International Conference on Recent Trends in Computational Methods, Communication and Controls (ICON3C 2012)
Proceedings published in International Journal of Computer Applications® (IJCA)

Weighted Fair Queue Scheduling Algorithm for IEEE802.16Wireless Networks

P.Saravanaselvi
Dept of ECE
Einstein College of engineering
Tirunelveli, India

P.Latha
Associate Professor/Dept of CSE
Government College of engineering
Tirunelveli, India

ABSTRACT

Worldwide Interoperability for Microwave Access (WiMAX), which is also known as IEEE 802.16 standards, supports last-mile broadband access wireless networks. WiMAX technology uses a number of scheduling techniques in the Medium Access Control layer, which is responsible for the utilization of available resources in the networks and distribute them among users in order to ensure the desired quality of service. In this work, the performance of Weighted Fair Queuing (WFQ) scheduling algorithm is evaluated to attain Quality of Service (QoS). The weight is assigned for a different service depends on the percentage of bandwidth utilization and priority of services. The proposed scheduling algorithm has designed and simulated using MatLab. The performance is evaluated with the other existing scheduling algorithm like First In First Out (FIFO), Priority queuing (PQ) and Weighted Round Robin (WRR).

Keywords

FIFO, QoS, PQ, WiMAX

1. INTRODUCTION

WiMAX was introduced by the IEEE 802.16 working group to facilitate broadband services on areas where cable infrastructure is inadequate. It provides triple play applications i.e. voice, data and video for fixed, mobile and nomadic applications. The key features of WiMAX including higher bandwidth, wider range and area coverage. WiMAX can provide wireless services up to 20 or 40 miles away from the base station. The requirement from IEEE 802.16 is to provide QoS for all possible applications in both (uplink and downlink) directions. The IEEE 802.16 Medium Access Control (MAC) specifies five types of QoS classes: 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) QoS classes.

Scheduling algorithms are responsible for distributing resources among all users in the network, and provide them with a higher QoS. Users request different classes of service that may have different requirements (such as bandwidth and delay), so the scheduling algorithm is to maximize the network utilization and achieve fairness among all users. The standard defines key parameters like minimum delay & high throughput to get maximum QoS.

Delay: We can measure delay by calculating difference between sending bytes and received bytes.

Throughput: We can find out the throughput by calculating number of bytes received per simulation time.

2. RELATED WORK

The packet scheduling algorithm is important within routers in high-speed integrated services packet networks for providing a wide range of quality-of service guarantees. Weighted fair queuing (WFQ) is an efficient packet scheduling algorithm for its delay and fairness properties. However, the timestamp computation in the WFQ scheduler is difficult. For that, many algorithms have been proposed to simplify the implementation of WFQ, such as Self Clocked Fair Queuing (SCFQ) [10], Frame-based Fair Queuing (FFQ) [8], Starting Potential Fair Queuing (SPFQ) [8], and Minimum Delay Self-Clocked Fair Queuing (MD-SCFQ) [9]. New Starting Potential Fair Queuing (NSPFQ) [7] is proposed which has O(1) complexity for virtual time computation and good delay and fairness properties.

In this work, the WFQ scheduling algorithm is evaluated with different weight assignment for different flows and compared it with some of the other scheduling algorithms that can be used to provide QoS. The remaining of this work is organized as follows. The overview of IEEE802.16 QoS services is presented in section III. In Section IV, the overview of scheduling algorithm in WiMAX networks is presented. The proposed system and its analysis are presented in section V. The detailed simulation results are explained in section VI.

3. IEEE 802.16 QoS Services

In IEEE 802.16, there are two models are defined. They are Point to Multipoint (PMP) networks and Multipoint to Multipoint (Mesh) networks [6]. PMP is an access network which includes a small number of Subscriber Stations (SSs) which are connected to a Base Station (BS). Multipoint to Multipoint is a network without centralized base station, and each subscriber station has the ability to connect directly to another subscriber station or via intermediate subscriber stations. The two independent channels are the Downlink Channel (from BS to SS) which is used only by the BS, and the Uplink Channel (from SS to BS) which is shared between all SSs, in Mesh mode, SS can communicate by either the BS or other SSs, in this mechanism the traffic can be routed not only by the BS but also by other SSs in the network, this means that the uplink and downlink channels are defined as traffic in both directions; to and from the BS.

The data transfer between the uplink and downlink directions with the help of Time Division Multiple Access (TDMA) in MAC protocol. The time is divided in to frames separated by time intervals. Each frame is divided between uplink sub frame and downlink sub frame. There are two fields for managing allocation of wireless communication. DL-MAP: downlink bandwidth allocation map to tell the SS of the timetable and physical layer transmission packets bursts. UL-MAP: uplink bandwidth allocation map. It controls the amount of time each SS is given access to the channel in the next uplink sub frames [3].

To support a wide variety of applications, WiMAX defines five QoS classes that should be supported by the BS:

3.1. Unsolicited Grant Service (UGS)

UGS is designed to support real-time data streams consisting of fixed-size data packets issued at periodic intervals. The BS provides fixed-size data grants at periodic intervals, like the case in E1and VOIP without silence suppression [5].

3.2. Real-Time Polling Service (rtPS)

rtPS is designed to support real-time data streams consisting of variable-sized data packets that are issued at periodic intervals. The BS provides periodic unicast (uplink) request opportunities, like the case in MPEG video transmission [5].

3.3. Extended Real-Time Polling Service (ertPS)

ertPS is suitable for variable rate real time applications that have data rate and delay requirements, like the case in VOIP without silence suppression. The IEEE 802.16e standard indicates that ertPS is built upon the efficiency of both UGS and rtPS. The BS provides unicast grants in an unsolicited manner like in UGS[5].

3.4. Non-Real-Time Polling Service (nrtPS)

nrtPS is designed to support delay tolerant data streams consisting of variable size data packets for which a minimum data rate is required, like the case in FTP traffic. The BS provides unicast uplink request polls on a regular basis, which guarantees that the service flow receives request opportunities even during network congestion.

3.5. Best Effort (BE)

BE is designed to support data streams for which no minimum service guarantees are required, like the case in HTTP traffic.

The two different approaches are defined for queues: Per-flow Handling approach which is to have a separate queue for each individual session or flow. It becomes very difficult with a large number of flows. The IntServ methods use per-flow handling of IP packets. The second approach is Aggregate Handling which is to classify packets into a few different generic classes putting each class in a different queue that is more scalable and reduces the maintenance and processing. DiffServ use aggregate traffic-handling mechanisms for IP and Ethernet.

4. SCHEDULING ALGORITHM

In general, schedulers can be characterized as work conserving or non work-conserving. A scheduler is work conserving if the server is never idle when a packet is buffered in the system. A non work-conserving server may remain idle even if there are available packets to transmit. A server may, for example, postpone the transmission of a packet when it expects a higher priority packet to arrive soon, even though it is currently idle.

Non work-conserving algorithms are used to control delay jitter by delaying packets that arrive early. Work conserving servers [3] always have lower average delays than non work-conserving servers. Examples of work-conserving schedulers include generalized processor sharing (GPS), WFQ, virtual clock, and weighted round robin (WRR). On the other hand, hierarchical round robin (HRR) and stop-and-go queuing are non work-conserving schedulers.

In this work the work conserving scheduler algorithm is analyzed. The WFQ, or packet-by-packet GPS (PGPS), algorithm is the packet-by-packet equivalent of GPS that is; it derives the system virtual time from the background simulation of a GPS server. The system virtual time $v(t)$ of WFQ evolves as that of the corresponding GPS system. Some of the scheduling algorithm is evaluated with WFQ.

4.1. FIFO scheduling algorithm

The most basic queue scheduling discipline is First In First Out (FIFO) algorithm. All packets are treated equally by placing them into a single queue, then servicing them in the same order that they were placed into the queue. The behavior of a FIFO queue is very predictable packets are reordered and the maximum delay is determined by the maximum depth of queue. The limitation of FIFO is that a single queue impacts all flows equally because the mean queuing delay for all flow increases as congestion increases. As a result, FIFO queuing can result in increased delay and jitter.

4.2. Priority queuing allgorithm

It is the basis for a class of queue scheduling algorithm that is designed to provide a relatively simple method of supporting differentiated service classes. Packets are first classified by the system and then placed into different priority queues. Packets are scheduled from the head of a given queue only if all queues of higher priority are empty. It allows routers to organize buffered packets and then service one class of traffic differently from other class of traffic. The limitation is that if the amount of high priority traffic is not policed of conditioned at the edges of the network. Lower priority traffic may experience excessive delay as it waits for unbounded higher priority traffic to be serviced.

4.3. Weighted Round Robin algoirthm

In this queuing, packets are first classified into various service classes and then assigned to a queue that is specifically dedicated to that service class. Each of the queues is serviced in a round robin order. The empty queues are skipped. It provides an efficient mechanism to support the delivery of differentiated service classes to reasonable number of highly aggregated traffic flows. The primary limitation of WRR is that it provides the correct percentage of bandwidth to each service class only if all of the packets in all of the queues are the same size or when the mean packet size is known in advance.

4.4 Weighted Fair Queuing Algorithm

It is the basis for a class of queue scheduling disciplines. Whenever a packet finishes transmission, the packet sent is the one with the smallest value of F_i^α. The finishing time is calculated by

$$F_i^\alpha = S_i^\alpha + \frac{P_i^\alpha}{\phi_\alpha} \text{--------------------(1)}$$

$$S_i^\alpha = \max[F_{i-1}^\alpha, R(\tau_i^\alpha)] \text{----------(2)}$$

The differing demands of different sources, the processor sharing discipline to allow for arbitrary capacity allocations. With Generalized Processor Sharing (GPS), each flow α is assigned a weight ϕ_α that determines how many bits are transmitted from that queue during each round. The effective packet length is to $1/\phi_\alpha$ times the true packet length. It is easy to see that, at any given time, the service rate g_i for a non empty flow i is calculated by using equation 3.

$$g_i = \frac{\phi_i}{\sum_j \phi_j} C \text{-----------------------(3)}$$

Where the sum is taken over all active queues and C is the out going link data rate. The maximum delay experienced by flow i, D_i is bounded by the equation 4.

$$D_i \le \frac{B_i}{R_i} \text{------------------------------(4)}$$

The set of flows that are defined by and limited to the token bucket specification. B_i and R_i are the bucket size and token rate respectively for flow i. The weight assigned to each flow equal the token rate. This is explained in figure 1.

All packets have the same size of one and link speed is one. The guaranteed rate of connection 1 is o.5 and the guaranteed rate for the other 10 connection is 0.05. Flow 1 sends 11 back to back packets starting at time and the other flows send a single packet at time 0. Under FIFO, one packet from each flow is transmitted and then the remaining 10 packets of flow 1. Under WFQ, the first ten packets of on flow 1 all have processor share finish times smaller than packets on other connections, the will transmit these packets first.

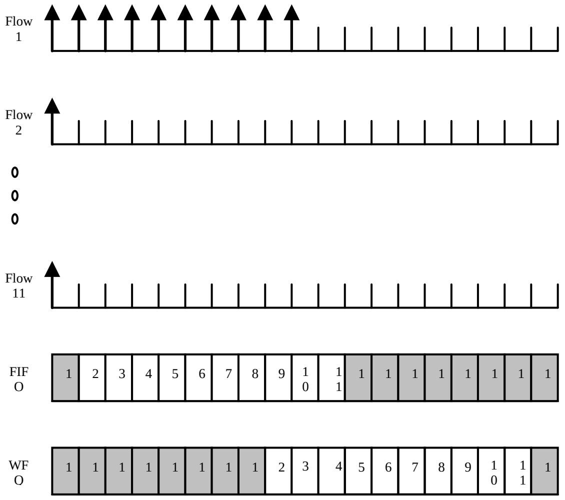

Figure.1. Comparison of FIFO and WFQ

The maximum delay experienced by flow i is modified and is given in equation 5.

$$D_i \le \frac{B_i}{R_i} + \frac{(K_i - 1)L_i}{R_i} + \sum_{m=1}^{K_i} \frac{L_{\max}}{C_m} \text{-----------(5)}$$

Where K_i is number of nodes in the path flow i through the internet, L_i is maximum packet size for flow i, L_{max} is maximum packet length for all flows through all nodes on the path of flow i and C_m is outgoing link capacity at node m.

The WFQ derives the system virtual time [9] from the background simulation of a GPS server. The system virtual time v(*t*) of WFQ evolves as that of the corresponding GPS system, whose derivative is given in equation 6.

$$\frac{dv(t)}{dt} = \frac{r}{\sum_{i \in B(t)} r_i} \text{--------------------- -----------(6)}$$

Where *B*(*t*) is the set of sessions that are backlogged in the GPS server at time *t*, and *r* is the server rate. WFQ achieves delay bound and fairness properties very close to those of GPS. However, since all *N* sessions can join or leave the set of backlogged sessions during a packet transmission time, the worst-case complexity of maintaining the system virtual time is O(*N*), which makes the algorithm not suitable for practical deployment in high-speed packet networks.

5. SIMULATION RESULT

The system model is considered with 11 flow sessions. The overall goal of this simulation study is to analyze the performance of different existing scheduling algorithm in Mobile WiMAX environment.

Eleven queues have been configured to avoid queuing packets of different service types into one queue. Even if the application sets a high precedence for its packets, they may be blocked by lower precedence packets in network queues. The precedence values corresponding for each queue are shown in the Table I.

Table 1. Traffic Class Vs

Mac Layer Services	Precedence value/Queue
BE	0
nrtPS	2
rtPS	3
ertPS	4
UGS	7

To evaluate the performance of scheduling algorithm, both qualitative and quantitative metrics are needed. In this work, focuses on the QoS most important metrics which are throughput and the average end-to-end delay. The five QoS classes have been compared in four different scheduling algorithms. The simulation model is considered for five different services with precedence value. The result is taken with constant data rate and service rate for varying packet length.

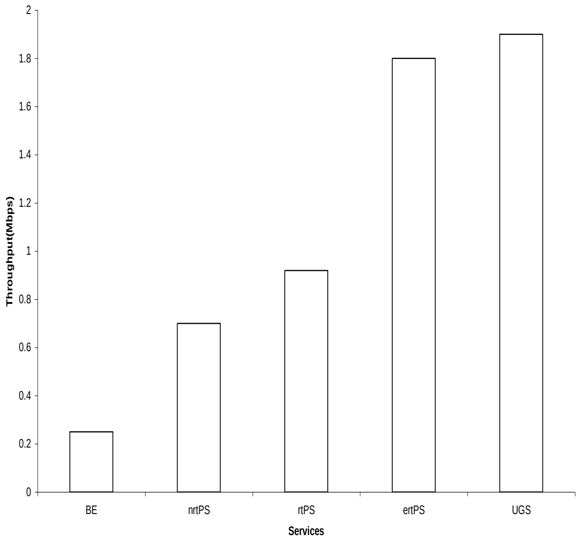

Figure.2. Throughput Computation for different

Figure.2.shows that higher priority service has high throughput for WFQ scheduling algorithm.

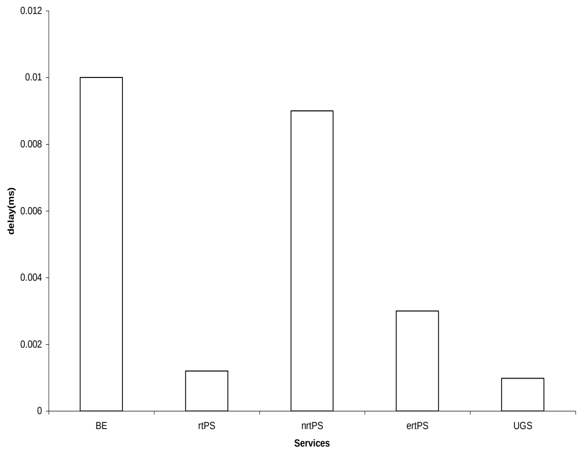

Figure.3. Delay computation for various Services

Figure.3.shows that higher priority service has low delay for WFQ scheduling algorithm. The delay is computed for WFQ algorithm based on the equation 5.

The simulation model is considered for four different scheduling algorithms. The result is taken with constant data rate and service rate for varying packet length.

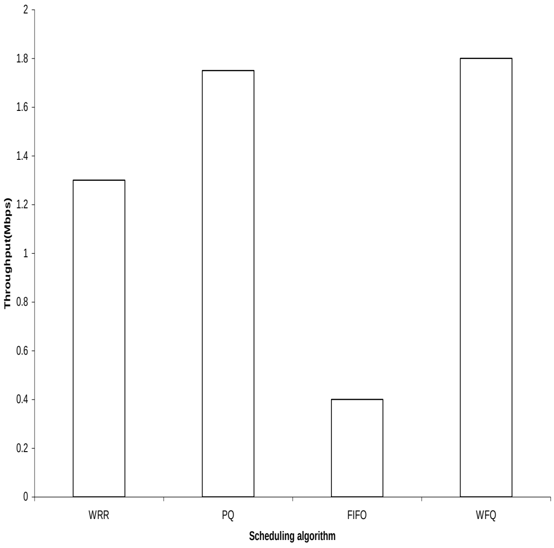

Figure.4.Throughput Comparison for different

Figure 4 shows that the WFQ algorithm performs well with less computational complexity and its throughput value is high compared to other scheduling algorithms.

Table.2.Comparison table

Services	Delay(ms)	Throughput(Mbps)
BE	0.01	0.25
rtPS	0.0012	0.92
nrtPS	0.009	0.7
ertPS	0.003	1.8
UGS	0.00098	1.9

6. CONCLUSION

The behavior of Weighted Fair Queuing scheduling algorithms in WiMAX has been investigated in this work. A simulation study was used to compare the performance of each scheduler on the different QoS classes. The simulations is verified that the Priority queuing has the highest throughput for high QoS classes. Both WFQ and WRR can control the performance of each class by assigning different weight to each queue.

7. REFERENCES

[1] Wail Mardini, “Modified WRR scheduling algorithm for WiMAX networks”, Network Protocols and Algorithm, Volume3, 2011.

[2] Harpreet Kaur et.al, “Implementation and Evaluation of scheduling algorithm in point to multipoint mode in WiMAX network”, IJCST, Volume 2, 2011

[3] Ahmed H.Rashwan et.al, “Comparative Assessments for Different WiMAX Scheduling Algorithms”, WCECS, 2009.

[4] Ala’a Z. Al-Howaide, Ahmad S. Doulat, Yaser M. Khamayse, “Performance evaluation of different Scheduling algorithms in wimax”, Proceedings of the World Congress on Engineering and Computer Science, Vol I, 2009.

[5] Nuaymi L., WiMAX: *Technology for Broadband Wireless Access*, John Wiley & Sons, 2007.

[6] Bo Han, Weijia Jia, Lidong Lin, “Performance evaluation of scheduling in IEEE 802.16 based wireless mesh networks”. Computer Communications, Vol. 30, 2006.

[7] Dong-Yong Kwak et.al, “A New Starting Potential Fair Queuing Algorithm with O(1) virtual time computation complexity”, ETRI Journal, Volume 25, 2003.

[8] D. Stiliadis and A. Varma, “Efficient Fair Queuing Algorithms for Packet-Switched Networks,” *IEEE/ACM Trans. Networking*, vol. 6, no. 2, Apr. 1998, pp. 175-185.

[9] F.M. Chiussi and A. Francini, “Minimum-Delay Self-Clocked Fair Queuing Algorithm for Packet-Switched Networks,” Proc. Of IEEE INFOCOM ’98, vol. 3, Mar. 1998, pp. 1112-1121.

[10] S.J. Golestani, “A Self-Clocked Fair Queuing Scheme for Broadband Applications,” Proc. of IEEE INFOCOM ’94, Apr. 1994, pp. 636-646.

[11] A.K. Parekh and R.G. Gallager, “A Generalized Processor Sharing Approach to Flow Control in Integrated Services Networks: The Single Node Case,” Proc. of IEEE INFOCOM’92, vol. 2, May 1992, pp. 915-924