On Dynamic Resource Allocation in WiMAX Networks

ROLA NAJA [12]
ABED AHMAD[3]

[1]Laboratoire PRiSM (CNRS 8144), Université de Versailles, 45 avenue des Etats-Unis, 78035 Versailles Cedex
[2]Département Informatique, ESIEE Paris,2 bd Blaise Pascal 93162 Noisy Le Grand Cedex Marne-la-Vallée
[3]Université Libanaise- Laboratoire LaMA-Liban
{r.naja@esiee.fr, abdahmad@hotmail.com}

Abstract. The emerging broadband wireless access WiMAX provides wireless users with ubiquitous access to different classes of services. One of the main challenging problems in WiMAX networks is the resource allocation and the quality of service provisioning. In this paper, we tackle this problem and apply a Double Movable Boundary Scheme (DMBS) which is based on dynamic sharing of resources between different traffic categories provided by a common resource pool. Besides satisfying each connection's QoS requirement, DMBS raises the throughput by dynamically adjusting resource allocation. Simulation results show that DMBS reduces real-time traffic delay and blocking probability and increases bandwidth usage while respecting non-real time traffic requirements.

Keywords: WiMAX, quality of service, scheduling.

(Received May 14, 2010 / Accepted September 15, 2010)

1 Introduction

Broadband wireless access networks have evolved to satisfy the increasing demands of users for ubiquitous and seamless access to the broadband service. The emerging IEEE 802.16e (WiMAX), is one of the most promising solutions for the last mile broadband access to support high data rate, high mobility and wide coverage at low cost.

However, due to the mobile station characteristics and several factors like quality of service requirements of real-time traffic, the provision of reliable data transmission and low-latency wireless communications is challenging for the network operator.

Even though the physical layer specifications and the Medium Access Control (MAC) protocol signalling are well defined in the standard [4], the resource allocation and admission control policies for the IEEE 802.16 air-interface remain as open issues. Thus, quality of service provisioning in WiMAX networks is an imperative and challenging problem to resolve.

In this paper, we will highlight quality of service methods adopted in WiMAX networks and propose a dynamic scheduling scheme that strives to meet the users' expectation.

In order to support Quality of Service (QoS) for different traffic types, MAC protocol defines several bandwidth request-allocation mechanisms and five classes of service: Unsolicited Grant Service (UGS), Real-Time Polling Service (rtPS), Extended Real-Time Polling Service (ertPS), Non-Real-Time Polling Service (nrtPS), and Best Effort (BE).

With UGS, the amount of allocated bandwidth is fixed, and explicit bandwidth request is not required. In the case of rtPS, the Base Station (BS) provides unicast request opportunities for a Subscriber Station (SS) to send its request at a predefined interval. In other words, the BS periodically polls the SS to allocate the Uplink (UL) bandwidth request. ErtPS is a combination of UGS and rtPS: Unsolicited periodic grants of bandwidth are provided but with flexibility in having dynamic data rates. The SS may request changing the size of the UL allocation.

As for nrtPS, the BS polls the SS less frequently than in rtPS. However, SSs are allowed to use contention request opportunities to send a bandwidth request message. With BE, bandwidth messages can only be transmitted through contention request opportunities, thus the performance achieved can vary sharply.

There have been several proposals for scheduling the above-mentioned classes of service in the literature. Some of the proposed algorithms were defined in an attempt to meet QoS requirements of classes of service or to provide fairness to the users or to strike a balance between bandwidth utilization and QoS. Other algorithms rely on channel characteristics for decision making processes.

In this paper, we propose to adopt a hierarchical scheduling algorithm deployed in two layers: Once legacy scheduling algorithm is applied within each class of service in the first layer, we apply in the second layer the Double Movable Boundary Scheme (DMBS) algorithm to the different classes of service.

The interest of the DMBS allocation scheme lies in the dynamic sharing of resources between different traffic categories provided by a common resource pool. This helps in relaxing the congestion conditions of a certain traffic at asymmetrical offered loads while permanently preserving a certain minimum number of resources for each category. Better channel utilization is then achieved while guaranteeing a certain quality of service for each traffic.

In this paper, we applied the DMBS algorithm in WiMAX networks and restricted our study to two classes of service (UGS, nrtPS).

Performance results show that the algorithm respects the priority of real-time connections and prevents least-priority classes starvation problem. In fact, we strive to achieve two major components: fairness to different classes of service and service differentiation.

The rest of this paper is organized as follows: In section 2, we introduce IEEE 802.16 scheduling algorithms proposed in the literature. Section 3 highlights the DMBS algorithm applied in WiMAX networks. In section 4, we present extensive simulation results to evaluate the performance of the proposed scheme. Finally, we conclude the paper in section 5.

2 Scheduling Algorithms

Several strategies have been proposed in the literature to allocate the radio channel bandwidth for transmitting different traffic types in WiMAX networks.

We have made a study of the existing scheduling algorithms and found that they fall into three categories: simple scheduling, cross-layer scheduling and hierarchical scheduling.

2.1 Simple scheduling algorithms

These are legacy scheduling algorithms that are considered for all classes in order to provide QoS, flow isolation and fairness. In [3], Hawa et al. proposed a bandwidth allocation scheme that divides the service classes into three types of connections. Each type has a different scheduling scheme: First In First Out (FIFO), Weighted Fair Queuing (WFQ), and Priority Queue (PQ).

Some research studies propose to have a PQ scheduling: The class with the highest priority (UGS) will be served until its queue gets empty. Then, comes the turn of lower priority class. The problem with PQ is that the lower priorities such as BE could go through bandwidth starvation whenever the highest priority flow continues for a long time.

Custom Queuing (CQ) solves the PQ problem by applying a fixed boundary scheme. Each traffic type has the same number of slots permanently allocated to it and hence encounters no competition from other types to share its resources. This policy may be inefficient in case resources are not fully utilized and hence are wasted by one traffic type while another is suffering from congestion.

In [2], authors define a Dynamic Allocating Priority Queue (DAPQ) scheduling scheme based on CQ in which each traffic type gets a portion of the resource. If an assigned bandwidth is unused, it will be allocated to another traffic type. This solution needs an efficient tuning of the bandwidth reservation percentages.

The Conventional Movable Boundary Strategy (CMBS) overcomes this drawback by allowing a limited sharing of resources. Non-real-time traffic can be allocated extra channels if they are not used by real time traffic but are pre-empted by the latter when it requests resources. Movable boundary strategies are found to achieve a reduction of queuing delay for real-time traffic compared to fixed boundary ones [6].

2.2 Cross-layer scheduling algorithms

Scheduling algorithms in this category are primarily exploiting the variability in channel conditions and modulation schemes based on SS locations.

An Opportunistic scheme is presented in [11], in which channel characteristics are used as parameters for decision making processes.

Sayenko algorithm satisfies each connection's minimal bandwidth requirement with considering adopted modulation, and then equally allocates the remaining bandwidth to each connection [10].

In [7], the authors proposed an algorithm that dynamically adjusts the downlink/uplink bandwidth ratio, satisfies connections QoS bandwidth requirements and allocates more bandwidth to the connections with better channel quality for promoting the throughput. Lack of fairness is the major drawback of this proposal. In order to resolve this problem, a downlink opportunistic fair scheduling scheme is proposed in [8]. A scheduler at the BS decides the order of downlink bursts to be transmitted. The decision is made based on the quality of the channel and the history of transmissions of each SS. It takes advantage of temporal channel fluctuations to increase the BS's throughput and maintain fairness by balancing the long term average throughput of SSs.

In [9], an opportunistic scheme is proposed for scheduling heterogeneous traffic in the downlink. The scheduler, located at the base station, uses the information of the channel and queue status of the users to schedule the traffic with different quality of service requirements and different arrival rates admitted to the queues in the base station. The scheduler deploys a differentiation technique, based on a notion of stability, to satisfy different service rate requirement of heterogeneous traffic types. This proposed solution offers prescribed delay for real-time traffic and rate guarantees for non-real time traffic.

2.3 Hierarchical scheduling algorithms

This category combines several queuing disciplines in more than one layer. Once bandwidth has been assigned to each class using the first layer scheduling algorithms, a legacy algorithm is executed within each class in the second layer. An important aspect of algorithms in this category is the overall allocation of bandwidth among the scheduling services.

Wongthavarawat in [12] adopted a two-layer scheduling scheme where bandwidth between different service classes is allocated with strict priority in the first layer and each service class has its own scheduling algorithm in the second layer. Chen and Jiang individually kept the Wongthavarawat approach in the second layer. However, they modified the first layer in an attempt to solve the starvation problem caused by adopting strict priority [1][5]. In the first layer, Chen used Deficit Fair Priority Queue (DFPQ) and Jiang used leaky bucket.

In this paper, we propose a hierarchical scheduling algorithm and provide service differentiation via the Double Movable Boundary Scheme.

H. Koraitem and S. Tohme, have proposed DMBS [6] and applied it in a satellite TDMA-based systems. They showed that it improves channel throughput achieved at high loading conditions and reduces blocking probability for real time connections.

We propose to implement DMBS in WiMAX networks. The main objective of our work is that DMBS strategy periodically modifies the resource allocation decision at the beginning of each frame to adapt it to the network loading conditions while always guaranteeing a minimum of resources for each traffic class at all loading conditions.

In detailing and analyzing the proposed DMBS scheme, we limit our discussion to two classes of services: connection oriented UGS and connectionless bursty nrtPS.

A very significant factor which limits the allocation decision is the very high ratio of the service rates of both traffic types. While UGS connections may last for minutes, data bursts durations are relatively short. The proposed technique permits data bursts to take possession of resources available in the UGS compartment in the absence of connection requests. UGS connections however have always the priority of their reserved resources when they arrive.

The key idea for this algorithm is the following: A common resource pool, the CRP, contains a number of resources which can be dynamically shared between traffic categories. The sharing process is tightly regulated to protect the relative quality of service parameters when the system is subject to congestion conditions.

The contribution of this algorithm in WiMAX networks is twofold:

- The proposed algorithm increases efficiency of bandwidth allocation without violating QoS requirements.
- The algorithm provides a service differentiation without implying a starvation for the least-priority class of service.

3 Applying Double Movable Boundary Scheme in WIMAX Networks

3.1 Context of the DMBS proposal

Our work considers point-to-multipoint architecture of WiMAX networks where transmission occurs between a BS and SSs and all transmission are based on a connection-oriented approach. The identification of each connection together with its service flow requirements provides the MAC layer with the basic QoS, which needs to be maintained for a specific connection.

In this paper, we restrict our study to the downlink. However, we intend to apply our algorithm in the uplink in future works.

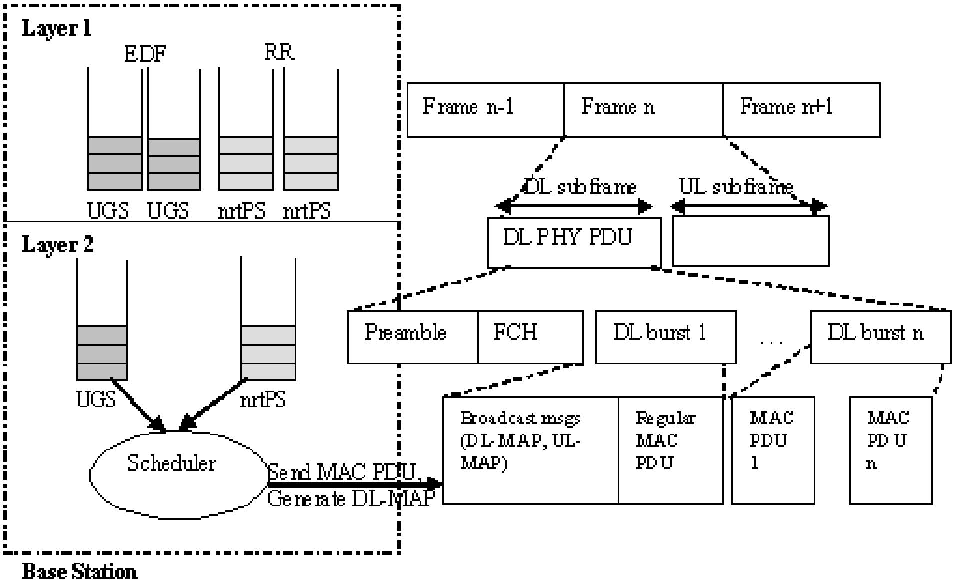

Figure 1: Base station downlink scheduler

3.2 Downlink scheduler

Each downlink connection has a packet queue at the BS (figure 1). In accordance with the set of QoS parameters and the status of the queues, the BS downlink scheduler selects from the downlink queues, on a frame basis, the next Service Data Units (SDUs) to be transmitted to SSs. Our proposed downlink scheduling algorithm works in two layers. In the first layer, connections of the same class of service are scheduled using an appropriate algorithm: different algorithms are employed for classes of service to meet their QoS requirements as illustrated in Figure1.

For all connections holding UGS traffic, the scheduler uses Earliest Deadline First (EDF), where packets with the earliest deadline will be scheduled first. For nrtPS connections, the scheduler uses RR queuing discipline, where each service gets a fair share of the allocated bandwidth in a RR fashion. RR assigns time slots without priority to each service flow in equal portions and in order.

In the second layer, the MAC scheduler proceeds as follows:

1. UGS and nrtPS queues are served according to the Double Movable Boundray Scheme as described in sub-section 3.4.

2. The BS notifies the SSs about their bandwidth allocations via the DL-MAP message which contains the timetable of the downlink grants in the forthcoming downlink sub-frame as specified next.

3.3 TDD frame structure

In TDD-based OFDM systems, which form by far the most common form of IEEE 802.16 implementation, a frame is divided into two sub-frames:

1. An uplink sub-frame, which consists of contention intervals scheduled for initial ranging and Bandwidth Request (BR) and one or multiple Uplink Physical PDU (UL PHY PDUs). Initial ranging is used to acquire the correct timing offset and power adjustments. BR purposes consist to make resource reservation.

2. A downlink sub-frame (figure 1), which consists of only one Downlink Physical PDU (DL PHY PDU). The DL PHY PDU carries the data of all the SSs. It begins with a preamble which helps synchronize all the terminals and mark the beginning of a new frame. The preamble is followed by the Frame Control Header (FCH) which carries the frame configuration information including the MAC Access Protocol (MAP) message length and the sub-channel information.

The broadcasted DL-MAP and UL-MAP MAC management messages will be in the burst following the FCH and define the access to the downlink and uplink information respectively. The DL-MAP is a MAC management message that defines burst start times on the downlink. Equivalently, the UL-MAP, following the DL-MAP message, is a set of information that defines the entire (uplink) access for all SSs during a scheduling interval.

Information in the DL-MAP is about the current frame. Information carried in the UL-MAP concerns a time interval starting at the Allocation Start Time (measured from the beginning of the current frame) and ending after the last specified allocation. Then, various SSs that are active send their data in the assigned time slots according to the scheduling algorithms as specified in next sub-section.

4 DMBS Model

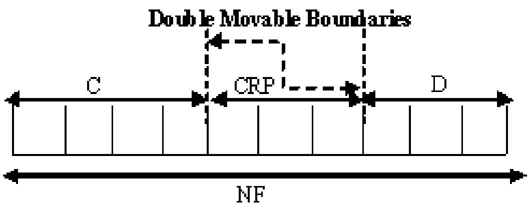

Figure 2: DMBS strategy

In the work at hand, we tackled the problem of the movable boundary allocation technique taking into account the different quality of service requirements of multimedia traffic.

The total number of resources slots, NF, on the downlink frame is divided into three distinct parts: one part is reserved for UGS traffic composed of a number C of slots, the second for nrtPS traffic composed of a number D of slots and the third part constitutes a Common Resource Pool CRP such that (figure 2):

$$CRP \quad NF - C - D \tag{1}$$

where C represents the total resources available for UGS calls at normal loading conditions to achieve a certain guaranteed blocking probability.

The CRP contains a number of common resources which can be dynamically shared between both traffic categories. The sharing process is tightly regulated to protect the relative quality of service parameters when the system is subject to congestion conditions.

nrtPS bursts can make use of the available resources in both the CRP and the UGS compartments when UGS is not fully utilizing them. In addition, UGS connection requests can be assigned resources from the CRP when nrtPS data queue is below a certain threshold.

Several measures of protection are however established to prevent nrtPS queues from building up when its load rises in the network. This will be explained in more details in the following sub-sections. The detailed procedure for sharing and allocating frame resources is explained in the next paragraphs.

UGS Allocation Procedure The DMBS allocation algorithm is initialized at the beginning of each control period and starts by examining and allocating channels to waiting UGS reservation requests. If resources are available in the UGS sub-frame, waiting UGS requests will be granted a channel slot throughout the call duration. If there are still more requests waiting to be satisfied after all resources on the UGS sub-frame are consumed, more channels can be allocated for these waiting calls from the CRP.

The previous step, however, is taken after monitoring the length of the data queue reservations requests:

- If the latter is less than a certain pre-specified threshold DQT, UGS calls are granted access to some resources in the CRP sub-frame.

- Conversely, if the data queue length exceeds the specified threshold, UGS requests are denied access to the CRP sub-frame. They are then made to wait for the release of resources either in the UGS sub-frame or in the CRP only when the data queue length goes below the threshold.

  The waiting room of UGS call connection requests, L, is finite to limit the call establishment delay. A call may also be blocked if the UGS queue is completely full when the UGS call request arrives.

nrtPS Allocation Procedure Next comes the turn for allocating resources to nrtPS reservation requests. Their requests are generally queued until resources are available on the frame. First bursts are allocated resources from the data sub-frame. If there are still waiting data requests, the CRP and the UGS sub-frames are searched for available channels which can then be assigned to data for the duration of the burst.

Occupying resources in the nrtPS sub-frame is not penalizing to UGS calls since burst durations are much smaller compared to UGS call durations.

nrtPS Queue Threshold The nrtPS queue threshold, DQT, choice has a considerable impact of the overall performance. Its value can be either permanently fixed or dynamically variable to follow the change in the network loading status.

In the first case, a certain value is specified for the queue threshold and maintained during the operational time of the scheme. In the dynamic threshold case, several techniques can be implemented to vary the DQT value during the operation of the allocation scheme. One of these is the CRP dependent technique.

The threshold is dynamically varied in each control period as a function of the number of UGS borrowed resources from the CRP. This will provide protection to the bursty data traffic against the continuous rise of UGS load.

The following equation permits to compute the DQT value:

$$DQT \quad \frac{Lth.C}{C \quad N} \tag{2}$$

where Lth is the initial threshold and N is the number of allocated CRP resources to UGS calls.

5 Performance Evaluation

In order to validate our model, we have made a simulation in MATLAB and adopted the parameters depicted in table1. We considered three schemes: CQ, PQ, CMBS and DMBS and made a performance comparison. With CQ and CMBS, we reserve D slots for nrtPS traffic. While the boundary is static with CQ, it is movable with CMBS.

We computed the following performance parameters : UGS blocking probability (Pb), mean UGS waiting delay (Dr) , mean nrtPS waiting delay (Dn) and bandwidth percentage usage (U). The UGS blocking probability refers to the ratio of blocked UGS calls number (due to unavailable resources) over the total UGS calls number.

UGS call process arrival is generated according to a Poisson process, with a call duration exponentially distributed. nrtPS calls arrive according to a Poisson. They consist of a sequence of packet bursts exponentially distributed. Each burst is divided into packets whose length fits in one frame slot.

We first exhibit performance results of PQ, CQ, DMBS and in a second stage, we compare performance of DMBS and CMBS for different values of UGS and nrtPs loads measured in Erlang (E).

Table 1: Simulation parameters

Frame duration	5ms
Channel Bandwidth	7MHz
Total Symbols number	34
Data Symbols number (F)	20
C	12
D(DMBS)	2
CRP	6
D(CQ,CMBS)	8
L	10
UGS mean call duration	3min
Symbol duration	0.144ms
Bits number per slot	192(QPSK :1/2) ; 576 (QAM-16 :3/4)
Total number of frames	50000

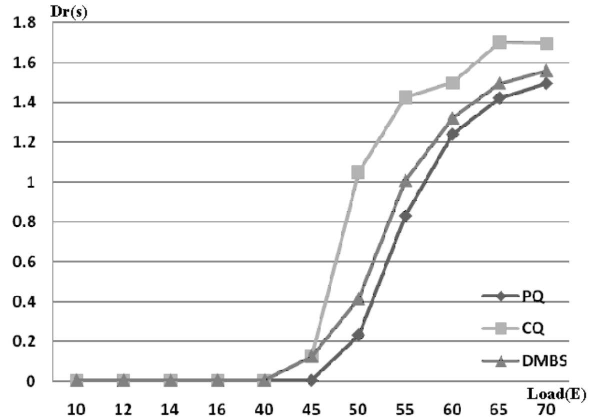

Figure 3: Mean UGS delay, nrtPS load=1E, UGS load variable

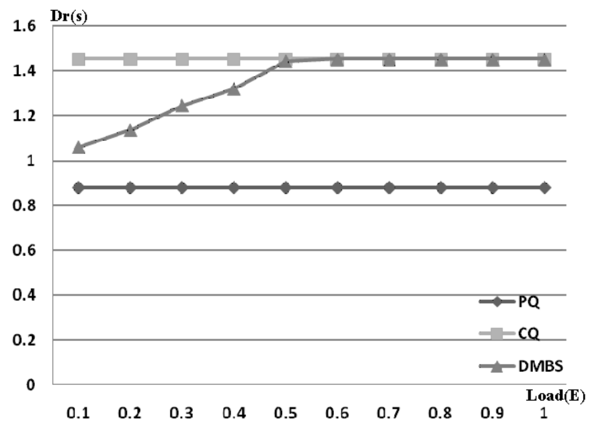

Figure 4: Mean UGS delay, UGS load=55E, nrtPS load variable

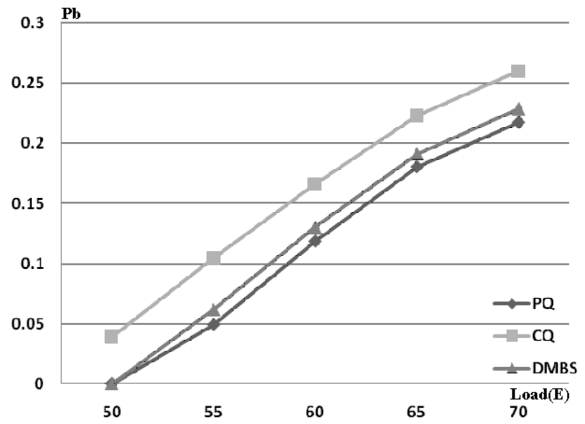

Figure 5: Blocking probability, nrtPS load=1E, UGS load variable

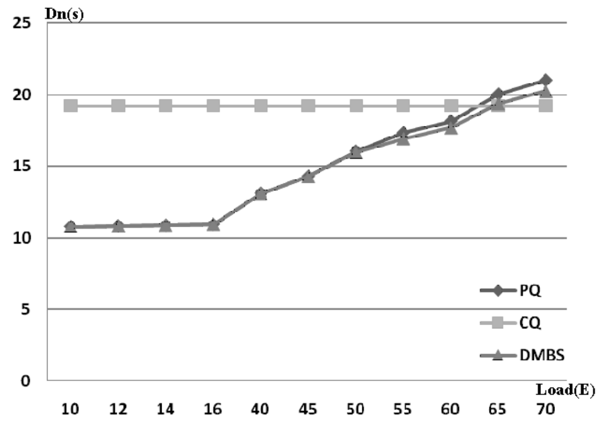

Figure 7: rtPS Mean delay, nrtPS load=1E, UGS load variable

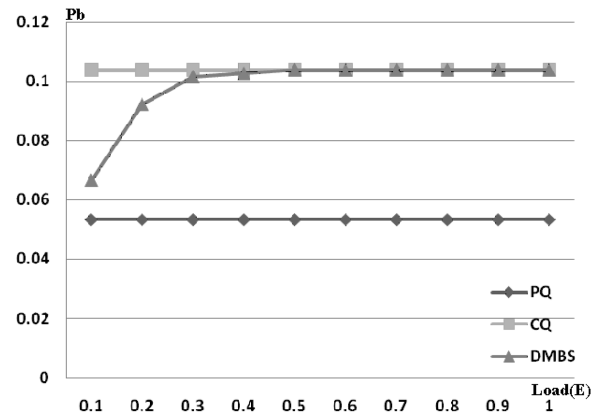

Figure 6: Blocking probability, UGS load=55E, nrtPS load variable

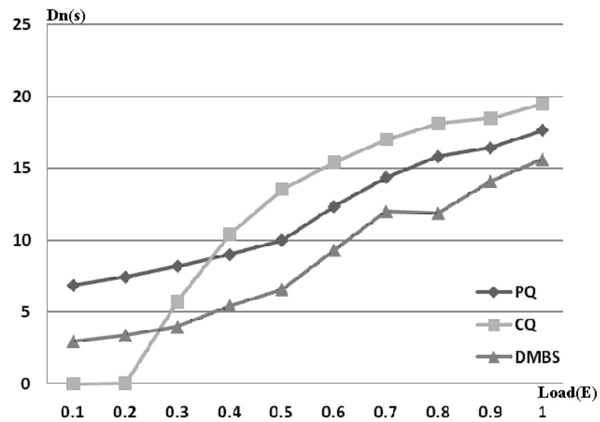

Figure 8: nrtPS Mean delay, UGS load=55E, nrtPS load variable

5.1 DMBS, CQ, PQ performance evaluation

This sub-section addresses the comparison issue between CQ, PQ and DMBS. Figure 3 shows mean delay of UGS calls with different values of UGS load and a nrtPS load equal to 1 Erlang (E). PQ achieves least delays due to the strict priority assigned for the UGS traffic, whereas CQ implies highest delays due to the static bandwidth allocation for UGS; this type of allocation will not be efficient for high UGS loads. DMBS achieves delays slightly higher than PQ: in fact, a part of the CRP pool will be assigned to the nrtPS packets with DMBS. This will impact the UGS delay at high UGS loads.Figure 4 shows mean delay of UGS calls with different values of nrtPS load and a UGS load equal to 55E. For a constant UGS load, CQ and PQ achieve a constant delay independent from the nrtPS load variation. As for DMBS, the UGS mean delay is less than that of CQ and higher than PQ. In fact, with high nrtPS load, nrtPS queue will build up and exceed threshold DTQ. Thus, CRP slots will be allocated to nrtPS. This will affect the mean UGS delay.

Figures 5 and 6 depict UGS blocking probabilities respectively with a contant nrtPS load and a constant UGS load. Similar conclusions are drawn : DMBS achieves better blocking probability than CQ. For a constant UGS load, PQ and CQ induce constant blocking UGS probabilities.

Figures 7 and 8 illustrate mean waiting nrtPS packet delay respectively for a constant nrtPS load and a constant UGS load. For a constant nrtPs load, CQ achieves a constant delay when the nrtPS load is constant: for a static bandwidth allocation, nrtPS traffic experiences same delay with a constant load. DMBS and PQ induce similar delay for a UGS load less than 50E. Then, DMBS achieves lower delays than PQ due to the CRP

slots.

Figures 9 and 10 illustrate the percentage of bandwidth usage respectively for a constant nrtPS load and a constant UGS load. PQ and DMBS achieve better usage than CQ. In fact both algorithms adopt a dynamic allocation of frame slots, whereas CQ has a static allocation scheme which does not adapt resource allocation to traffic fluctuation.

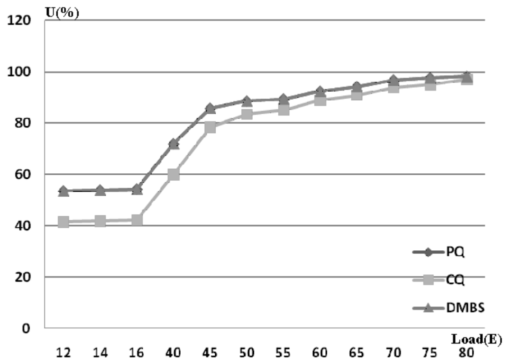

Figure 9: Bandwidth usage percentage, nrtPS load=1E, UGS load variable

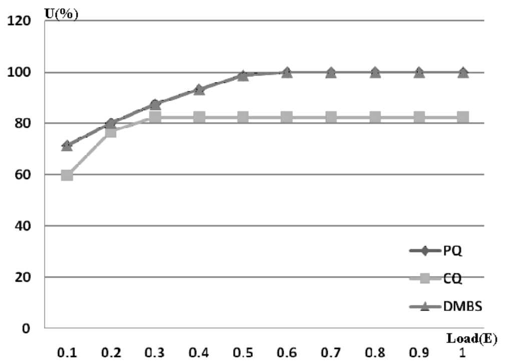

Figure 10: Bandwidth usage percentage, UGS load=50E, nrtPS load variable

5.2 DMBS and CMBS performance evaluation

This section is devoted to performance evaluation of DMBS and CMBS schemes. Figure 11 shows that DMBS achieves better bandwidth usage than CMBS. In fact, CMBS adopts a movable boundary in one direction: Whereas nrtPS packets use available slots in the UGS compartment, UGS requests do not have the permission to be served by unused slots in the nrtPS compartment. This is not the case for DMBS scheme. Therefore, UGS packets experience lower blocking probability with DMBS than with CMBS (figure 12).

Figures 13 and 14 illustrate mean nrtPS waiting delay and mean UGS delays for variable nrtPS load. We can see that DMBS achieves slightly higher nrtPS delays and significant less UGS delays than CMBS. This result is a direct consequence from the double movable boundary in which both UGS and nrtPS traffic profit from a common resource pool.

Available slots in UGS frame part are used by nrtPS packets with CMBS and thus the boundary moves in one direction only, contrarily to DMBS in which UGS packets use slots in the common pool. The DMBS dynamic adaptive boundary permits to reduce UGS delays and slightly increase nrtPS delays. The increase of nrtPS delay is tuned by the nrtPS queue threshold.

Several factors have their impact in determining the performance of this allocation technique. These include the data queue threshold value and the fraction of the borrowed resources from the CRP. The data queue threshold value has a considerable influence on the real-time resource allocation performance and consequently the efficiency of the DMBS policy.

A low data queue threshold favours data over UGS traffic by reducing its waiting delay in the queue at the expense of risking a rise in the blocking probability and call setup delay for UGS traffic. A high threshold value on the other hand increases the chance of accepting a UGS reservation request while increasing mean data delay. The DQT choice depends then on the type of offered services, their relative priorities for the operator and the quality of service parameters guaranteed by the network for each traffic category.

Another important factor which largely influences the performance of the DMBS policy is the number of resources allocated to UGS calls from the CRP. Obviously, allocating a large number of CRP resources to UGS traffic at a time can dramatically degrade the delay performance of bursty data.

On the contrary, if a smaller fraction of the CRP resources are allocated to UGS per control period, this can leave a better chance for data to benefit from the left resources in the CRP as its load goes up. Together with the dynamic data queue threshold DQT value, these two factors must be properly tuned in order to protect nrtPS and UGS traffic against large queuing delays.

6 Conclusion

In this paper, we tackled the problem of MAC downlink scheduling at a WiMAX base station. We proposed to apply the DMBS scheme which enables a dynamic adaptive resource allocation. The performance of DMBS was compared to that of CQ, PQ and CMBS schemes. The DMBS scheme was shown to be more efficient particularly in asymmetrical loading conditions compared to the conventionally considered movable boundary techniques. The efficiency is reflected in the improved channel throughput achieved at high loading conditions and the reduced blocking probability for real-time connections. This fact is proven by simulation models and analysis developed to evaluate the performance parameters of DMBS.

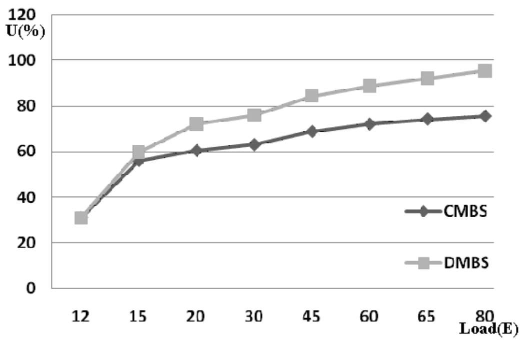

Figure 11: (Bandwidth usage percentage, nrtPS load=0.15E, UGS load variable

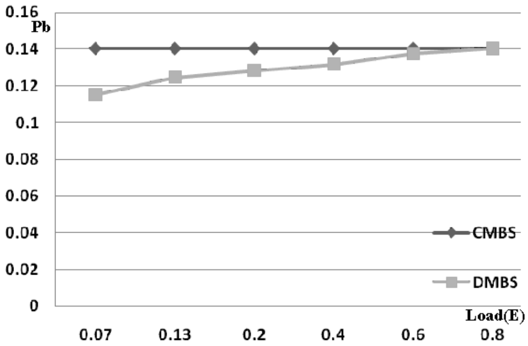

Figure 12: Blocking probability, UGS load=60E, nrtPS load variable

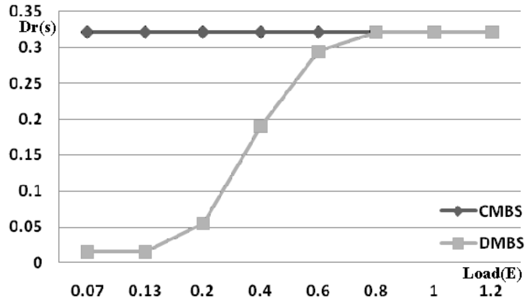

Figure 13: Mean rtPS delay, UGS load=45E, nrtPS load variable

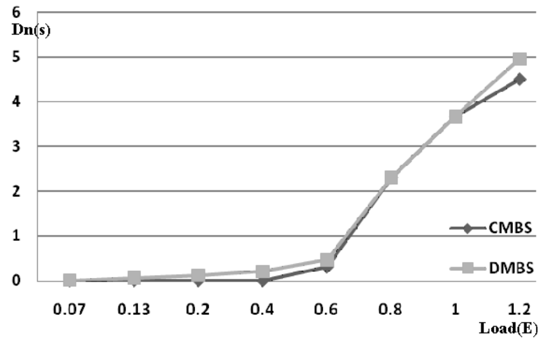

Figure 14: Mean nrtPS delay, UGS load=45E, nrtPS load variable

It is noteworthy that data queue threshold value and the fraction of the borrowed resources from the CRP largely influence the performance of the DMBS policy. Both factors should be well defined and dynamically tuned to reach an optimum operation. In our future works, we intend to evaluate the performance of DMBS strategy while testing the influence of each of its determining parameters. Finally, this paper tackles the resource allocation in the downlink. Nevertheless, DMBS algorithm can be applied in the uplink as well: this matter will be another perspective that we envision to investigate.

References

[1] Chen, J., Jiao, W., and Wang, H. A service flow management strategy for ieee 802.16 broadband wireless access systems in tdd mode. *IEEE ICC*, 5:3422–3426, 2005.

[2] Esmailpour, A. and Nasser, N. Packet scheduling scheme with quality of service support for mobile wimax networks. *IEEE WLN*, pages 1040–1045, 2009.

[3] Hawa, M. and Petr, D. W. Quality of service scheduling in cable and broadband wireless access system. *IEEE International Workshop on Quality of Service*, pages 247–255, 2002.

[4] IEEE. Std 802.16e-2005: Ieee standard for local and metropolitan area networks-part 16: Air interface for fixed and mobile broadband wireless access systems. 2005.

[5] Jiang, C. H. and Tsai, T. C. Token bucket based cac and packet scheduling for ieee 802.16 broadband wireless access network. *IEEE CCNC*, 2006.

[6] Koraitim, H. and Tohmé, S. Resource allocation and connection admission control in satellite networks. *IEEE Journal on Seleted Areas in Communications*, 17(2):360–372, 1999.

[7] Lai, Y. and Chen, Y. A channel quality and qos aware bandwidth allocation algorithm for ieee 802.16 base stations. *IEEE AINA*, pages 472–479, 2008.

[8] Mehrjoo, M., Dianati, M., Shen, X., and Naik, K. Opportunistic fair scheduling for the downlink of ieee 802.16 wireless metropolitan area networks. *ACM Quality of Service in Heterogeneous Wired/wireless networks*, 191, 2006.

[9] Mehrjoo, M. and Shen, X. An efficient scheduling scheme for heterogeneous traffic in ieee 802.16 wireless metropolitan area networks. *IEEE IST*, pages 263–267, 2008.

[10] Sayenko, A., Alanen, O., Karhula, J., and Hamalainen, T. Ensuring the qos requirements in 802.16 scheduling. *MSWiM '06*, 2006.

[11] Tarchi, D., Fantacci, R., and Bardazzi, M. Quality of service management in ieee 802.16 wireless metropolitan area networks. *IEEE ICC*, 4:1789–1794, 2006.

[12] Wongthavarawat, K. and Ganz, A. Packet scheduling for qos support in ieee 802.16 broadband wireless access systems. *International Journal of Communication Systems*, 16(1):81–96, 2003.