Proceedings of the World Congress on Engineering and Computer Science 2009 Vol I
WCECS 2009, October 20-22, 2009, San Francisco, USA

A Packet Scheduling To Enhance Quality of Service in IEEE 802.16

Chirayu Nagaraju and Mahasweta Sarkar, *Member IEEE*

Abstract --- **Wireless Broadband Network like the standardized IEEE 802.16, popularly referred to as WiMAX is known for its differentiation of Quality-of-Service (QoS) feature based on the traffic requirement. However the QoS differentiation process is often unfair especially to traffic classes which do not require high QoS. Such traffic classes are often starved to the extent that they experience extremely low throughput. In this paper, we look into a scheduling algorithm for providing QoS in Wireless Broadband Networks which alleviates this unfairness. The algorithm implements two queuing theory principles - Earliest Deadline First (EDF) and Weighted Fair Queuing (WFQ) to assure QoS to various classes of traffic. We propose an enhancement to the EDF queuing principle to ensure that lower priority traffic do not starve. We simulated the algorithm in dot net framework using the programming language C#. Our simulation results show considerable improvement in the throughput of low priority traffic while retaining the packet delay constraint of the high priority traffic classes.**

Keywords -- **WiMAX, Quality of Service, EDF, WFQ, Scheduling**

I. INTRODUCTION.

Worldwide Interoperability for Microwave Access or WiMAX is a wireless communication technology based on IEEE 802.16 standard. It is a data-on-the-go alternative to cable and DSL and a standards-based broadband wireless access technology for enabling the last-mile delivery of information.

Earliest Deadline First (EDF) is one the most widely used scheduling algorithms in WiMAX for real-time applications as it selects Subscriber Station (SSs) based on their delay requirements. The algorithm assigns deadline to arriving packets of a SS.

Since each SS specifies a value for the maximum latency parameter, the arrival time of a packet is added to the latency to form the tag of the packet [2],

The information module needs to find the real time polling service (rtPS) deadline information. Based on this deadline information, the Uplink Packet Scheduling (UPS) will know exactly when to schedule packets such that packets' delay requirements are met.

This work has been supported by the National Science Foundation grant award DUE-737048

Chirayu Nagaraju is with the Department of Electrical and Computer Engineering San Diego State University, San Diego
(e-mail: chirayunraj@yahoo.com).

Mahasweta Sarkar is now with the Department of Electrical and Computer Engineering San Diego State University, San Diego, CA, USA
(e-mail: msarkar2@mail.sdsu.edu).

In this paper we see, while providing for rtPS packets making use of EDF, we see to it that portion of the Best Effort (BE) packet is also transmitted along with the rtPS packets (and not starved as usually is the case). As a result we see an improvement in traffic throughput for Best Effort traffic type. The remainder of the paper is structured as follows. In section II we look into a brief overview on WiMAX. In section III we discuss our scheduling scheme. We look into our simulation setup and results in section IV. Finally we conclude the paper in section V.

II. OVERVIEW ON WiMAX.

WiMAX is a wireless communication technology based on IEEE 802.16 standard. WiMAX provides fixed, nomadic portable and mobile wireless broadband connectivity without the need for direct line-of-sight connection between a base station and a subscriber station. The physical layer operates at 10-66 GHz (IEEE 802.16) and 2-11 GHz (IEEE 802.16a) with data rates of 32-130 Mbps depending on the width of the channel frequency and modulation technique. WiMAX uses either of two types of duplex methods: Time division duplex (TDD) and Frequency division duplex (FDD).

A. WiMAX Architecture

IEEE 802.16 architecture consists of two kinds of fixed (non-mobile) stations: subscriber stations (SS) and a base station (BS). The BS regulates all the communication in the network. Thus there is no peer-to-peer communication between the SSs. The communication path between SS and BS has two directions: uplink (from SS to BS) and downlink (from BS to SS). When the system uses time-division multiplexing (TDM) for uplink and downlink transmissions, the frame is subdivided into an uplink sub frame and a downlink sub frame (Figure 1). The duration of these sub frames is dynamically determined by the BS. Each sub frame consists of a number of time slots. SSs and BS have to be synchronized in time and transmit data into predetermined time slots. Reference [1] suggests that IEEE 802.16 can support multiple communication services (data, voice, and video) with different QoS requirements. On the downlink (from BS to SS), the transmission is relatively simple because the BS is the only one that transmits during the downlink sub frame. The data packets are broadcasted to all SSs and an SS only picks up the packets destined for it. For uplink traffic, BS determines the number of time slots that each SS will be allowed to transmit in an uplink sub frame.

This information is broadcasted by the BS through the Uplink Map message (UL-MAP) at the beginning of each frame. UL-MAP contains Information Element (IE) which includes the transmission opportunities, i.e. the time slots in which the SS can transmit during the uplink sub frame. After receiving the UL-MAP, each SS will transmit data in the pre-defined time slots as indicated in IE. The BS uplink scheduling module determines the IEs using bandwidth request PDU (BW-request) sent from SSs to BS. In IEEE 802.16 standard, there are two modes of transmitting the BW-Request: contention mode and contention-free mode (polling). In contention mode, SSs send BW-Request during the contention period. Contention is resolved using back-off resolution. In contention-free mode, BS polls each SS and SSs reply by sending BW-request. Due to the predictable signaling delay of the polling scheme, contention-free mode is suitable for real time applications [1], [2].

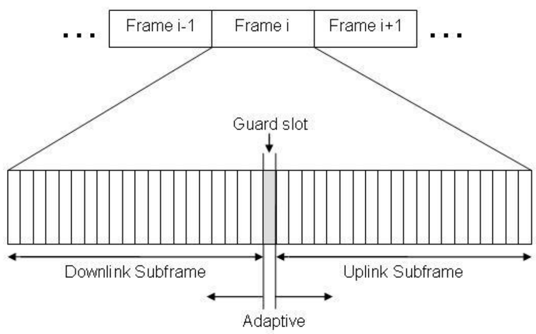

Figure 1: TDD Frame Structure.

B. QoS in WiMAX

IEEE 802.16 defines four types of service flows and each one with different QoS requirements:

1) Unsolicited Grant Service (UGS) - Designed for services which require constant bit rate. E.g.: T1/E1.
2) Real-Time Polling Service (rtPS) - designed to support real-time services which generate variable size data packets but also must meet the delay constraints imposed by application. E.g.: MPEG.
3) Non-Real Time Polling Services (nrtPS) Designed for services which require good data rate transfer performance support but are tolerant to delay. E.g.: FTP
4) Best Effort Service (BE) : Designed for services which do not require guarantee in QoS .E.g. HTTP.

As depicted in Figure 2, IEEE 802.16 QoS Architecture includes:

1) Applications which initiate at SS (subscriber station) send a connection request to BS (Base Station) with its QoS requirements.

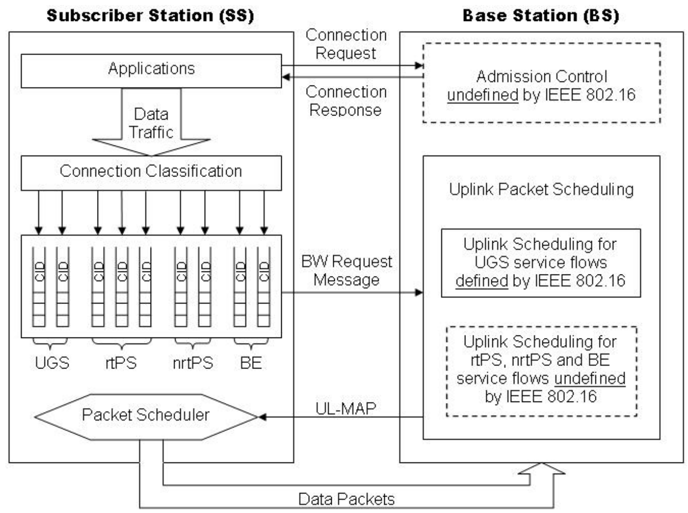

Figure 2: WiMAX Traffic Classification

2) Admission Control is a module at the BS which verifies if this new connection can be supported for its QoS requirements. This module is left undefined by IEEE 802.16
3) If BS accepts an application request, the application starts streaming packets to SS, which are mapped into one of the existing classes (UGS, rtPS, nrtPS, BE) by Connection Classifier module at SS and put into an appropriate queue
4) Based on the current queue size of each class, SS makes a BW (Bandwidth) request to BS.
5) On receiving BW request from all SSs, BS allocates bandwidth to each SS based on the queue size of each SS and their QoS requirements. The scheduling is defined in IEEE 802.16 only for UGS class. Scheduling for other classes is left undefined.
6) After scheduling BW among all SSs, BS broadcasts this information to all SS in a data structure called the Uplink Map (UL-MAP).
7) Based on the UL-MAP, each SS will transmit packets according to the BW allocated to them [1].

III. THE SCHEDULING SCHEME.

Our scheduling scheme is based on a scheme proposed in [1] to which we have added substantial enhancements. In this section we first briefly present the basic scheme as outlined in [1]. We then describe our enhancements to that scheme.

A. Uplink Packet Scheduling

We look into a combination of both Earliest Deadline First (EDF) & Weighted Fair Queue (WFQ) technique to allocate bandwidth among SSs. For Unsolicited Grant Service (UGS) connections, the standard has specified to allocate fixed bandwidth as required by the connection. For rtPS, EDF technique is used which schedules packets with earliest deadlines first. For nrtPS class, WFQ technique and finally whatever remaining bandwidth is available is to be equally allocated to the different applications of the BE traffic class.

The algorithm has been divided into three modules - Information Module, Scheduling Database Module & Service Assignment Module. All these three modules work in the BS. In information module as designed in [1], for an rtPS connection, the number of packets generated in slot (k+1) at the service station is given by (1):

$$q(k+1) - [q(k) - s(k)] \qquad (1)$$

Where $q(k+1)$ is the queue size at frame$(k+1)$, $q(k)$ is the queue size at frame k and $s(k)$ is the number of packets served at frame k. The queue size is obtained at the beginning of each time slot which includes the packets generated during the previous time slot. This formula computes correctly the number of packets generated at frame k, as long as there were no dropped packets in that frame for that queue. For example, if $q(k) = 7$, $q(k+1) = 8$, $s(k) = 4$, then the number of packets generated as given by their formula is $8 - 7 + 4 = 5$, which is correct, since after serving 4 packets, the queue size would be equal to 3 (i.e. $7 - 4$), and if current queue size is 8, then the number of packets generated would be equal to 5 (i.e. $8 - 3$).

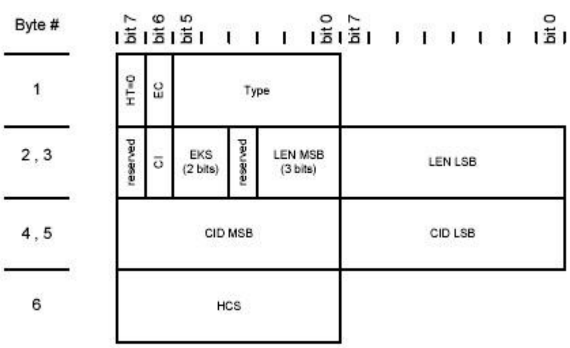

Figure 3: Generic MAC header for the generic MPDU

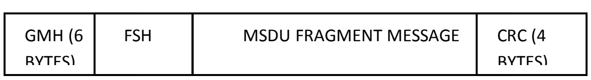

Figure 4: A fragment sub-header is added at the start of a payload to tell the system to divide the MSDU into fragments.

On the other hand, if say 1 packet was dropped as it missed its deadline, then the queue size after transmitting 4 packets & dropping 1 packet would be 2 (i.e. $7 - 4 - 1$). In this case the number of packets generated would be equal to 6 (i.e $8 - 2$), but using the above formula we would still get 5 packets as generated, which is incorrect. Hence we modify the formula, as shown below in (2). Thus the number of packets generated in slot (k+1) is given by:

$$q(k+1) - [q(k) - s(k) - d(k) \qquad (2)$$

where $d(k)$ represents the number of packets dropped in frame k. Applying this formula in the above example, we get $8 - 7 + 4 + 1 = 6$, which is correct.

At this point, it is crucial to understand in brief the basic functionalities of the medium access control (MAC) layer.

In an 802.16 system, the MAC communicates using MAC protocol data units (MPDUs) that are carried by the PHY. The generic MAC header (GMH) contains details of the MPDU. Each of the generic MPDU begins with a GMH [4].

An MSDU may be divided into fragments that are transmitted independently. To signal this, a fragment sub-header (FSH) is included at the start of the payload.

As discussed above a GMH is appended to every payload from the above layer to form a MAC level frame which is required to be transmitted through a wireless channel. Each sub frame (uplink and downlink) in the WiMAX architecture consists of a number of time slots.

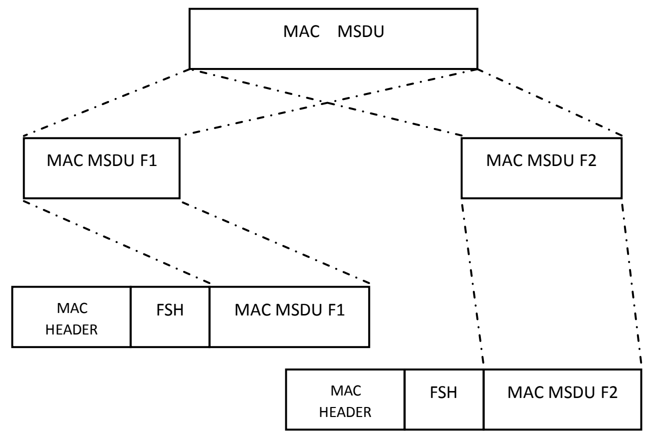

Figure 5: Fragmentation into sub-frames.

SSs and BS have to be synchronized and transmit data into predetermined fixed size time slots. Thus if the original payload with the MAC header is too big to be transmitted in a single time slot then the payload is fragmented and combined with MAC header and fragmented sub-header to form a MAC level sub frame which can be transmitted Within the duration of a time slot as depicted in Figure 5.

B. The Earliest Deadline First (EDF)Algorithm

We observe that the EDF algorithm tends to starve the BE traffic class if the number of rtPS packets are high. Hence we propose a solution to service rtPS class using EDF, but at the same time ensuring that BE traffic class is not starved. Since each frame is divided into time slots, and SS are required to transmit packets in these slots, the original packets generated at the application level are fragmented to ensure that these packets fit into and can be transmitted in a time slot. When we fragment a packet, the last fragmented packet might be of any length from 1 byte to the maximum size which can be transmitted in a slot. If it so happens that the last fragment contains lesser number of bytes than the maximum allowable fragment size, then we can stuff a part of a BE packet into this empty section. In this way, two or three such empty slots might be enough to transmit a complete BE packet to the BS. The BS buffers these bytes and awaits the reception of the remaining bytes to completely reconstruct the packet. This proposed solution does introduce some additional overheads. The additional overheads and our justification for each are listed as follows:

A. **Additional buffer at BS:** BS needs to buffer small parts of BE data until it receives a complete BE packet. Though this requires extra memory at BS, the total extra memory required might not be too much as BS only needs one buffer which is as big as a BE packet size, for each SS in the network.

b. **Additional computation at BS & SS:** Since a BE packet now needs to be fragmented at SS and put into an rtPS packet, and then defragmented and assembled back to a complete BE packet at the BS, it definitely requires extra computation effort both at the SS and the BS. Since these extra computation is only required at the SS and the BS and not on the end nodes, it is not too much of an overhead as both BS and SS can be supplied with abundant resources like power, processors etc.

c. **Extra header information:** Extra header information is required to indicate if a current packet is a hybrid packet (i.e. it contains both rtPS and BE data bytes), if yes then how much BE data is stuffed into it, and the offset of it. We propose a solution which would reduce the extra space required on each packet to just 1 bit. Every packet would need to include 1 bit in the header to indicate if it is a hybrid packet or a complete rtPS packet. If it is a hybrid packet, then we would put the extra header information in the empty section of the packet. This can be done by putting the header information at the bottom of the packet as shown in Figure 6.

Thus the extra header information required would be accommodated in the empty section of the packet, eliminating the need to reserve space for this in the header of every packet. When the frames are too big to be transmitted in a single time slot, the frames are fragmented and FSH is appended to the fragmented part and then transmitted. FSH (fragment sequence header) follows right after the MAC header, as seen in Figure 6. FSH has 3 reserved bits. We make use of one of the reserved bit to indicate that the packet is a hybrid one, eliminating the need of introducing additional bit spaces. The BE packet portion which is also fragmented has its own header information which includes the length of the data being transmitted (11 bits), offset of the Best effort data in packet (11 bits), FSN (Fragment Sequence Number) (11 bits) and FC (Fragment Control) (2 bits)present in all FSH.

IV. SETUP'S & SIMULATION RESULTS.

Some of the parameters used to run the simulation.

i) Data rate: number of bits per second.
ii) Frame length: total number of slots (uplink and downlink subframe)
iii) time slots: time duration of each slots in a frame, .

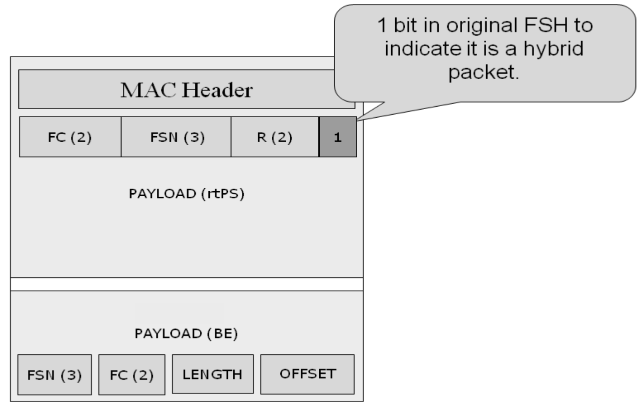

Figure 6: Hybrid packet

iv) frame time: product of frame length and time slots,
v) traffic rate: number of packets of each of the services like (rtPS, nrtPS, BE) per frame,
vi) packet size- number of bytes in the packet, Values of the parameters considered can be changed.

We have considered some of the parameter to have fixed value, which can also be changed initially before the simulation could be conducted.

i) maximum downlink slots – 50 slots
ii) contention slots – 20 slots
iii) BE packets overhead - 100 bytes
iv) frame length considered - 125 slots
v) data rate – 40Mb/s
vi) Maximum bytes per packets – (time slots*data rate) / 8000.

While simulating we have setup two scenarios, Long duration and the short duration with different traffic condition. During long duration the simulation time considered is 10000000 ms, whilst the short duration we considered 10000 ms. The traffic condition chosen for each of scenarios vary, when its high traffic its maximum of 10 (rtPS, nrtPS, BE) packets per frame, when its low traffic it's a minimum of 2 (rtPS, nrtPS, BE) packets per frame.

We see in (figure 7) the obtained graph is only showing rtPS and nrtPS services, in this scenario enhanced version is not having significant improvement.

In (figure 8), we have considered a long duration of high traffic data. Since there is high traffic, our scheduling scheme comes into play and we see a significant improvement in the Enhanced (BE) as shown in the graph. In this scenario along

ISBN:978-988-17012-6-8

with BE, rtPS and nrtPS services are also simulated and BE is plotted as shown.

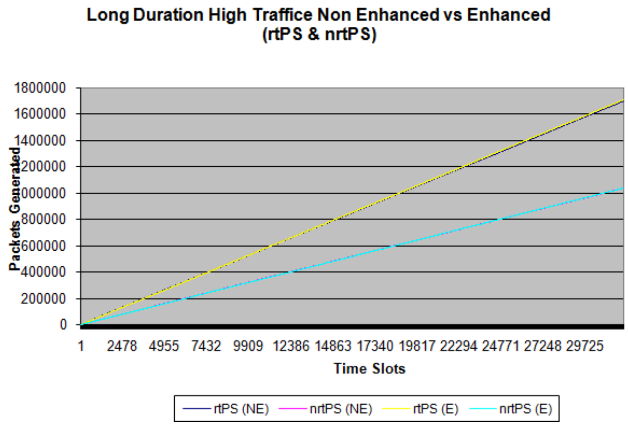

Figure 7: Long Duration High Traffic Non- Enhanced Vs. Enhanced (rtPS, nrtPS).

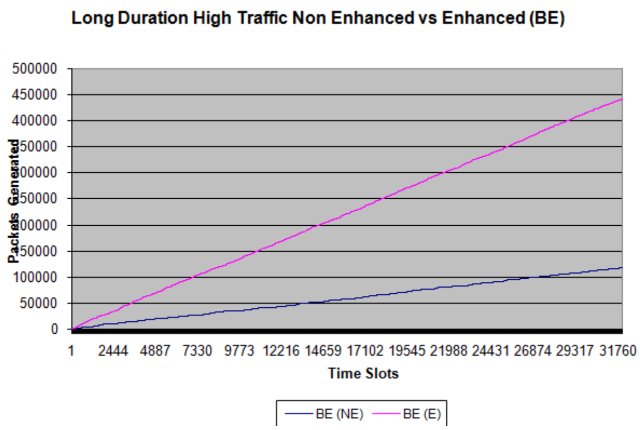

Figure 8: Long Duration High Traffic Non- Enhanced Vs. Enhanced (BE).

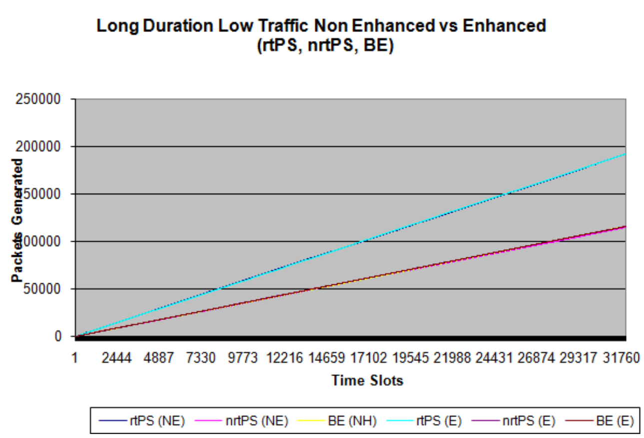

Figure 9: Long Duration Low Traffic Non Enhanced Vs. Enhanced (rtPS, nrtPS, BE).

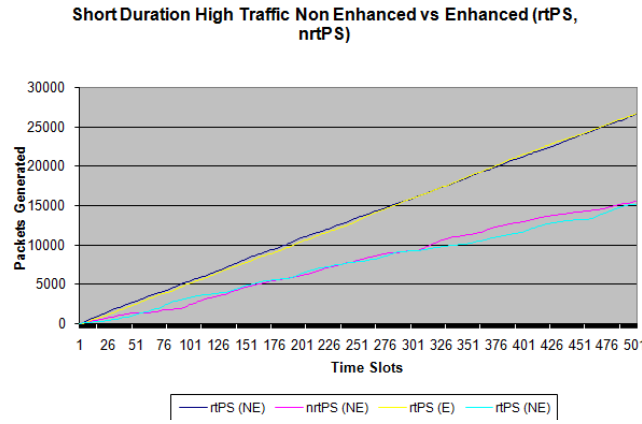

Figure 10: Short Duration High Traffic Non- Enhanced Vs. Enhanced (rtPs, nrtPS).

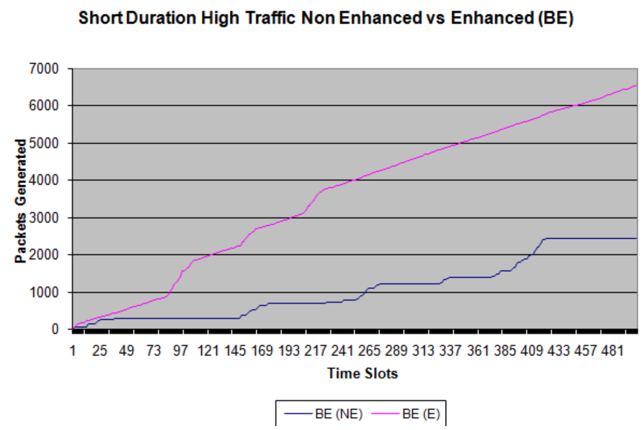

Figure 11: Short Duration High Traffic Non- Enhanced Vs. Enhanced (BE).

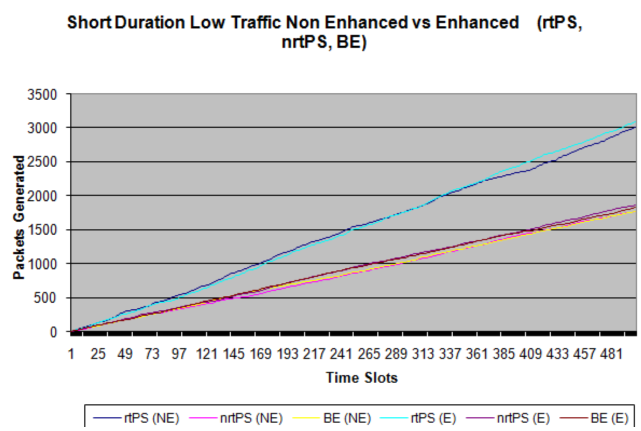

Figure 12: Short Duration Low Traffic Non- Enhanced Vs. Enhanced (rtPS, nrtPS, BE).

In (figure 11) , we see that the traffic is high but for a short duration, we see a significant improvement in the Enhanced (BE) as shown in the graph, In this scenario BE, rtPS and nrtPS services are also simulated and BE is plotted as shown.

V. CONCLUSION

We looked into various scheduling algorithms proposed as part of providing QoS in WiMAX. We studied a combination of EDF + WFQ algorithm in detail and simulated it in dot net using C#. This algorithm starved BE class traffic in presence of rtPS & nrtPS classes and hence we proposed a unique way of assuring that BE packets are also transmitted. More the rtPS traffic, more are the chance that BE packets will find empty spaces to be transmitted. If rtPS traffic is less, then BE packets anyways are transmitted as seen in simulation results. With our proposed enhancements, we achieved significant improvements in the BE traffic class when traffic of rtPS class is high.

REFERENCES

[1] K. Wongthavarawat, and A. Ganz, "Packet scheduling for QoS support in IEEE 802.16 broadband wireless access systems", *International Journal of Communication Systems*, vol. 16, issue 1, pp. 81-96, February 2003

[2] Pratik Drona, "A Performance Study of Uplink Scheduling Algorithms in Point to Multipoint WiMAX Networks"

[3] T.Tsai, C.Jiang and C.Wang, "CAC and Packet Scheduling Using Token Bucket for IEEE 802.16 Networks", *Journal of Communications*, vol.1, no2., pp.30-37, May 2006.

[4] .Eklund, R.B.Marks, S.Ponnuswamy, K.L.Stanwood, N.J.M. van Waes, WirelessMAN, IEEE Standards Wireless Networks Series, IEEE Press.

[5] C.Cicconetti, A.Erta, L.Lenzini and E.Mingozzi, "Performance Evaluation of the IEEE 802.16 MAC for QoS Support", *IEEE Transactions on Mobile Computing*, vol.6, no.1, pp.26-38, January 2007

[6] N.Ruangchaijatupon, L.Wang and Y.Ji, "A Study on the Performance of Scheduling Schemes for Broadband Wireless Access Networks", *Proceedings of International Symposium on Communications and Information Technology*, pp. 1008-1012, October 2006

[7] M.Settembre, M.Puleri, S.Garritano, P.Testa, R.Albanese, M.Mancini and V.Lo Curto, "Performance analysis of an efficient packet-based IEEE 802.16 MAC supporting adaptive modulation and coding", *Proceedings of International Symposium on Computer Networks*, pp.11-16, June 2006.

[8] D.Niyato and E.Hossain, "A Queuing-Theoretic Optimization-Based Model for Radio Resource Management in IEEE 802.16 Broadband Wireless Networks", *IEEE Transactions on Computers*, vol.55, no.11, pp. 1473-1488, November 2000