Comparative Study Of Different Scheduling Algorithms For Wimax MAC Scheduler Design

Gajanan M. Galshetwar*, Prof. Amutha Jayakumar **, Yashika Mittal*

*(M.Tech.(Electronics)

Department of Electrical Engineering, VJTI, Matunga, Mumbai.

** (Professor, Department of Electrical Engineering, VJTI, Matunga, Mumbai.

Abstract

Increased user mobility and the need for data all the time has lead to increase the interest in Broadband Wireless Access (BWA). Unlike wireless LANs, WiMAX networks incorporate several quality of service (QoS) mechanisms at Media Access Control (MAC) level for guaranteed service of voice, data and video. There are some Design factors a designer must consider in order to provide guaranteed QoS. This paper discusses different design factors that are essential for designing a MAC Scheduler and comparison of different algorithms based on these design factors.

Keywords - WiMAX, QOS, UGS (Unsolicited grant service), rtps (real time polling service),ertps (extended real time polling service)

I. INTRODUCTION

IEEE 802.16 is a set of telecommunications technology standards aimed at providing wireless access over long distances in a variety of ways - from point-to-point links to full mobile cellular type acces. It covers a metropolitan area of several kilometers and is also called WirelessMAN. Theoretically, a WiMAX base station can provide broadband wireless access in range up to 30 miles (50 kms) for fixed stations and 3 to 10 miles (5 to 15 kms) for mobile stations with a maximum data rate of up to 70 Mbps compared to 802.11a with 54 Mbps up to several hundred meters, EDGE (Enhanced Data Rates for Global Evolution) with 384 kbps to a few kms, or CDMA2000 (Code-Division Multiple Access 2000) with 2 Mbps for a few kms.

IEEE 802.16 standards group has been developing a set of standards for broadband (high-speed) wireless access (BWA) in a metropolitan area. Since 2001, a number of variants of these standards have been issued and are still being developed. Like any other standards, these specifications are also a compromise of several competing proposals and contain numerous optional features and mechanisms. The Worldwide Interoperability for Microwave Access Forum or WiMAX Forum is a group of 400+ networking equipment vendors, service providers, component manufacturers and users that decide which of the numerous options allowed in the IEEE 802.16 standards should be implemented so that equipment from different vendors will inter-operate. Several features such as unlicensed band operation, 60 GHz operation, while specified in the IEEE 802.16 are not a part of WiMAX networks since it is not currently in the profiles agreed at the WiMAX Forum. For an equipment to be certified as WiMAX compliant, the equipment has to pass the inter-operability tests specified by the WiMAX Forum. For the rest of this paper, the terms WiMAX and the IEEE 802.16 are used interchangeably.

II. DESIGN FACTORS

To decide which queue to service and how much data to transmit, one can use a very simple scheduling technique such as First In First out (FIFO). This technique is very simple but unfair. A little more complicated scheduling technique is Round Robin (RR). This technique provides the fairness among the users but it may not meet the QoS requirements. Also, the definition of fairness is questionable if the packet size is variable. In this section, we describe the factors that the scheduler designers need to consider.

1. QoS Parameter:

The first factor is whether the scheduler can assure the QoS requirements for various service classes. The main parameters are the minimum reserved traffic, the maximum allowable delay and the tolerated jitters. For example, the scheduler may need to reschedule or interleave packets in order to meet the delay and throughput requirements. Earliest Deadline First (EDF) is an example of a technique used to guarantee the delay requirement. Similarly, Largest Weighted Delay First (LWDF) has been used to guarantee the minimum throughput.

2. Throughput Optimization

Since the resources in wireless networks are limited, another important consideration is how to maximize the total system throughput. The metrics here could be the maximum number of supported MSs or whether the link is fully utilized. One of the best ways to represent throughput is using the goodput, which is the actual transmitted data not including the overhead and lost packets. The overheads include MAC overhead, fragmentation and packing overheads and burst overhead. This leads to the discussion of how to optimize the number of bursts per frame and how to pack or fragment the SDUs into MPDUs. The bandwidth request is indicated in number of bytes. This does not translate straight forwardly to number of slots since one slot can contain different number of bytes depending upon the modulation technique used. For example with Quadrature Phase-Shift Keying 1/2 (QPSK1/2), the number of bits per symbol is 1. Together with PUSC at 10 MHz system bandwidth and 1024 Fast Fourier transform (FFT),that leads to 6 bytes per slot. If the MS asks for 7 bytes, the BS needs to give 2 slots thereby consuming 12 bytes. Moreover, the percentage of packet lost is also important. The scheduler needs to use the channel state condition information and the resulting bit error rate in deciding the modulation and coding scheme for each user.

3. Fairness

Aside from assuring the QoS requirements, the left-over resources should be allocated fairly. The time to converge to fairness is important since the fairness can be defined as short term or long term. The short-term fairness implies long term fairness but not vice versa.

4. Energy Consumption and Power Control

The scheduler needs to consider the maximum power allowable. Given the Bit Error Rate (BER) and Signal to Noise Ratio (SNR) that the BS can accept for transmitted data; the scheduler can calculate the suitable power to use for each MS depending upon their location. For mobile users, the power is very limited. Therefore, MS scheduler also needs to optimize the transmission power.

5. Implementation Complexity

Since the BS has to handle many simultaneous connections and decisions have to be made within 5 ms WiMAX frame duration, the scheduling algorithms have to be simple, fast and use minimum resources such as memory. The same applies to the scheduler at the MS.

6. Scalability

The algorithm should efficiently operate as the number of connections increases.

III. CLSSIFICATION OF SCHEDULARS

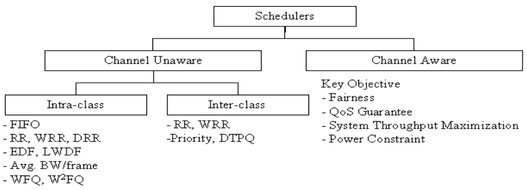

Fig. 1 Classification Of Wimax Schedulers

In this section, we present a survey of recent scheduler proposals for WiMAX. Most of these proposals focus on the scheduler at BS, especially DL-BS (Downlink Base station) scheduler. For this scheduler, the queue length and packet size information are easily available. To guarantee the QoS for MS at UL-BS(Uplink Base station) scheduler, the polling mechanism is involved. Once the QoS can be assured, how to split the allocated bandwidth among the connections depends on the MS scheduler. Recently published scheduling techniques for WiMAX can be classified into two main categories: channel-unaware schedulers and channel-aware schedulers as shown in Fig.1. Basically, the channel-unaware schedulers use no information of the channel state condition in making the scheduling decision. Channel-unaware schedulers generally assume error-free channel since it makes it easier to prove assurance of QoS. However, in wireless environment where there is a high variability of radio link such as signal attenuation, fading, interference and noise, the channel-awareness is important. Ideally, scheduler designers should take into account the channel condition in order to optimally and efficiently make the allocation decision.

A. Channel-Unaware Schedulers

This type of schedulers makes no use of channel state conditions such as the power level and channel error and loss rates. These basically assure the QoS requirements among five classes - mainly the delay and throughput constraints. Although, jitter is also one of the QoS parameters, so far none of the published algorithms can guarantee jitter.

1) Intra-class Scheduling:

Intra-class scheduling is used to allocate the resource within the same class given the QoS requirements.

a. Round Robin (RR) algorithm: Aside from FIFO, roundrobin allocation can be considered the very first simple scheduling algorithm. RR fairly assigns the allocation one by one to all connections. The fairness considerations need to include whether allocation is for a given number of packets or a given number of bytes. With packet based allocation, stations with larger packets have an unfair advantage. Moreover, RR may be non-work conserving in the sense that the allocation is still made for connections that may have nothing to transmit. Therefore, some modifications need to be made to skip the idle connections and allocate only to active connections. However, now the issues become how to calculate average data rate or minimum reserved traffic at any given time and how to allow for the possibility that an idle connection later has more traffic than average? Another issue is what should be the duration of fairness? For example, to achieve the same average data rate, the scheduler can allocate 100 bytes every frame for 10 frames or 1000 bytes every 10th frame. Since RR cannot assure QoS for different service classes, RR with weight, Weighted Round

Robin (WRR), has been applied for WiMAX scheduling. The weights can be used to adjust for the throughput and delay requirements. Basically the weights are in terms of queue length and packet delay or the number of slots. The weights are dynamically changed over time. In order to avoid the issue of missed opportunities, variants of RR such as Deficit Round Robin (DRR) or Deficit Weighted Round Robin (DWRR) can be used for the variable size packets. The main advantage of these variations of RR is their simplicity. The complexity is O(1) compared to O(log(N)) and O(N) for other fair queuing algorithms. Here, N is the number of queues.

b. Weighted Fair Queuing algorithm (WFQ):

WFQ is an approximation of General Processor Sharing (GPS). WFQ does not make the assumption of infinitesimal packet size. Basically, each connection has its own FIFO queue and the weight can be dynamically assigned for each queue. The resources are shared in proportion of the weight. For data packets in wired networks with leaky bucket, an end-to-end delay bound can be provably guaranteed. With the dynamic change of weight, WFQ can be also used to guarantee the data rate. The main disadvantage of WFQ is the complexity, which could be O(N). To keep the delay bound and to achieve worst-case fairness property, a slight modification of the WFQ, Worst-case fair Weighted Fair Queueing (WF2Q) was introduced. Similar to WFQ, WF2Q uses a virtual time concept. The virtual finish time is the time GPS would have finished sending the packet. WF2Q, looks for the packet with the smallest virtual finishing time and whose virtual start time has already occurred instead of searching for the smallest virtual finishing time of all packets in the queue. The virtual start time is the time GPS starts to send the packet.

In achieving the QoS assurance, procedure to calculate the weight plays an important role. The weights can be based on several parameters. Aside from queue length and packet delay we mentioned above, the size of bandwidth request can be used to determine the weight of queue (the larger the size, the more the bandwidth). The ratio of a connection's average data rate to the total average data rate can be used to determine the weight of the connection. The minimum reserved rate can be used as the weight. The pricing can be also used as a weight. Here, the goal is to maximize service provider revenue.

c. Delay-based algorithms: **This set of schemes is** specifically designed for real-time traffic such as UGS, ertPS and rtPS service classes, for which the delay bound is the primary QoS parameter and basically the packets with unacceptable delays are discarded. Earliest Deadline First (EDF) is the basic algorithm for scheduler to serve the connection based on the deadline. Largest Weighted Delay First (LWDF) chooses the packet with the largest delay to avoid missing its deadline. Delay Threshold Priority Queuing (DTPQ) was proposed for use when both real-time and non real-time traffic are present. A simple solution would be to assign higher priority to real-time traffic but that could harm the non realtime traffic. Therefore, urgency of the real-time traffic is taken into account only when the head-of-line (HOL) packet delay exceeds a given delay threshold. This scheme is based on the tradeoff of the packet loss rate performance of rtPS with average data throughput of nrtPS with a fixed data rate. Rather than fixing the delay an adaptive delay threshold-based priority queuing scheme is used which takes both the urgency and channel state condition for real-time users adaptively into consideration. Variants of RRs, WFQs and delay based algorithms can resolve some of the QoS requirements. However, there are no published papers considering the tolerated delay jitter in the context of WiMAX networks. Especially for UGS and ertPS, the simple idea is to introduce a zero delay jitter by the fragmentation mechanism. Basically, BS transfers the last fragmented packet at the end of period. However, this fragmentation increases the overhead and also requires fixed buffer size for two periods. Compared to EDF, this simple technique may require more bursts. This needs to be investigated further.

2) Inter-class Scheduling: As shown in Fig. 1,

RR, WRR and priority-based mechanism have been applied for interclass scheduling in the context of WiMAX networks. The main issue for inter-class is whether each traffic class should be considered separately, that is, have its own queue. For example, in rtPS and nrtPS are put into a single queue and moved to the UGS (highest priority) queue once the packets approach their deadline. Similarly in UGS, rtPS and ertPS queues are combined to reduce the complexity. Another issue here is how to define the weights and/or how much resources each class should be served. There is a loose bound on service guarantees without a proper set of weight values.

a. Priority-based algorithm (PR):

In order to guarantee the QoS to different classes of service, priority-based schemes can be used in a WiMAX scheduler. For example, the priority order can be: UGS, ertPS, rtPS, nrtPS and BE, respectively. Or packets with the largest delay can be considered at the highest priority. Queue length can be also used to set the priority level, e.g., more bandwidth is allocated to connections with longer queues. The direct negative effect of priority is that it may starve some connections of lower priority service classes. The throughput can be lower due to increased number of missed deadlines for the lower service classes' traffic. To mitigate this problem, Deficit Fair Priority Queuing (DFPQ) with a counter was introduced to maintain the maximum allowable bandwidth for each service class. The counter decreases according to the size of the packets. The scheduler moves to another class once the counter falls to zero. DFPQ has also been used for inter-class scheduling. To sum up, since the primary goal of a WiMAX scheduler is to assure the QoS requirements, the scheduler needs to support at least the five basic classes of services with QoS assurance. To ensure this, some proposed algorithms have indirectly applied or modified existing scheduling disciplines for each WiMAX QoS class of services. Each class has its own distinct characteristics such as the hard-bound delay for rtPS and ertPS. Most proposed algorithms have applied some basic algorithms proposed in wired/wireless networks to WiMAX networks such as variations of RR and WFQ. For example, to schedule within a class, RR and WFQ are common approaches for nrtPS and BE(best effort) and EDF for UGS and rtPS. The priority-based algorithm is commonly used for scheduling between the classes. For example, UGS and rtPS are given the same priority which is also the highest priority. Moreover, “two-step scheduler” is a generic name for schedulers that try first to allocate the bandwidth to meet the minimum QoS requirements - basically the throughput in terms of the number of slots or subcarrier and time duration and delay constraints. Then, especially in WiMAX networks (OFDMA-based) in the second step, they consider how to allocate the slots for each connection. This second step of allocating slots and subcarriers is still an open research area. The goal should be to optimize the total goodput, to maintain the fairness, to minimize the power and to optimize delay and jitter.

B. Channel-Aware Schedulers

The scheduling disciplines we discussed so far make no use of the channel state condition. In other words, they assume perfect channel condition, no loss and unlimited power source. However, due to the nature of wireless medium and the user mobility, these assumptions are not valid. For example, a MS may receive allocation but may not be able to transmit successfully due to a high loss rate. In this section, we discuss the use of channel state conditions in scheduling decisions. The channel aware schemes can be classified into four classes based on the primary objective: fairness, QoS guarantee, system throughput maximization, or power optimization. A comparison of the scheduling disciplines is presented in Table II. Basically, the BS downlink scheduler can use the Carrier to Interference and Noise Ratio (CINR) which is reported back from the MS via the CQI channel. For UL scheduling, the CINR is measured directly on previous transmissions from the same MS. Most of the purposed algorithms have the common assumption that the channel condition does not change within the frame period. Also, it is assumed that the channel information is known at both the transmitter and the receiver. In general, schedulers favor the users with better channel quality since to exploit the multiuser diversity and channel fading, the optimal resource allocation is to schedule the user with the best channel or perhaps the scheduler does not allocate any resources for the MS with high error rate because the packets would be dropped anyway. However, the schedulers also need to consider other users’ QoS requirements such as the minimum reserved rate and may need to introduce some compensation mechanisms. The schedulers basically use the property of multi-user diversity in order to increase the system throughput and to support more users. Consider the compensation issue. Unlike the wireless LAN networks, WiMAX users pay for their QoS assurance. Thus, in the argument of what is the level of QoS was brought on due to the question whether the service provider should provide a fixed number of slots. If the user happens to choose a bad location (such as the basement of a building on the edge of the cell) the provider will have to allocate a significant number of slots to provide the same quality of service as a user who is outside and near the base station. Since the providers have no control over the locations of users, they can argue that they will provide the same resources to all users and the throughput observed by the user will depend upon their location. A generalized weighted fairness (GWF) concept, which equalizes a weighted sum of the slots and the bytes. WiMAX equipment manufacturers can implement generalized fairness. The service providers can then set a weight parameter to any desired value and achieve either slot fairness or throughput fairness or some combination of the two. The GWF can be illustrated as an equation below:

$$Total\_Slots = \sum_{i=1}^{N} S_i$$

$$wS_i + \frac{(1-w)B_i}{M} = wS_j + \frac{(1-w)B_j}{M}$$

$$B_i = b_i \times S_i$$

For all subscribers i and j in N. Here, Si and Bi are total number of slots and bytes for subscriber i. bi is the number of bytes per slot for subscriber i. N is the number of active subscribers. M is the highest level MCS size in bytes. w is a general weight parameter. It has been observed that allowing unlimited compensation to meet the QoS requirements may lead to bogus channel information to gain resource allocations. The compensation needs to be taken into account with leading/lagging mechanisms. The scheduler can reallocate the bandwidth left-over either due to a low channel error rate or due to a flow not needing its allocation. It should not take the bandwidth from other well-behaved flows. In case, there is still some left-over bandwidth, the leading flow can also gain the advantage of that left-over. However, another approach can be by taking some portion of the bandwidth from the leading flows to the lagging flows. When the error rate is high, a credit history can be built based on the lagging flows and the scheduler can allocate the bandwidth based on the ratio of their credits to theirs minimum reserved rates when the error rate is acceptable. In either case, if and how the compensation mechanism should be put into consideration are still open questions.

1) Fairness:

This metric mainly applies for the Best Effort (BE) service. One of the commonly used baseline schedulers in published research is the Proportional Fairness Scheme (PFS). The objective of PFS is to maximize the long-term fairness. PFS uses the ratio of channel capacity (denoted as Wi(t)) to the long-term throughput (denoted as Ri(t)) in a given time window Ti of queue i as the preference metric instead of the current achievable data rate. Ri(t) can be calculated by exponentially averaging the ith queue's throughput in terms of Ti. Then, the user with the highest ratio of Wi(t)/Ri(t) receives the transmission from the BS. Ti affects the fluctuation of the throughput. There are several proposals that have applied and modified the PFS. For example, given 5 ms frame duration, setting Ti to 50 ms is shown to result in an average rate over 1 second instead of 10 seconds with Ti = 1000 ms.

3) QoS Guarantee:

Modified Largest Weighted Delay First (M-LWDF) can provide QoS guarantee by ensuring a minimum throughput guarantee and also to maintain delays smaller than a predefined threshold value with a given probability for each user (rtPS and nrtPS) and it is provable that the throughput is optimal for LWDF. The algorithm can achieve the optimal whenever there is a feasible set of minimal rates area. The algorithm explicitly uses both current channel condition and the state of the queue into account. The scheme serves the queue j for which "ρiWj(t)rj (t)" is maximal, where ρi is a constant which could be different for different service classes (the difficulty is how to find the optimal value of ρi). Wi(t) can be either the delay of the head of line packet or the queue length. ri(t) is the channel capacity for traffic class i. There are several proposals that have used or modified MLWDF. For example, in [11], the scheduler selects the users on each subcarrier during every time slot. For each subcarrier k, the user (i) selection for the subcarrier is expressed by

max[channel gain(i, k) × HOL delay(i) × a(i) d(i)]

In this equation, a is the mean windowed arrival and d is mean windowed throughput. "a" and "d" are averaged over a sliding-window. HOL delay is the head of line delay. The channel state information is indirectly derived from the normalized channel gain. The channel gain is the ratio of the square of noise at the receiver and the variance of Additive White Gaussian Noise (AWGN). Then, the channel gain and the buffer state information are both used to decide which subcarriers should be assigned to each user. The buffers state information consists of HOL delay, a and d. Similar to M-LWDF, Urgency and Efficiency based Packet Scheduling (UEPS) was introduced to make use of the efficiency of radio resource usage and the urgency (time-utility as a function of the delay) as the two factors for making the scheduling decision. The scheduler first calculates the priority value for each user based on the urgency factor expressed by the time-utility function (denoted as U' i (t) the ratio of the current channel state to the average (denoted as Ri(t)/R' i(t)).

After that, the subchannel is allocated to each selected user I where:

i = max|U 'i (t)| × Ri(t) /R'i(t)

Another modification of M-LWDF has been proposed to support multiple traffic classes. The UEPS is not always efficient when the scheduler provides higher priority to nrtPS and BE traffic than rtPS, which may be near their deadlines. This modification handles QoS traffic and BE traffic separately. The HOL packet's waiting time is used for QoS traffic and the queue length for BE traffic.

3) System Throughput Maximization:

[10] focus on maximizing the total system throughput. In these, Max C/I (Carrier to Interference) is used to opportunistically assign resources to the user with the highest channel gain. Another maximum system throughput approach is the exponential rule in that it is possible to allocate the minimum number of slots derived from the minimum modulation scheme to each connection and then adjust the weight according to the exponent (p) of the instant modulation scheme over the minimum modulation scheme. This scheme obviously favors the connections with better modulation scheme (higher p). Users with better channel conditions receive exponentially higher bandwidth. Two issues with this scheme are that additional mechanisms are required if the total slots are less than the total minimum required slots. And, under perfect channel conditions, connections with zero minimum bandwidth can gain higher bandwidth than those with non-zero minimum bandwidth. Another modification for maximum throughput was proposed using a heuristic approach of allocating a subchannel to the MS so that it can transmit the maximum amount of data on the subchannel. Suppose a BS has n users and m subchannels, let i be the total uplink demand (bytes in a given frame) for its UGS connections, Rij be the rate for MSi on channel j (bytes/slot in the frame), Nij be the number of slots allocated to MSi on subchannel j, the goal of scheduling is to minimize the unsatisfied demand, that is,

$$Minimize \sum_{1\leq i\leq n} [\lambda_i - (\sum_{1\leq j\leq m} R_{ij}N_{ij})]$$

subject to the following constraints: $\sum_{1\leq i\leq n} N_{ij} < N'_j$ and $\sum_{1\leq j\leq m} R_{ij}N_{ij} \leq \lambda_i$

Here, N' j is the total number of slots available for data transmission in the jth subchannel. A linear programming approach was introduced to solve this problem, but the main issue is the complexity, which is O(n3m3N). Therefore, a heuristic approach with a complexity of only O(nmN) was also introduced by assigning channels to MSs that can transmit maximum amount of data.

4) Power Constraint:

The purpose of this class of algorithms is not only to optimize the throughput but also to meet the power constraint. In general, the transmitted power at a MS is limited. As a result, the maximum power allowable is introduced as one of the constraints. Least amount of transmission power is preferred for mobile users due to their limited battery capacities and also to reduce the radio interference. Link-Adaptive Largest-Weighted-Throughput (LWT) algorithm has been proposed for OFDM systems. LWT takes the power consumption into consideration. If assigning nth subcarrier to kth user at power pk,n results in a slot throughput of bk,n, the algorithm first determines the best assignment that maximizes the link throughput (max_bk,n). The bit allocation is derived from the approximation function of received SNR, transmission power and instantaneous channel coefficient. Then, the urgency is introduced in terms of the difference between the delay constraint and the waiting time of HOL packets. After that, the scheduler selects the HOL packet with the minimum value of the transmission time and the urgency. The main assumption here is that the packets are equal length. Integer Programming (IP) approach has also been used to assign subcarriers. However, IP complexity increases exponentially with the number of constraints. Therefore a suboptimal approach was introduced with fixed subcarrier allocation and bit loading algorithm. The suboptimal Hungarian or Linear Programming algorithm with adaptive modulation is used to find the subcarriers for each user and then the rate of the user is iteratively incremented by a bit loading algorithm, which assigns one bit at a time with a greedy approach to the subcarrier. Since this suboptimal and iterative solution is greedy in nature, the user with worse channel condition will mostly suffer. A better and fairer approach could be to start the allocation with the highest level of modulation scheme. The scheduler has to try to find the best subcarriers for the users with the highest number of bits. This is also a greedy algorithm in a sense of the algorithm is likely to fill the un-allocated subcarriers to gain the power reduction. To minimize the transmit power, a horizontal and vertical swapping technique can also be used. The bits can be shifted horizontally among subcarriers of the same user if the power reduction is needed. Or, the swapping can be done vertically (swap subcarriers between users) to achieve the power reduction. IEEE 802.16e standard defines Power Saving Class (PCS) type I, II and III. Basically PSC I increases the sleep window size by a power of 2 every time there is no packet (similar to binary backoff). Sleep window size for PSC type II is constant. PSC III defines a pre-determined long sleep interval without the existence of the listen period. Most of the proposals on this topic concentrate on constructing the analytical models for the sleep time; to figure out the optimal sleep time with guaranteed service especially delay (the more the sleep time, the more the packet delay and the more the buffer length). The models basically are based on the arrival process such as in [12] Possion distribution is used for arrival process. Hyper-Erlang distribution is used for self-similarity of web traffic. In order to reduce waking period for each MS, Burst scheduling was proposed in [10]. A rearrangement technique for unicast and multicast traffic is used so that a MS can wake up and received both type of traffic at once if possible [11]. In [13] a hybrid energy-saving scheme was proposed by using a truncated binary exponential algorithm to decide sleep cycle length for VoIP with silence suppression (voice packets are generated periodically during talk-spurt but not generated at all during the silent period).

IV. CONCLUSION

The MAC scheduler in WiMAX technology is a crucial issue to design. Fulfilling design constraints like guarantee of QOS, minimum throughput are utmost important. In this paper we discussed different scheduling algorithm that serves these design constraints out of which Weighted Fair Queuing Algorithm (WFQ) is found to be the best suited for WIMAX MAC scheduler. Although complex in design, it guarantees throughput, delay and fairness with proper and dynamic weight allocation to each queue of connections. Its dynamic nature makes it more meaningful then other algorithms.

REFERENCES

[1] Teng,Green,Bean"An Architecture for IEEE 802.16 MAC Scheduler Design"pp. 89-94,2007 IEEE.

[2] Valera Loscri"A queue Based dynamic approach for the coordinated distributed scheduler of the IEEE 802.16" pp. 423-428, 2008 IEEE.

[3] Darmawaty M. Ali, K. Dimyati "Performance Study of the WiMAX Uplink Scheduler" p. 831-835, December 2009 IEEE .

[4] Isha Khirwar, Anjulata Yadav, Preeti Trivedi"Comparative Assessment of WiMAX Scheduler in Fixed and Mobile WiMAX Networks for VoIP using QualNet" pp. 15-21, 2010 IEEE.

[5] Nawal Fishway, Mohamed Zahra"Modified cross-layered Scheduling for Mobile WiMAX Networks" April 2011 NRSC.

[6] Amani Sabri,Khaled Shehata"Video Scheduling in the Uplink over WiMAX Networks by the Novel concept 'Adaptive minimum service Interval'" pp. 3-8, 2010 IEEE.

[7] Cuicui Zhao, Jinlog Hu"Channel Quality Dependant Rate- Limited Scheduling algorithm for IEEE 802.16 Wireless Networks" pp. 402-407, 2009 IEEE.

[8] Guannan Qu, Zhiyi Fang"A New QoS Guarantee Strategy in IEEE 802.16 MAC Layer"pp. 157-163, 2008 IEEE.

[9] V. Singh and V. Sharma, "Efficient and Fair Scheduling of Uplink and Downlink in IEEE 802.16 OFDMA Networks," in Proc. IEEE Wireless Communication and Networking Conf., Las Vegas, NV, 2006, vol. 2, pp. 984-990.

[10] P. Parag, S. Bhashyam, and R. Aravind, "A subcarrier allocation algorithm for OFDMA using buffer and channel state information,".in Proc. IEEE Vehicular Technology Conf., Dallas, TX, 2005, vol. 1, pp. 622-625.

[11] Z. Yan, "Performance Modeling of Energy Management Mechanism in IEEE 802.16e Mobile WiMAX," in Proc. IEEE Wireless Communication and Networking Conf., Hong Kong, 2007, pp. 3205-3209.

[12] H. Choi and D. Cho, "Hybrid Energy-Saving Algorithm Considering Silent Periods of VoIP Traffic for Mobile WiMAX," in Proc. IEEE Int. Conf. Communications., Glasgow, Scotland, 2007, pp. 5951-5956.