EBSA provides real-time services with very low access delays even during congestion periods.

Introduction

Broadband Wireless Access (BWA) has become the best way to meet residential and small business demand for high speed Internet access and multimedia services. As an emerging technology for broadband access, it provides the following advantages over its wired competitors: 1) rapid deployment and ease to implement, a BWA network can be installed rapidly without extensive underground cable infrastructure as in the case of Cable or DSL networks, 2) high scalability, carriers can expand the BWA network as subscribers demand for bandwidth grows by adding channels, or cells, 3) lower maintenance and upgrade costs and 4) higher data rates. However, the wide-scale adoption of BWA systems will be determined by its ability to overcome cost and performance barriers. If BWA can meet these challenges it could easily be the next revolution in wireless networks systems such as WLANs.

As for the performance is concerned, BWA systems will provide services for diverse traffic classes, with different Quality of Service (QoS) requirements. Although the IEEE 802.16 MAC protocol includes QoS guarantees, it does not provide a complete solution and does not tell how to schedule traffic to fulfill QoS requirements specifically. In recent years, several scheduling algorithms for BWA networks were published [1-4]. In this paper we explore the combination of various scheduling mechanisms in order to closely match QoS service classes defined in IEEE 802.16 over its air interface. To the best of our knowledge, there is no proposed packet scheduling solution specifically targeted for IEEE 802.16. In this paper we present a scheduling algorithm that supports diverse traffic classes, such as Constant Bit Rate (CBR) and Variable Bit Rate (VBR) with different QoS requirements. The scheduling algorithm combines Prioritization, Early Deadline First [5], Round Robin [6], and Weighted Fair Queueing [7] strategies to closely match realtime and non-realtime traffic services, such as Voice over IP (VoIP), multimedia services and high speed Internet access. We call this scheduling mechanism the **E**DF-**B**WA **S**cheduling **A**lgorithm (***EBSA***).

The rest of the paper is structured as follows, In section 2, we provide an overview of the IEEE 802.16 MAC protocol and provide detailed information of the four different types of QoS agreements supported by the standard. In Section 3, we describe the proposed scheduling algorithm. In Section 4, we present a performance analysis of *EBSA* for different traffic scenarios. Finally in Section 5, we present a conclusion and future work.

IEEE 802.16 Standard Overview

The first version of the IEEE 802.16 standard (also termed as the WirelessMAN^{TM} air interface standard) addressed line-of-sight environments at high frequencies bands from 10-66 GHz [8] and data rates from 32 to 130 Mbps depending of the channel bandwidth and modulation technique (64-QAM, 16-QAM or QPSK). The recently adopted amendment, the IEEE 802.16a standard [9] is designed for non-line-of-sight environments at lower frequencies bands operating in the 2-11 GHz range. Currently the IEEE 802.16 Working Group is defining a mobile amendment, the IEEE 802.16e. The present paper focuses on the first version of the IEEE 802.16 standard.

IEEE MAC Protocol

The IEEE 802.16 MAC protocol regulates uplink (UL) channel access using Time Division Multiple Access (TDMA). Upon entering the BWA network, each Subscriber Station (SS) has to go throughout the Initialization and Registration setup, described as follows:

[1] The research presented in this paper is founded by the National University of Mexico. Project: UNAM IN110805.

Upon power up, subscriber stations need to synchronize with a downlink channel (DL-Ch) and an uplink channel (UL-ch). When a SS has tuned to a DL-ch, it gets the frame structure of the UL-ch, called a UL-MAP frame, then the ranging procedure is performed, where the round-trip delay and power calibration are determined for each SS, so that SS transmissions are aligned to the correct mini-slot boundary. The next step is to establish IP connectivity, the Base Station (BS) uses the DHCP mechanisms in order to obtain an IP address for the SS and any other parameters needed to establish IP connectivity. Then, the SS establishes the time of the day, which is required for time-stamping logged events and key management. In the next step, the SS establishes a security association and transfers control parameters via TFTP, these parameters determine the BS and SS capabilities, such as QoS parameters, fragmentation, packing, among others. Finally, the registration process is performed; the SS must be authorized to forward traffic into the network once it is initialized, authenticated and configured.

Once the Initialization and Registration setup is complete, a SS can create one or more connections over which their data are transmitted to and from the BS. Subscriber stations request for transmission opportunities on the UL channel. The BS gather these requests and determines the number of time slots (grant size) that each SS will be allowed to transmit in the UL-Frame. This information is broadcasted in the DL channel by the BS using the UL-MAP message at the beginning of each DL-Frame, as illustrated in Figure 1. The UL-MAP contains Information Elements (IE) which describe the transmission opportunities in the UL channel, such as initial maintenance, station maintenance, contention and reservation access. After receiving a UL-MAP, a SS will transmit data in the predefined transmission indicated in the IE, these transmission opportunities are assigned by the BS using the following QoS service agreements.

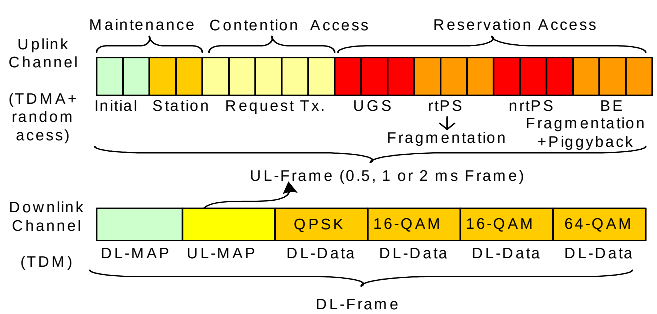

Figure. 1. UL and DL channel structure.

QoS in IEEE 802.16

The IEEE 802.16 defines four different QoS types, which are based on those defined in the DOCSIS v.1.1 standard [10]. These classes of services are described below.

Unsolicited Grant Service (***UGS)***: This service is oriented for the support of real-time service flows that generate fixed- size data packets on a periodic basis (CBR-like services), such as T1/E1, VoIP or videoconference. At the beginning of the connection setup, a SS indicates the BS about its requirements for this service, such as grant size (*G*), grant interarrival time (λ), tolerated grant jitter (*j*) and *Poll* bit. The UGS service also includes Activity Detection (AD) to examine the flow state. If the state is inactive, then the UGS-AD Service sets the *Poll* bit to 1 and provides periodically a *unicast transmission opportunity* (*utxop*), in which a SS can indicate the BS to reestablish its UGS service, thus saving bandwidth.

Real-Time Polling Service (***rtPS***): This service is oriented for the support of real-time service flows that generate variable size data packets on a periodic basis (VBR-like services), such as MPEG video streams. The *rtPS* service offers periodic *utxop*, which meet the flow's real-time needs and allow the SS to specify the size of the desired channel reservation. A SS should indicate to the BS at the beginning of the session about its requirements for this service, such as polling interval (λ) and tolerated poll jitter (*j*).

Non Real-Time Polling Service (***nrtPS***): This type of service is like the *rtPS*, however polling will typically occur at a much lower rate and may not necessarily be periodic. This applies to applications that have no requirement for a real time service but may need an assured high level of bandwidth. An example of this may be a bulk data transfer (via FTP) or an Internet gaming application. The parameters required for this service are the polling interval (λ), minimum and maximum sustained data rate.

Best Effort (***BE***): This kind of service is for standard Internet traffic, where no throughput or delay guarantees are provided.

The IEEE 802.16 MAC protocol can identify the type of service flow required by a SS using the following fields of the IEEE 802.16 protocol stack: source or destination MAC address, EtherType, source and destination IP address or network, IP protocol type, source or destination port number, IP type of service bits and any combination thereof. A simple example of how a classification might be used would be to match VoIP traffic from a particular source IP address and UDP port and to direct that traffic into a dynamically created service flow that had a QoS parameter set providing a UGS mode of data transmission.

Once the service flows have been identified, the BS uses two modes of operation to allocate grants: 1) Grants per Connection (GPC) and Grants per Subscriber Station (GPSS). In the first case, the BS grants bandwidth explicitly to each connection, whereas in the second case the bandwidth is granted to all the connections belonging to the SS. The latter case (GPSS) allows smaller uplink maps and allows more intelligent SSs to make last moment decisions and perhaps utilize the bandwidth differently than it was originally granted by the BS. This may be useful for real-time applications that require a faster response from the system.

EDF-BWA Scheduling Algorithm (EBSA)

The *UGS*, *rtPS* and *nrtPS* have specific requirements, for instance, *UGS* and *rtPS* have a deadline, (jitter), and late packets that miss the deadline will be useless, but these two services can tolerate packet loss. However, for *nrtPS* packet loss is not permitted but accommodates larger delays. In order to guarantee these three types of service with their specific requirements, we have implemented a scheduling algorithm to match CBR-like and VBR-like traffic. For all types of service, when a request for bandwidth is received at the BS, the QoS policy of the SS is first analyzed to make sure it is not violating the QoS contract (e.g. maximum bandwidth requirement). For all types of request, EBSA will provide to the SS a *UGS*, *rtPS*, *nrtPS* or *BE* service using classifiers as defined in [8].

If the request is for a *UGS* service, EBSA will provide periodic grants to the SS. These grants are allocated in a *UGS* queue and ordered using the EDF principle. The tolerated grant jitter is taken as the ordering parameter. If the request is for a *rtPS* or n*rtPS* service, EBSA will provide periodic *utxop* to the SS, as defined in section II. These *utxop* are allocated in an *rtPS* or *nrtPS* queue, respectively, and ordered using EDF. For these two services the tolerated poll jitter is taken as the ordering parameter. For *BE* requests, the grants are ordered using a FIFO scheme as illustrated in Figure 2.

Finally, EBSA will dispatch the grants or *utxop* using a WFQ scheme where the rations of reservation for each type of service (W_{ugs}, W_{rtp} , W_{nrtps}, and W_{be}) could be configured by the network operator or dynamically assigned by EBSA using a reservation ratio calculator, based on the current network traffic. We base our study, using fixed ratios.

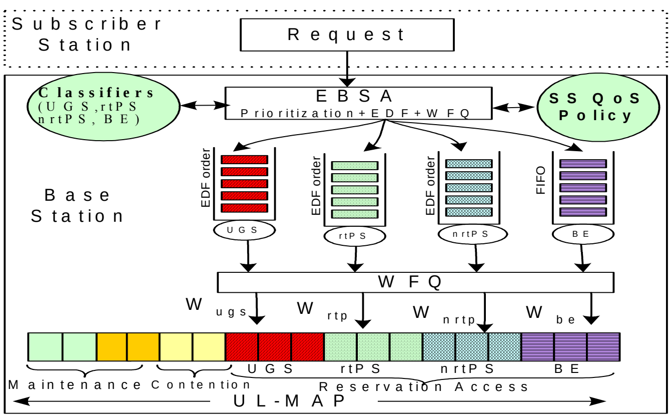

Figure 2. UL and DL channel structure.

In general, EBSA is easy to implement and is compatible with IEEE QoS requirements. Bellow we present 6 steps to implement EBSA.

1) We have identified the SS requirements by the following vector that is registered at connection setup or updated as the SS asks for bandwidth.

$$SS_n = \{QoS_i, s, l, j, Poll, p, G\} \quad (1)$$

where '***QoS**$_i$*' is the service type required by a SS defined by

$$QoS_i = \{0 \text{ for } UGS, 1 \text{ for } rtPS, 2 \text{ for } nrtPS\} \quad (2)$$

'l' is the interarrival time or polling interval measured in ***UL-Frames***, (e.g. if interarrival time or polling interval = 10 ms and *UL-Frame* = 2 ms then *l = 10ms/2ms = 5 UL-Frames), **'j'*** is the jitter measured also in *UL-Frames, **'p'*** is the user transmission (tx) priority, '***G***' is the grant size and '***s***' is the state of the connection. For every UL-MAP sent, the BS increases the user state (SSs_n) by 1. When SSs_n becomes equals to the interarrival time (SSs_n=SSl_n), it is time for the BS to send a grant or unicast transmission opportunity (*utxop*) to user SS_n depending on the '***Poll'*** bit .

2) Get the maximum number of priorities 'Max_p_i' that will be used for service type QoS_i as follows:

$$Max\_p_i = \frac{\max(SSj_1\_QoS_i, SSj_2\_QoS_i, ..., SSj_n\_QoS_i)}{UL\_Frame} \quad (3)$$

where $SSj_n\_QoS_i$ refers to the jitter of user *SSn* requiring QoS_i.

3) For each SS_n with QoS_i get its tx priority '$SS\boldsymbol{p}_n\_QoS_i$' using expression (4).

$$\begin{array}{ll}
\text{if } SSs_n_QoS_i - SSl_n_QoS_i = 0 \text{ and Poll} = 1 & (\text{a utxop is required}) \\
\quad SSp_n_QoS_i = Max_p_i & (\text{set SSp to highest tx priority}) \\
\text{else if } SSs_n_QoS_i - SSl_n_QoS_i > 0 & (\text{a grant is required}) \\
\quad (\text{get tx priority based on current state and tolerated delay}) & (4) \\
\quad SSp_n_QoS_i = SSs_n_QoS_i - SSj_n_QoS_i - SSl_n_QoS_i + Max_p_i & \\
\textit{else} \quad (\text{a grant is not required in the current UL - Frame}) & \\
\quad SSp_n_QoS_i = 0 & (\text{set tx priority to } 0)
\end{array}$$

4) Schedule first any maintenance regions as indicated in [8], then schedule the minislots that will be used in contention access (for bandwidth request transmissions), using a Contention Slot Allocator (CSA). We have presented a good approximation for this in [11]. If there are some SS_n requiring a *utxop* in the current UL-Frame, schedule these slots first. Users with a *UGS* service and the *Poll* bit set to 1 will use the *utxop* to indicate to the BS to reestablish its service. Users with a *rtPS* or *nrtPS* service always have the *Poll* bit set to 1, and these subscriber stations use the *utxop* for bandwidth requirements.

5) In the remaining space of the UL-Frame, schedule the SS_n according to its tx priority, in the following order *UGS*, *rtPS, nrtPS, BE*. In order to avoid scheduling unfairness, use the weighted rations (W_{ugs}, W_{rtp}, W_{nrtps}, and W_{be} respectively) for bandwidth allocation. A high priority have precedence upon a low priority, where priority ($SSp_n\_QoS_i$)= 1 is the lowest and priority = Max_p_i is the highest. If there are two or more SS_n with the same priority, use the technique of Round Robin to grant these users. For *rtPS* services, if the available space in the current UL-Frame < G then the BS should use the fragmentation technique and send continuous grant opportunities until G is complete. For *nrtPS* services, if the available space in the current UL-Frame < G then the BS should use the fragmentation technique and users may request for further bandwidth using piggyback requests. For every grant allocated in the UL-MAP for SS, update its state as follows.

$$SSs_n\_QoS_i = SSs_n\_QoS_i - SSl_n\_QoS_i \quad (5)$$

6) Finally, schedule any *BE* request using fragmentation + piggyback for a better utilization of the UL channel.

In the following section we demonstrate that our scheduling algorithm is well suited for the support of QoS providing very low access delays for VBR-like and CBR-like traffic for IEEE 802.16 based networks.

OPNET Model Description

This section shows the key functional breakdown of the OPNET model implementation. The model can be partitioned into two major parts: 1) Subscriber Stations and 2) a Base Station, as shown in Figure 3.

Subscriber Station

The Subscriber Station access node consists of a traffic generator, a Media Access Control unit and two RF modules for transmission and reception as pointed out in Figure 4.

The traffic generator module produces three different traffic types according to the QoS agreed with the BS. These traffic types are described in the following section.

The RF modules (ant_rx, from_link_rx, to_link_tx, and ant_tx) are responsible for accepting packets from or transmitting packes to the radio access network according to the propagation model described in [12].

The MAC module is responsible for processing packets of higher layers and transmitting packets to the radio access network according to its QoS level. The MAC module uses a primary Finite State Machine (FSM) and a secondary FSM. The primary FSM (Figure 5) is responsible of the initialization and registration procedure as well as the reception and processing of synchronization packets, UCD, UL-MAP and DL-MAP frames from the BS. All packets received from higher layers, are processed at the application traffic process and sent to the secondary FSM for transmission on the UL channel, as shown in Figure 6.

Upon receiving a traffic packet from the Primary FSM, the secondary FSM process this packet according to its QoS level. If the packet is for a Best effort service (e.i. Internet packet), the transmit opportunity state (Tx_Opp_Proc) looks for a contention opportunity (either in the current or in the following UL-MAP frame) and transmits a reservation request to the BS. In case this request results in collision with other contention transmissions, the collision resolution (Collition_Res) state takes care to resolve it according to the exponential backoff algorithm. If the packet demands a UGS or rtPS service (i.e. for voice or video packets, respectively) the transmit opportunity state sends a Dynamic Service Addition (DSA) request to the based station, indicating its type of service needed for this connection. If this request is accepted, the No_Request_Outstanding process takes care of receiving the corresponding grants and to indicate to the Tx_Opp_Proc when to transmit these voice or video packets.

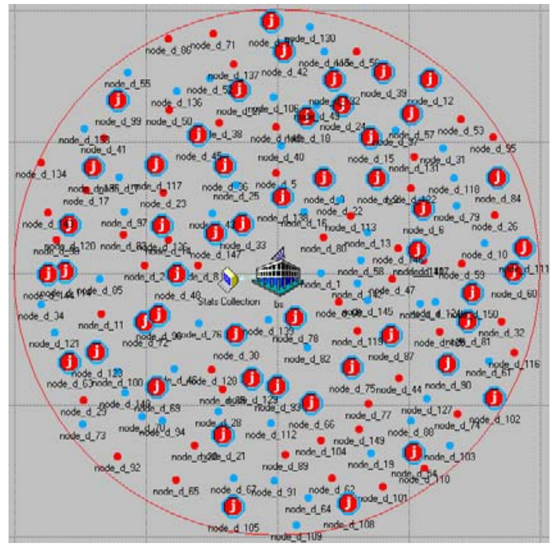

Figure 3. Network Model.

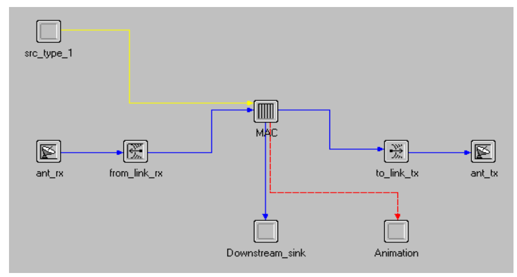

Figure 4. SS access node model.

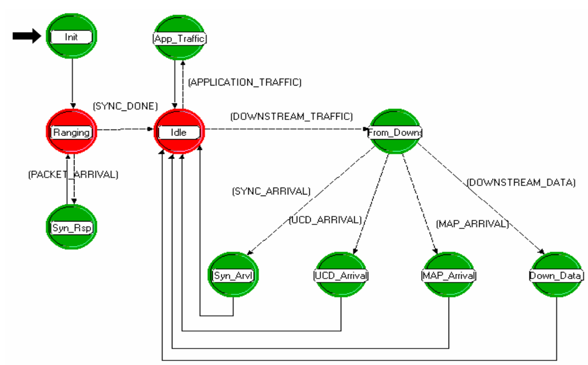

Figure 5. Primary SS´s FSM.

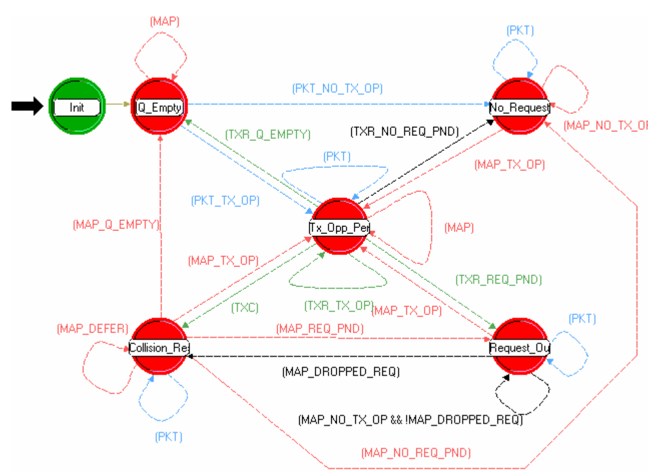

Figure 6. Secondary SS´s FSM.

For nrtPS packets (e.g. for FTP traffic), the same procedure is carried out as in the previous case, with the exception that the Request_Outstanding process is responsible for receiving the grants from the BS. But if no-grants are allocated for this service, these packet can still be transmitted in the radio access network using a Best Effort service.

Base Station

The BS is in charge of SS´s identification and to provide to SS with a QoS level. It is used for point to multipoint links and is the main gate for incoming and outgoing packets. Figure 7 shows the BS node model, which is composed mainly of a MAC unit, three reception modules (for QPSK/from_link_rx, 16-QAM/ from_link_rx or 64-QAM/from_link_rx_1 modulation) and a transmission module (to_link_tx).

The BS can support several channels with different data rates according to the modulation type negotiated. For example using a bandwidth of 6 MHz for UL-Channels, a data rate of 9.6, 19.2 or 28.8 Mbps can be obtained with QPSK, 16-QAM or 64-QAM modulation, respectively.

The MAC module is responsible of providing SS´s with the right QoS level, and guarantees the correct transmission opportunities. In order to provide theses transmission opportunities, the MAC module uses also two FSMs.

Basically, the primary FSM, illustrated in Figure 8, performs the following three functions: 1) takes care of the initialization and registration procedure, which is done by the ranging, rng_rcvd and rng_complete states. 2) Based on SS´s request, the Primary FSM creates the signaling MAPs, which describe the maintenance region (using the Mtn_MAP state), as well as contention and reservation access (using the MAP-Time state). 3) Provides with synchronizations and UCD information to SSs, as show in Figure 8.

All frames produced in the primary FSM, are passed to the secondary FSM, which takes care of transmitting theses frames in the correct DL channels as shown in Figure 9.

Performance Analysis

In all simulations, one UL channel with a capacity of 9.6 Mbps and one DL channel with a capacity of 22 Mbps were used. For the performance analysis, we have considered the simulations parameters given in Table I. Two novel traffic scenarios were considered for the performance of the proposed scheduling algorithm. These scenarios are based on the following traffic sources.

Traffic scenarios

1) ***VoIP- G.723-UGS***. This traffic type emulates a speech codec "G.723.1", which according to the ITU, IETF and the VoIP Forum is the preferred speech codec for Internet telephony applications. This codec generates a data rate of 5.3 kbps or 6.3 kbps depending on the mode, where 20-byte data packets are generated and encoded every 30 ms. By adding the complete set

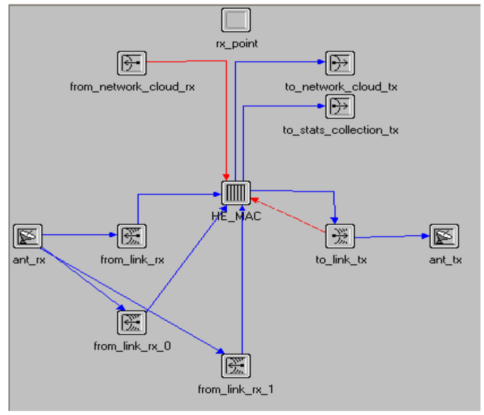

Figure 7. BS node model.

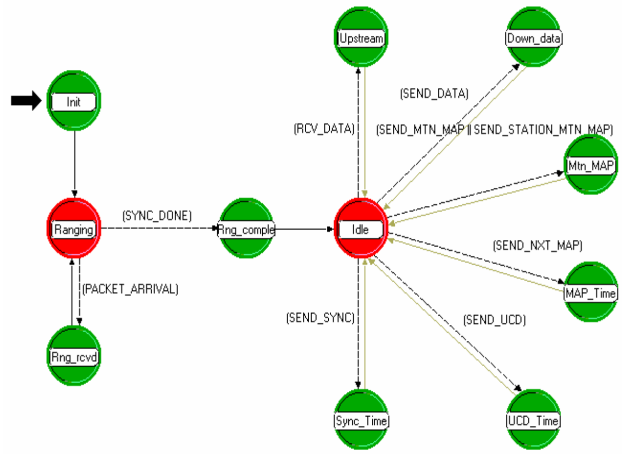

Figure 8. Primary BS´s FSM.

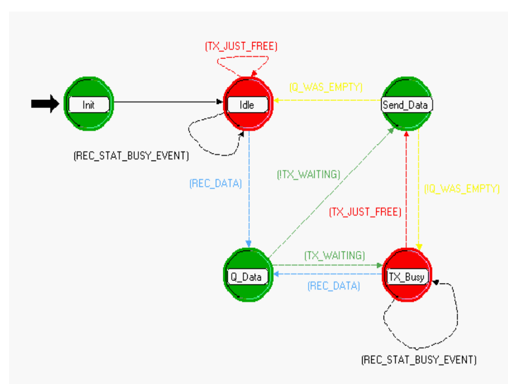

Figure 9. Secondary BS´s FSM.

Table I. Simulation Parameters

Parameter	Value
UL data rate (QPSK)	2.816 Mbps
DL data rate (16-QAM)	22.4 Mbps
Minimum contention slots per UL-Frame	8 slots
UL minislot size	16 bytes
UL-Frame Duration	2 ms =44 minislots
Simulation time for each run	60s
Distance from nearest/farthest SS to the BS	0.1 – 2..3 km
Reed Solomon (short grants/long grants)	6 bytes/ 10 bytes
Limit between short and long grants	245 bytes
Maximum number of users in the network	200

of headers as illustrated in Table II, one obtains a VoIP packet which demands 9 minislots in the UL-channel, obtained as follows: 20-byte voice frame + 12-byte RTP + 8-byte UDP + 20-byte IP + 3-byte LLC + 5-byte SNAP + 18-byte Ethernet MAC (needed when the 802.16 SS works as an external bridge) + 6-byte 802.16 MAC + 24-byte FEC (this is the Reed Solomon codification which includes 6 bytes of codification for each 30-bytes of data) + 4-byte padding (needed to make the last codeword of 12 bytes) + 10-byte Preamble and Guard band = 130 bytes ≈ 9 minislots). Thus at the PHY layer, this results in a VoIP stream of (9slots*16bytes*8bits/30ms Frame size) = 38.4 kbps.

2) ***Vo IP- G.711***. Codec G.711 was considered to stress the BWA network and also because this codec will be used for quality voice calls. G.711 is the mandatory codec according to the ITU-T H.323 conferencing standard, which uses Pulse Code Modulation (PCM) to produce a data rate of 64 kbps. This audio codec creates and encapsulates a 80-byte VoIP frame every 10 ms, and demands a stream of 166.4 kbps at the PHY layer

3) ***MPEG4-rtPS***. For this type of service we used the traces of 10 MPEG4 movies as defined in [13]. The traces were digitalized using QCIF (Quarter Common Intermediate Format) with 176*144 pls, at 25 fps (frames per second). Table III shows the selected MPEG-4 movies. The MPEG-4 movie to transmit is selected by the SS in a random fashion.

4) ***Internet Traffic-IP:*** The Internet traffic distribution utilized is the one introduced by the IEEE 802.14 working group [11]. The message size distribution is as follows: 64-byte Pk. 60%, 128-byte Pk. 6%, 256-byte Pk. 4%, 512-byte Pk. 2%, 1024-byte Pk. 25% and 1518-byte Pk. 3%. The inter-arrival times are set in such a way that the Internet offered load per active station is 38.4 kbps at the PHY layer.

The first analysis includes Internet and voice traffic, since it is considered that most of the traffic transmitted on residential zones would be of this type and for the second analysis we studied a mixed traffic configuration to stress the network, including voice, video and Internet traffic.

Performance of EBSA for Voice and Internet Traffic

In this scenario, the network was configured so that one third of the network population was transmitting VoIP-G723 traffic (eg: SS 1, 4, 7, etc), the second third VoIP-G711 traffic (SS 2, 5, 8, etc), and the last third Internet traffic (SS 3, 6, 9, etc). We used W_{ugs} = 1 and W_{be} = 0, thus all *BE* traffic will be scheduled with the lowest priority. The performance analysis of *EBSA* will be provided in terms of throughput, access delays and cumulative probability.

Figure 10 presents the UL channel utilization and the throughput achieved by each service flow. In the Figure we can appreciate that the maximum system utilization is of 9.2 Mbps, which corresponds to 95 % of the UL channel capacity (of 9.6 Mbps), the other 5% was assigned to contention access. From this utilization, only 66 % (6.3 Mbps) was used for data

Table II. VoIP Codecs: G.711 and G.723.1.

	G.711 - 64 kbps	G.723.- 5.3 kbps
Frame size [ms]	10	30
Voice frame [bytes]	80	20
RTP [bytes]	12	12
UDP [bytes]	8	8
IP [bytes]	20	20
LLC [bytes]	3	3
SNAP [Bytes]	5	5
Ethernet MAC [bytes]	18	18
IEEE 802.16 MAC	6	6
PHY: (Prea+GB+FEC)	10+FEC	10+FEC
Total PacketSize	202bytes or G=13 slots	86 bytes or G=9 slots
Net rate at MAC / PHY	116.6 / 166.4 kbps	22.9 / 38.4 kbps

Prea = Preable, GB = Guarband, and FEC = 6* No_CodeWords

Table III. MPEG-4 Movies.

Number	Movie Name	Mean video Frame (bytes)	Grant Size (Slots)	Video frame Rate (kbps)
1	Aladdin	297.61	25	80
2	Die Hard III	587.06	44	14.08
3	Futurama	1106.30	70	224
4	Jurassic	684.74	50	160
5	Mr. Bean	437.91	34	108.8
6	Robin_Hood	460.18	36	115.2
7	Silence	1871.20	130	416
8	Start Trek	209.22	19	60.8
9	Starwars	530.59	40	128
0	The Simpsons	1464.60	103	329.6

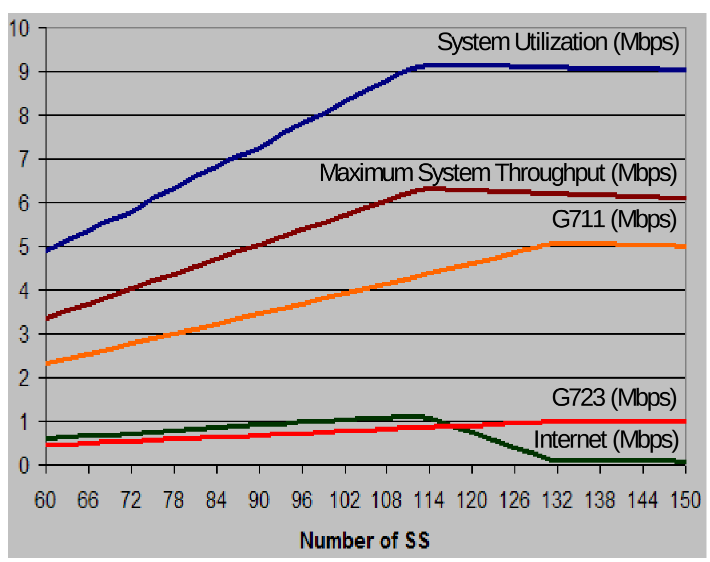

Figure 10. Throughput for all traffic types (60-150 SS).

transmission, and corresponds to the maximum system throughput. The rest (29%) was consumed by the Radio Link Protocol (RLP) of the IEEE 802.16, which consists of MAC and PHY headers. The maximum number of supported SS in a 6-MHz UL channel can be estimated by using a simple expression, as follows:

$$MaxSS = 3 * \frac{Utilization}{VoIP_{G711} + VoIP_{G723} + Internet} = 3 * \frac{9150}{166.4 + 38.4 + 38.4} \approx 112 \quad (6)$$

In Figure 10 we can appreciate that throughput for Internet traffic begins to decrease when the maximum system throughput is achieved with 112 users in the BWA network (37-G711 streams, 38-G723 streams and 37 Internet users). This is to be expected since EBSA always schedules UGS traffic with a higher ratio due to the Wugs factor set to 1.

This also results in lower access delays for UGS service flows as appreciated in Figure 11. In Figure 11a we can observe that 100% of VoIP frames were transmitted under 3.5 ms for G711 streams and 7.5 ms for G723 streams (Figure 11b). However, only 17% of Internet frames had access delays under 10 ms (Figure 11c). The direct consequence of having the ratio factor of BE traffic set to 0 is that on high congestions periods, Internet traffic gets a decreased scheduling priority in order to guarantee very low transmission delays for UGS. However this unfairness can be controlled by the network operator.

From (6), the maximum VoIP streams (without Internet traffic) that a UL channel can support is 44-G.711 and 44-G.723 streams. Figure 12 shows that for VoIP streams, 100% of frames transmitted on the BWA network had access delays under 5ms for G.711 streams (Figure 12a) and 10ms for G.711 streams (Figure 12b). All VoIP were transmitted without packet loss. Figure 12c shows the performance when the BWA network is over-loaded with 45-G.711 streams, 45-G.723 streams and 45 Internet users. Here, simulations results reported a packet loss of 3.5 %, due to late packets, which is still acceptable for the support of VoIP streams.

Performance of EBSA for Voice, Video and Internet Traffic

In this traffic scenario we study how the proposed algorithm scales with an increased number of traffic sources producing voice, video and Internet traffic. The network was configured in groups of 6 SS, in the following manner: 1-G711 stream (SS 1, 7, 13 etc), 1-G723 stream (SS 2, 8, 14, etc), 1 MPEG-4 stream (SS 3, 9, 15, etc) and 3 Internet users (SS 4-6, 10-12, etc). For this traffic scenario the weighted ratios were configured as follows: W_{ugs} =0.60 , W_{rtp} = 0.38, W_{nrtps}, and W_{be} = 0.02.

In Figure 13 we can appreciate that Internet throughput starts to decrease when the network is saturated with 112 SS. At this point, the proportion of Internet traffic transmitted was of 23% (55-SS*38.4kbps/9152kbps) of the bandwidth for reservation access of 9.152Mbps (=143*9.6Mbps/150). This proportion was much higher than W_{be} = 2%, this is because EBSA allocated un-scheduled traffic of *UGS* and *rtPS* to *BE* traffic. However on

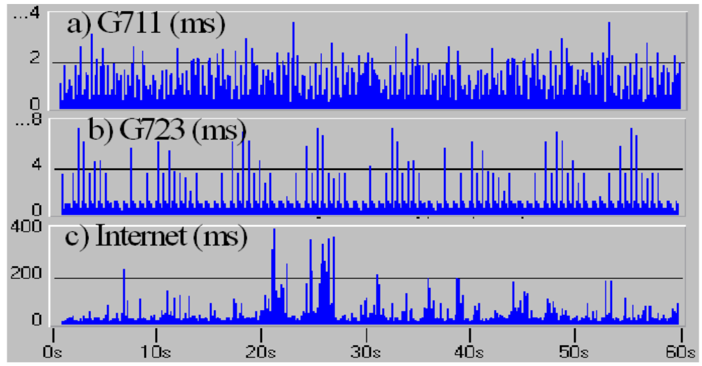

Figure 11. Access delay vs time (112 SS).

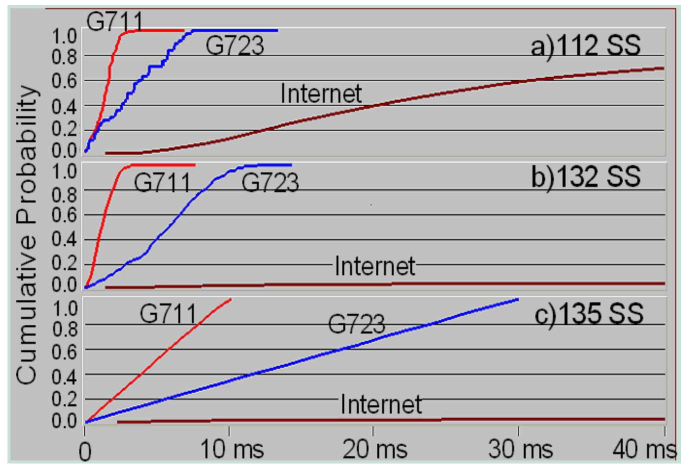

Figure 12. Cumulative Probability vs Access delay.

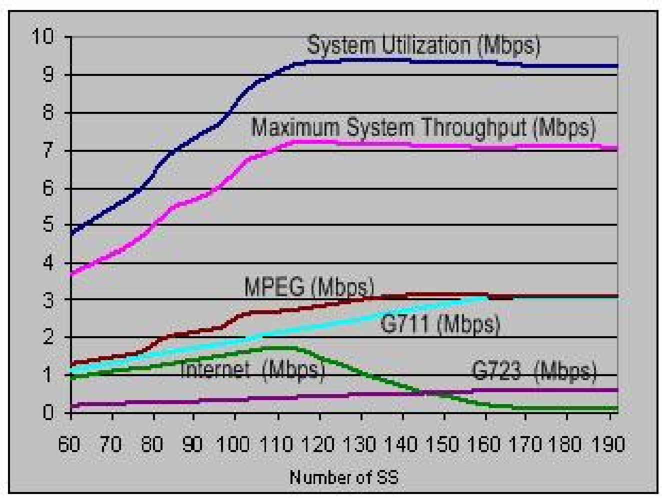

Figure 13. Throughput for all traffic types (60-150 SS).

high congestions periods (> 165 SS), EBSA makes sure that each service type gets its proportion assigned (2% for *BE*, 38% for *rtPS* and 60% for *UGS*). In terms of mean access delays, EBSA always provides balanced access delays, according to the tolerated jitter. The mean access delays for VoIP-G711, VoIP-G723 and MPEG-4 on high congestions periods were of 2 ms, 8ms and 95 ms, respectively, as shown in Figure 14. These access delays were lower than the tolerated jitter for each *UGS* and *rtPS* stream.

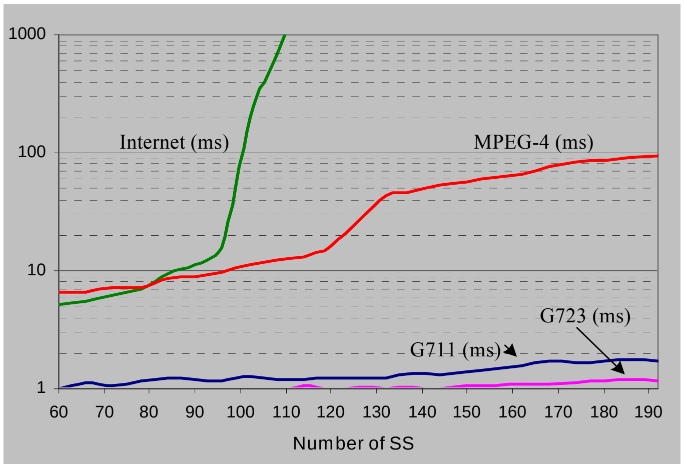

Figure 14. Mean Access Delays.

Due to space limitation we did not include a discussion about the impact of channel errors in the previous analysis of EBSA. Channel errors can degrade the QoS observed by SS in various ways depending of the particular service class being considered.

For *UGS* and *ntPS* classes for instance, losing packets in the air due to channel errors may represent a violation of the QoS flow agreement. We are currently investigating ways to overcome this problem within the EBSA framework. One simple solution already proposed by other researchers [14] is to assign additional transmission opportunities (slots in our case) to flows facing channel errors in order to keep up with QoS requirements.

Conclusion

In this paper we have presented a scheduling algorithm for IEEE 802.16 based networks. The proposed algorithm is practical, compatible with IEEE QoS requirements, and easy to implement. The purpose of *EBSA* is to provide a higher transmission priority to service flows with minimum tolerated jitter. It provides tight delays guarantees for *UGS* and *rtPS* and minimum bandwidth reservations for *nrtPS* and *BE* flows. Simulation results of *EBSA* show that real-time services, such as VoIP, can be supported with very low access delays even on high congestion periods. The performance of *EBSA* with mixed traffic sources, (*UGS*, *rtPS*, *nrtPS* and *BE*) will be further investigated through simulations and theoretical analysis. The results of such performance analysis will be provided in future publications.

References

[1] C. Fragouli, V. Sivaraman, and M. Srivastava, "Controlled multimedia wireless link sharing via enhanced class-based queuing with channel-state dependent packet scheduling," in Proc. of IEEE INFOCOM 98, vol. 2, March 1998, pp 572-580.

[2] G. Chu, D.Wang, S. Mei "A QoS Architecture for the MAC protocol of IEEE 802.16 BWA System," IEEE 2002 ICC, Circuits and Systems, vol. 1, July 2002, pp. 435 – 439.

[3] K. Wongthavarawat, A. Ganz, "IEEE 802.16 based last mile broadband wireless military networks with QoS support," in Proc. of IEEE MILCOM 2003, vol. 2, Oct. 2003, pp. 779 – 784.

[4] M. Hawa, D. W. Petr, "Quality of service scheduling in cable and broadband wireless access systems," X IEEE International Workshop on Quality of Service 2002, May 2002, pp 247 – 255.

[5] M. Andrews, "Probabilistic End-to-End Delay Bounds for Earliest Deadline First Scheduling," In Proc. of IEEE INFOCOM 2000, vol. 2, March 2000, pp. 603 – 612.

[6] L. Lenzini, E. Mingozzi, G. Stea, "Tradeoffs Between Low Complexity, Low Latency, and Fairness With Deficit Round-Robin Schedulers," IEEE-ACM Transactions on Networking, vol. 12, 2004, pp. 681-693.

[7] J. C. R. Bennet, H. Zhang. "WF2Q: Worst-case Fair WFQ", in Proc. of IEEE MILCOM 1996, San Francisco, March 2003.

[8] IEEE 802.16-2001, "IEEE Standard for Local and Metropolitan Area Networks - Part 16: Air Interface for Fixed Broadband Wireless Access Systems," April 2002.

[9] IEEE 802.16a, "IEEE Standard for Local and Metropolitan Area Networks - Part 16: Amendment 2: Medium Access Modifications and Additional Physical Specifications for 2–11 GHz," April 2003.

[10] DOCSIS 1.1, "Data-Over-Cable Service Interface Specifications - Radio Frequency Interface 1.1 Specification," July 2000.

[11] V. Rangel and R. M. Edwards, "Performance Evaluation of Adaptive Contention Slot Allocators for CATV Networks based on the European Cable Communications Protocol DVB/ DAVIC", Journal of the SCTE, vol. 24, Sep. 2002, pp 24-30.

[12] Vinko Erceg et al, "An empirically based path model for wireless channels in suburban environments," IEEE Journal on Selected Areas in Communications, vol. 17, Jul. 1999, pp 1205-1211.

[13] D. Hertrich, "MPEG4 video transmission in wireless Basic QoS support on the data link," Minor Thesis, Faculty of EE and CS, Technical University of Berlin, October 2002.

[14] S. Lu, V. Bharghavan, and R. Srikant. Fair Scheduling in Wireless Packet Networks.