A Cross-Layer Qos Based Scheduling Algorithm WRR Design in Wimax Base Stations

Hattab Guesmi1,2*, Sassi Maaloul3

1Laboratory of Electronic and Micro-electronic, Monastir University, Monastir, Tunisia
2Faculty of Science and Arts Addayer, Jazan University, Jazan, KSA
3Unité de Recherche Médiatron de SupCom, Carthage University, Tunisia
*Corresponding author: Hattab.guesmi@fsm.rnu.tn, hguesmi@jazanu.edu.sa

Received December 23, 2012; Revised January 26, 2013; Accepted February 26, 2013

Abstract The IEEE 802.16 standard defines a wireless broadband access network technology called Wimax. It introduces several advantages, one of which is the support for Quality of Service (QoS) at the MAC level. To ensure meeting the QoS requirements, the 802.16 base stations must run some algorithms to allocate slots between connections. Call admission and scheduling are the strongest tools in our hand to ensure QoS. We propose an efficient design architecture that is capable of allocating slots based on the QoS requirements, bandwidth request sizes, and the 802.16 network parameters. To test the proposed solution, we have implemented a cross layer between the 802.16 MAC and the network layers in the NS-2 simulator and, then in RTL level with VHDL to be designed in FPGA. Several simulation scenarios are presented. According to the simulation results, the proposed scheduling solution ensures the QoS requirements of all 802.16 service classes. The solution shares free resources weighted fairly and demonstrates work-conserving behavior. The proposed design in this paper was analyzed: simulation results show a significant performance improvement in terms of overall throughput and delay when compared to recently published work.

Keywords: Cross-layer, design, IEEE 802.16, scheduling, QoS, WiMax, WRR

1. Introduction

IEEE 802.16 is a standard for wireless broadband access network. The main advantages of 802.16 when compared to other wireless network access technologies, such as 802.11, are the longer transmission range and more sophisticated support for QoS at the MAC level. The MAC is designed to support convergence of various application and service types can be used in Wimax networks. The standard defines two basic operational modes: point-to-multipoint (PMP) and Mesh. While a subscriber station (SS) can communicate with other stations and with the base station (BS) in the Mesh mode, it is allowed to communicate only through the BS in the PMP mode. An important principle of 802.16 is that it is connection oriented. This means that an SS must register with the base station before it can start to send or receive data. During the registration process, an SS can negotiate the initial QoS requirements with the BS. These requirements can be changed later, and a new connection may also be established on demand.

The basic approach for providing the QoS guarantees in the 802.16 network is that the BS does the scheduling for both the uplink and downlink directions. However, in other words, an algorithm at the BS has to translate the QoS requirements of SSs into the appropriate number of slots within the 802.16 frame. When the BS makes a scheduling decision, it informs all SSs about it by using the UL-MAP and DL-MAP messages at the beginning of each frame. These management messages explicitly allocate slots to each SS in the uplink and downlink directions. However, the scheduling policy (an algorithm to allocate slots) is not defined in the 802.16 specification, but rather is open for alternative implementation. There are several research works on 802.16 QoS scheduling that present architecture and scheduling disciplines to guarantees QoS. However, in [1,2] the authors have focused mainly on the scheduling issues and components of the QoS architecture without presenting any exact method. Several research works propose using complex scheduler, such as earliest deadline first (EDF), deficit round robin (DRR) [3], weighted fair queuing (WFQ) [4], weighted fair queuing with decoupled bandwidth [5], and even a hierarchy of schedulers. However, it is a challenging task to use a hierarchy of schedulers because the per-connection QoS requirements must be translated into the scheduler configuration at each level. Furthermore, it is not enough to calculate the scheduler configuration only once when an SS joins or leaves the network. As SSs send data, their request sizes change all the time. As a result, the scheduler at the BS should reassign slots for every 802.16 frame to achieve an accurate and fair resource allocation. For instance, if there are 400 FPS (frames per second) [1], then the BS must make 400 scheduling decisions per second. The OFDMA PHY specification allows the sending of up to 500 FPS. This is precisely reason why we suggest using one level with a simple scheduling that is based conceptually on the weighted round robin (WRR) approach. A simpler
solution is better because there is not much time for the scheduling decision. This paper presents a scheduling solution for the 802.16 base stations. When compared to previous research, our solution supports all the 802.16 service classes. The 802.16d MAC provides QoS differentiation for different types of applications that might operate over 802.16 networks that define the following types of services: Unsolicited Grant Service, Real-time Polling Service, Non-real-time Polling Service, and Best Effort.

2. Overview of the QoS Management

QoS corresponds to all mechanisms that allow a network to distribute equitably and according to requirements of applications all the resources offered by networks, to provide the need quality. Also, it can be characterized by different performance criteria that include basic availability, the loss rate, throughput, average delay, security, etc… To provide an efficient level of QoS support, many mechanisms can be managed tightly coupled and plays an essential role in determining the network performance:

Admission control is a network Quality of Service (QoS) procedure. Admission control determines how bandwidth and latency are allocated therefore need to be implemented between network SS and BS to control the traffic entering the network. The role of CAC is to control the number of connection flows into the network. A new connection request is progressing only when sufficient resources are available at each successive network element to establish the connection through the whole network based on its service category, traffic contract and QoS, while the agreed QoS of all existing connections are still maintained. Admission control is useful in situations where a certain number of connections may all share a link, while an even greater number of connections cause significant degradation in all connections to the point of making the BS useless such as in congestive collapse.

Buffer management refers to any particular discipline used to regulate the occupancy of a particular queue where packets may be held (or dropped). Buffer is set to improve link utilization and system performance, but it also increases packet's queue delay. With the increase of user demands for service quality, providing stable and low delay was the primary requirement of real-time services. The most important and easy controls part of total delay is queue delay. So how to set the capacity of the buffer, how to control efficiently buffer length while network circumstance is dynamic and how to achieve the tradeoff between throughput and queue delay. These are the important problems to be solved in buffer management and QoS control of whole networks.

Packet scheduling refers to the decision process used to choose which packets should be served, it is the process of resolving contention for bandwidth. The target of a scheduling algorithm has to determine the allocation of bandwidth among the users and their transmission order. One of the most important tasks of a scheduling scheme resides in satisfying the Quality of Service (QoS) requirements, while efficiently utilizing the available bandwidth.

At the same time, a scheduling algorithm can also decide when to send any given packet (or flow’s packet), a choice that can help guarantee (or bound) packet latencies through the router. The idea of scheduling is to adapt the policy of transmission of packets in buffers, according to the requirements of QoS for flows. Scheduling has a significant impact not only on the average delay but also on the buffer size. Scheduling is used to control the resources distribution between the classes of service. Many legacy scheduling algorithms, able to provide certain guaranteed QoS, have been developed for wireless networks. The characteristics of wireless communication pose special problems that do not exist in wire line networks. They include:

- High error rate and bursty errors;
- Location-dependent and time-varying wireless link capacity;
- Scarce bandwidth;
- User mobility;
- Power constraint of the mobile hosts.

All of the above characteristics make developing efficient and effective scheduling algorithms for very challenging wireless networks. WiMAX networks provide services for heterogeneous classes of traffic with different quality of service (QoS) requirements. Currently, there is an urgent need to develop new technologies for providing QoS differentiation and guarantees in WiMAX networks. Among the most important technical issues that need to be resolved, there is packet scheduling in WiMAX networks. In this sub-section, we assess proposed scheduling algorithms for QoS support in WiMAX networks thoroughly with respect to the characteristics of the IEEE 802.16 MAC layer and PHY layer [6,7,8]. With respect to the nature of scheduling, algorithm mechanism can be classified as per-flow, per-class and per-packet scheduling algorithms. Representative schemes in each of these categories will be discussed after.

So far, several proposals have been insightful in the present literature of QoS support for wireless networks. One can find several such as FIFO, RR, CBQ, FQ, DRR, WRR and WFQ. Algorithms implemented and simulated in this architecture is FIFO, Weighted Round Robin, Fair Queuing and Weighted Fair Queuing.

![Figure 1](image-url)

Figure 1 shows the scheme of the QoS management mechanism. At the beginning of transmission, each SS
through a phase of registration connect with the BS. Thus, the BS allocates a CID to each connection and the backup settings according to their service class. Uplink packet processing is handled by the Base Station (BS) through signaling process to the Subscriber Station (SS). Each SS has a scheduler to decide which packet will be sent from the queue to the network in the proper time interval as defined in Uplink Map Message (UL-MAP) sent by BS in subsidies allocated data in BS. Two types of scheduling are localized in the BS: Downlink scheduler determines which packets of the upper layer will be next downlink. This decision is based on the QoS requirements and situation of the queue. Uplink scheduler decides which SS can transmit next uplink, and the number of slots that can use SS. This decision is based on the quality requirements and service demands of bandwidth sent by the SS.

3. Related Works

In the recent years, many proposals were enriching the literature of QoS support for wireless networks. Most of the researches are based on IEEE 802.16 standard. An overview of some of these approaches can be found in [3-8]. They can be divided into categories to define the requirements, such as parameters that indicate quality of service and mechanisms that act over these parameters. In the first category QoS analyzes in WiMAX networks. Rohit and Mohammad [1] studied the quality of service architecture and analyzed parameters that indicate quality of service, such as, throughput, packet loss, average jitter and average delay. They concluded that VOIP traffic can be served with UGS service flow and that the rtPS service flow is designed for applications such as streaming audio and streaming video. Aymen and Louifi [9] work consists in the addition of the QoS classes as well as the QoS management requirements, unicast and contention request opportunities mechanisms, and scheduling algorithms for the UGS, rtPS and BE QoS classes. This work is conducted to show that the behavior of UGS, rtPS, and BE schedulers fits with the QoS specifications of the IEEE 802.16 standard. Another category which is based on the scheduling algorithms, many algorithms are specifically designed for WiMAX. The major aim of the study of Jin-Cherng et al. [2] a description and performance evaluation of scheduling algorithms. They show through simulation the best performance for different queue scheduling in the network. It seems that WFQ are not the best selection in performance, even though WFQ give weight to control bandwidth allocation and priority. If the system needs the best throughput performance in this networks transport environment, RIO scheduling scheme is the best scheduling. As well, Najah et al. [10], Tsai et al. [11] present a performance study of uplink scheduling algorithms in point-to-multipoint WiMAX networks. They concluded that there is no single scheduling scheme that provides the desired performance with respect to all QoS requirements and characteristics of the IEEE 802.16 MAC layer. A scheduling algorithm needs to be selected based on the requirements and traffic profiles of the network.

In our study, we focus our work on the implementation of various scheduling algorithms for IEEE 802.16 wireless MAN using network simulator, where a brief description and performance evaluation to provide the better packet scheduling to various application. In this setting, we implemented architecture able to transport the data belonging to several classes of traffic, to provide better packet scheduling to each application. We study and analyze the behavior of each scheduling algorithms in time and when the number of nodes increases. Then, we evaluate the algorithms using parameters that indicate quality of service such as: average delay, average jitter, packet loss and throughput, for all service classes.

4. Performance Evaluation: Experimental Results

4.1. Module WiMAX Structure Simulated

To analyze QoS in a network it is necessary to study real life scenarios. The network topology described in the Figure 2 is used to simulate the proposed scheduling algorithms in order to evaluate the architecture performance that supports QoS. It consists in two parts: the first part defines the sources or the subscriber stations and the second part defines the base station. The scheduling algorithms show interesting results when they are studied under different mix of traffic. Thus, we have created four sources emitting each one a specific type of traffic. The type of traffic based on UDP protocol uses the CBR applications assigned for the UGS class. The second source is the rtPS class that uses the VBR applications based on the UDP protocol. The third source defines the nrtPS class that uses the FTP applications based on TCP protocol. Finally the fourth source uses the FTP applications based on TCP protocol assigned for the BE class. The simulation time is fixed to 100 seconds to analyze and evaluate each algorithm to guarantee the QoS for all types of applications. During the simulation we increase the number of sources (nodes) to evaluate the behavior of each algorithm also when the network is congested [7,12,13].

![Figure 2. Topology of the simulated network](image)

In the current implementation, our contribution consist in implement some scheduling schemes in WiMAX model to simulate real network and conclude from the simulation results to show QoS requirement. Therefore, we assume that subscriber stations cooperate with different types of traffic service and achieve several scheduling algorithms (Drop Tail, FQ, DRR and WFQ) in NS2. The main simulation parameters are listed in Table 1.
Table 1. Parameter simulation

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency bandwidth</td>
<td>5 MHz</td>
</tr>
<tr>
<td>Modulation Type</td>
<td>OFDM_QPSK</td>
</tr>
<tr>
<td>Length of frame</td>
<td>5 ms</td>
</tr>
<tr>
<td>Time simulation</td>
<td>100s</td>
</tr>
<tr>
<td>Routing Protocol</td>
<td>OSDV</td>
</tr>
<tr>
<td>Mean length of packet</td>
<td>1024 bytes</td>
</tr>
<tr>
<td>Number of nodes</td>
<td>24</td>
</tr>
</tbody>
</table>

4.2. Simulation Results

These simulations analyze the parameters of QoS which are strongly related to the applications performances. The same architecture, the same parameters and the same traffics are used for various scheduling algorithms. The numbers of Subscriber Station (SS) used with each class of service are varied.

4.2.1. The Performance Evaluation of Each Class of Service:

A series of simulations have been done where the nodes number is incremented each time, to see the behavior of the algorithms and its influence on the QoS of each service class. Figure 3 to Figure 14 show the result for QoS parameters (delay, jitter ...) variation for UGS, rtPS, nrtPS, and BE service classes depending on the number of nodes.

The delay for different classes of service rises to the peak, this increase is due to queues saturation. If the number of nodes exceeds 16, the average time decreases, this fall reflects the reduction in the number of packets received after the onset of congestion in queues for rtPS, nrtPS and BE flows, as shown in Figure 3. This simulation presents that average delay for UGS application is guaranteed using the WFQ algorithms, but the minimum delay is obtained using WRR algorithm for all kinds of applications.

Figure 7, 8, 9 and 10 shows that the jitter is minimal when the network is not loaded, increasing the number of nodes causes an increase of inter packet delay. This could be explained by the increase of delay in queues. The average jitter is important for rtPS, nrtPS and BE applications, also the low jitter is guaranteed using the WRR algorithm especially for the UGS application.
Based on Figure 11, 12, 13 and 14, there is a degradation of throughput depending on growth in the number of nodes, but it remains stable for UGS flows. The throughput for certain classes is relatively higher than other classes, because of the sizes of packets generated by the traffic generator and the throughput designed to each application... All presented algorithm have the same behavior to guarantee the throughput so if the network is congested only the UGS flows are not affected. This proposed architecture guarantees all parameters that manage QoS for UGS flows using the WRR algorithm.

4.2.2. The Performance Evaluation of the WRR Algorithm:

These scenarios present the performance evaluation of the WRR algorithm. The average delay depends on network state (number of nodes) is presented in Figure 15. The average delay of UGS flows remains constant, while nrtPS flow delay reaches a peak 1,173 (s) and a slight increase in delay for the rtPS stream. This is due to the increase in numbers of received packets of each flow. As soon as the number of nodes exceeds 16, the average delay decreases, this reflects the reduction in the number of packets received after the onset of congestion in queues (Figure 18). This reduction is the result of a request for throughput reduction at sources. The great number of nodes which represent the network congestion, affects the behavior of jitter in rtPS, nrtPS and BE flows and then a steady state for the UGS flow (Figure 16), this could be explained by the added delay in the queues caused by congestion. The performance in terms of average throughput of rtPS, nrtPS and BE flowsis decreased due to the growth of the number of nodes, but the average throughput of UGS flows remains constant (Figure 17).
The variation of delay for various flows according to time at fixed number of nodes is almost constant and identical. The Figure 19 presents a differentiation in delay between flows. BE delay increases from time $t = 20$ (s), this presents the network behavior in the case of saturation that respects constraints of other service classes using WRR algorithm. This variation starts at the transmission of UGS, rtPS and nrtPS flows. Figure 21 presents that the jitter remains stable for different flows. Figure 20 and 22 describe the network behavior to respect constraints (jitter and delay) using FIFO algorithm.

This paper, presents simulation results of scheduling algorithms in WiMAX network concerning the QoS.
parameter for various techniques. After analyzing the simulation results which shows the variations of delay, throughput, jitter and packet received, FIFO algorithm has the same behavior to respect QoS to all service classes without differentiation of service, also the other algorithm guarantee required constraints of different flows. But this guarantee is limited to the WRR and FQ for not congested networks, so if the number of nodes increases, they not affect gradually the QoS. As well, these results show that the WRR algorithm ensures good quality of services especially for multimedia applications. It maintains low latency for applications with high temporal constraints. Each connection with the associated queue can be treated as a separate session, which the BS has to allocate for each connection based on its QoS requirements is the weight value of the WRR scheduler. Since all the slots are of the same size, there is no need to use complex scheduling disciplines, such as weighted fair queuing (WFQ).

5. Cross-layer Architecture

Generally, the cross-layer maintains a logical queue for each class of service to each connection. Each queue is implemented in a FIFO with head and tail pointers pointing to its head of line (HOL) and tail of line (TOL) packets. When a packet arrives at the system, it is stored in the corresponding queue. The scheduler queue prioritizes all HOL packets, or all eligible HOL packets if a shaper-scheduler is implemented, based on their finish times. It then chooses the packet with the smallest finish time to transmit first. This requires fast sorting or searching operations and it is one of the challenges in designing a packet scheduler. In general, all HOL packets are first stored in the FIFO queue. Only those that are currently eligible can be moved to the scheduler queue. Suppose the scheduler queue selects the HOL packet of queue; it determines the head pointer associated with queue and then reads out the packet using the head pointer.

All recently proposed packet-scheduling algorithms for output-buffered switches that support quality of services (QoS) transmit packets in some priority order, e.g. according to dead-lines, virtual finishing times, eligibility times, or other time stamps that are associated with a packet [2,11,14]. Since maintaining a sorted priority queue introduces significant overhead, much emphasis on QoS scheduler design is put on method to simplify the task of maintaining a priority queue. The two main metrics for measuring the performances of a scheduling algorithm: throughput and delay. In our architecture we use fast and scalable pipelined priority queue architecture for use in high-performance switches with support for fine-grained quality of services guarantees. In the cross layer each down link is maintained using an output port manager. The output port manager implements mechanisms that support QoS such as buffer management and scheduling [10,15,16].

5.1. Implementation of the Architecture

The architecture design of the cross layer includes the following blocks: a classifier unit, a cross-layer manager and a queue controller as depicted in Figure 23 [17,18,19].

5.1.1. Classifier Module (CM)

This unit classify packets according to their destination and their class of service required to the application (Figure 24). When a packet is received this unit determines its MS and then determines the corresponding FIFO reserved to the class of service that stores the packets. A FIFO is reserved to each class of service in order to guarantee the QoS to every application. The FIFO selection is based on QoS parameters to every packet.

5.1.2. Cross-layer Manager Module (CLMM)

The design of the cross-layer manager unit uses 4 separated FIFO memories, one FIFO per class of service for every active MS to track resource demands. This unit execute the algorithm presented in Figure 25 to write/read a packet from the corresponding queue. The write operation is controlled by the classifier unit, and the read operation is controlled by the queue controller queue. When the cross-layer manager receives a request signal, it
tests if the request is received from the classifier unit. Then, if the request is from a new connection the cross-layer create a new MS and manages a write operation to store the packet in the selected FIFO. If the request is received from the queue controller, so the cross-layer unit manage a read operation to send the packet. If it is the last packet this unit release the MS. While the queue state is updated when a queue or de-queue operation is executed and the MS state is updated if it is necessary [20,21,22].

5.1.3. Queue Controller Module (QCM)

Having the information on resource demands and the QoS requirements, the BS can make a scheduling decision that results in allocated slots in the downlink and uplink directions. This module is used to control all packets related to the cross-layer management and to guarantee the QoS. The design of the queue controller module implements the algorithm described in Figure 24 which control the QoS parameters (bandwidth, delay, gigue). The queue controller chooses packets to be send based in the scheme of the WRR scheduling algorithm. The WRR weight value only specifies the number of packets to send from each queue during around. The most important one is that WRR behaves work-conservingly skipping empty queues and starting to serve the next queue once all the packets from the current one have been sent.

5.2. Design Results of the Proposed Architecture

A description in VHDL is carried out at the RTL level and is simulated using Model Sim simulator to implement and validate our architecture. This approach has permitted to evaluate the behavior of each component alone as well as the interaction of the overall architecture with all interconnected components. The design is simulated and synthesized using ISE design foundation. Simulation results are presented in Figure 26 shows that the received data in signal din_nod is stored in selected FIFO. The classifier activates signals wr1_11, then wr2_22 and finally wr3_33. So, packets are stored in FIFO01 then FIFO02 and FIFO03. In parallel the WRR algorithm selects FIFO 1, then 2 and 3, so the signal rd1_11, then the signal rd2_22 and finally the signal rd3_33 are activated to send data out based on the WRR scheme. Simulation chronograms show also that the architecture receives and sends data packets in parallel manner.

Figure 26. Simulation results of the architecture using Model Sim

6. Conclusion

In this paper, we present the design of a cross-layer QoS based scheduling algorithm WRR in Wimax Base stations. Firstly, we have studied the performance evaluation of scheduling algorithms in Wimax network using the NS2 simulator. These simulations describe the behavior of each algorithm to respect the QoS of each service class. The studied parameters have a direct influence on the flows nature such as the delay, throughput, jitter and loss rate. We conclude that WRR algorithm guarantee the QoS of each service class in Wimax networks. This algorithm provides required parameters to each application also when the network is congested. The WRR algorithm is easy to be implemented in hardware thus gaining a performance advantage over software-based alternatives. Finally, we present the design results of the cross layer implementation between the network layer and MAC of the base station. This cross layer implements the scheduling algorithm which guarantees the QoS to all Wimax applications.

References


