A Dynamic Traffic Shaping Technique for a Scalable QoS in ATM Networks

Francis Ogwu[1], Mohammad Talib[1], Ganiyu Aderounmu[2], and Olufade Onifade[3]
[1]Department of Computer Science, University of Botswana, Gaborone, Botswana
[2]Department of Computer Science and Engineering, Obafemi Awolowo University, Nigeria
[3]Department of Computer Science, University of Ibadan, Nigeria

Abstract: *Traffic shaping function becomes imperative for the new broadband services that are being deployed in order to avoid information loss, to provide the end users multiple traffic options in terms of bandwidth and to ensure optimal use of the communication channels. To simultaneously manage the amount of cell loss and delay experienced by two or more classes of service categories constant bit rate/ variable bit rate, we developed a new buffer partitioning scheme tagged complete sharing with gradual release. The proposed model was combined with a scheduling method known as weighted round robin with absolute increment. An analytical model was developed for the proposed buffer partition to dynamically monitor and determine the output mean rate of the classes of service present, and the individual mean rate of the class of service. The model was simulated and performance evaluation carried out. The result thus obtained depicts a better performance as a method of traffic shaper in a multi-quality of service traffic over asynchronous transfer mode networks.*

Keywords: *Asynchronous transfer mode, quality of service, complete sharing/ gradual release, traffic shaper, constant bit rate, variable bit rate.*

Received July 25, 2006; accepted May 31, 2007

1. Introduction

Asynchronous Transfer Mode (ATM) network has been accepted as a transport platform for the next-stage tele-communication network [3]. Supporting existing, future and even unknown applications now behooves on ATM-networks to be able to integrate a variety of service with diverse traffic characteristics and Quality of Service (QoS) requirements. Traffic management in ATM networks consists of several functions acting at different levels from cell level to connection level [9]. An attempt to guarantee QoS requirements for all connection therefore should entail control mechanisms like Connection Admission Control (CAC), Usage Parameter Control (UPC), traffic scheduling, traffic shaping, congestion control and flow control for handling different characteristics and their point of occurrence [4]. To realize this, ATM standards defined Virtual Path Connections (VPCs), which contain Virtual Channel Connections (VCCs). A VCC (or Virtual circuit) is the basic unit, carrying single stream of cells from user to user. These can be bundled together into a VPC. A VPC can be created from end-to-end across ATM networks [12]. Therefore, cells belonging to a particular virtual path are rooted the same way through the ATM virtual network, thus making possible faster recovery in the occurrence of major failures.

The above has necessitated the need for buffering in/ around the switches to delay the onward transmission of cells onto the output link. Buffering techniques, location and buffer size therefore becomes a determining factor in the design of ATM traffic shaping/ switch architecture. The research therefore focuses on the development of a dynamic traffic shaping techniques for a suitable multi-QoS over ATM switches system.

2. Blocking and Non-Blocking Traffic Control Measures

Broadly speaking, the objectives of ATM traffic management are to deliver QoS guaranteed for the multimedia application and provide overall optimization of network resources. Meeting these objectives enables different classes of service offering the potential for service differentiation and increased revenues, while at the same time simplifying network operations and reducing network cost [5]. The traffic management and its various functions can be categorized into three distinct elements based on timing requirements [2, 3]. These elements are: nodal-level control, operating in real time, network-level control operating in non-real time and network engineering capabilities operating in non-real time. To control these dynamic situations, we employ the preventive control mechanism and the reactive control mechanisms.

Traffic control mechanism has two major components: Call Admission Control (CAC) and Usage Parameter Control (UPC) [1]. Preventive control attempts to manage congestion by preventing it before it occurs. It is targeted primarily at real-time traffic [3]. Reactive traffic control is useful for service class like Available Bit Rate (ABR), which allows sources to use bandwidth unutilized by cells in other classes. This is because, bandwidth usage of this class is dynamically changing based on its availability in the network. There exist two major classes of reactive traffic control: rate-based and credit based [10]. Common examples include the Enhanced/ Proportional Rate Control Algorithm (EPRCA) [7] and Quantum Flow Control (QFC) algorithm for rate-based and credit-based traffic control algorithm respectively.

2.1. ATM Traffic Shaper

Traffic shaping is a major factor for the operation of the ATM networks. This is to allow for the full realization of the multiplexing ability of ATM networks. Several attempts had been made from inception to actually monitor this, among which are: the switch architecture that engulfs the blocking and the non-blocking switches [3]. The above were marred by the cells contention arising from the transmission of the multiplexed cells. The next stage was the buffering stage. This is not blocking, and various locations have been proposed for the placement of the buffer within the switch architecture. However, this development suffered the fact that most cells were delivered out of sequence, not minding the intended QoS negotiated from inception. This is as shown in Figure 1.

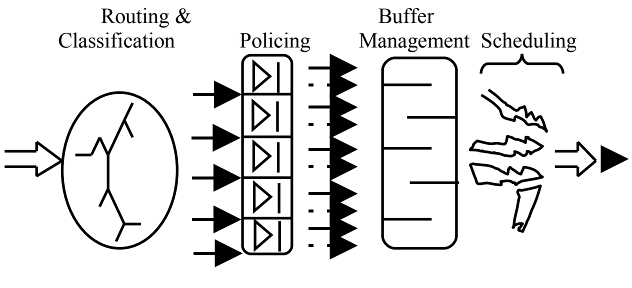


Figure 1. Combination of scheduling and buffer management in ATM traffic shaping.

Traffic shaping is a mechanism that alters the traffic characteristics of a stream of cells on a connection to achieve better network efficiency while meeting the quality of service objectives, or ensuring conformance at a subsequent interface [11]. Traffic shaping must also maintain cell sequence integrity on a connection. The design of a traffic shaper using a priority scheduling and buffer management was discussed in [8]. An attempt was made to balance the amount of cell loss and delay associated with burst input traffic. The scheduling was a dynamically weighted priority with the aim of reacting to input of the traffic using equation 1.

$$P(t) = Uj / [W_j(t)]^{\beta} \quad (1)$$

where $P(t)$ is the instantaneous priority index at time t, U is the associated fixed priority number, (a lower U means higher priority, and $W(t)$ is the amount of time the oldest cell in the j^{th} class has waited in queue j). The workability of the above is based on the continue recalculation of the priority index for each queue at every output time slot (based on the outbound link). The above present's two major problems: The continuous recalculation of the priority index for the determination of the queue to be permitted to transmit a cell at a particular time is not adaptively dynamic. And the choice of β would pose a serious problem in a non-deterministic environment.

The partitioning employed in the above is the limiting type, i.e., sharing with minimum allocation, although efficient but a class of service like VBR with ON-OFF transmission resulted in non-optimal utilization of the minimum allocation during the OFF period. It is clearly evident that increase in input traffic load is inversely proportional to the network efficiency after a critical occupancy level is attained [8]. There exists many congestion control mechanisms aimed at correcting, limiting or avoiding this situation. Applicability of the above is determined by the duration of the congestion to be controlled. This could either be by CAC on dynamic routing to limit congestion caused by added network connection. Others are: dynamic compression and end-to-end feedback [3]. Congestion as a result of short burst of traffic on existing connections can be limited using buffering. Buffering partitioning delineates/ demarcates the amount of buffer space available to a given queue and defines how the queue is to be shared amongst different queues. This can be generally classified into three types [6, 8]. Complete sharing schemes: -this is very efficient, but common problems that usually arise are the issue of fairness. In a situation where a single queue consumes the entire buffer [1]. Complete partitioning scheme and sharing with minimum allocation scheme. The discarding policy determines whether an incoming cell is to be dropped or placed in a buffer space [8]. The buffer partitioning employed however was dynamic. The above is aimed at presenting a situation where there's no cell-discard until all buffers space is completely consumed; therefore when one class is idle, the other class can make use of the entire buffer optimally. This research therefore feature a new approach to traffic shaper design using a flavor of Weighted Round-Robin with Absolute Increment (WRR/ AI) and a newly developed buffer partitioning scheme known as Complete Sharing with Gradual Release (CS/ GR), to eliminate the state of idleness during the OFF period and at the same time guaranteeing servicing of all classes without prejudice

3. Proposed Model

Dynamic Level of Threshold (DLT) dealt with cell-loss by prioritizing a particular type of traffic at the expense of other thereby incurring appreciable delay and sometimes resulting into cell loss of other traffic types. A critical study of the DLT scheme under the PBS depicts that intermittent burst of traffic cells cannot be properly handled by the scheme because of the deterministic pattern of buffer size growth taking into consideration that ATM traffics comprises of both deterministic and non-deterministic pattern of transmission. There is no prior-provision for the occurrence of burst traffic, which is a major factor to be considered in ATM networks congestion control. To this end, occurrence of burst will lead to significant loss of the low priority cells as the buffer begins to discriminate in the admittance of cells during network congestion resulting from bursts.

Again, reduction of loss with increase in buffer size is a common phenomenon. However, it should be pointed out that not all transmitting sources are usually active all the time; if this is taking into consideration, then there will be wastage of resources (buffer) when the transmission is not utilizing the maximum capacity provided by the DLT. CS/ GR however, address the issue of buffer capacity utilization through the result of equation 6 using the determination of the mean rate of transmission prior to buffer space allocation, i.e., δ.

On the same vein, admittance/ non-admittance of the low-priority cells at a particular threshold is another drawback of the DLT scheme against transmission and space allocation fairness. Indeed, some classes of service can endure delay, but not indefinitely. Delay leading to loss of cells as recorded by the DLT on the part of the low-priority cells is quite significant considering the size of the buffer. With increasingly growing buffer, the implementation of DLT is likely going to be limited to particular environment/ switch architecture. This is due to the fact that not all point of traffic control can guarantee the stupendous volume of buffer required to support the implementation of the DLT scheme. The performance evaluation of CS/ GR is aimed at providing means of addressing the above stated drawbacks of DLT as a buffer space management scheme in an ATM networks.

3.1. Complete Sharing/ with Gradual Release

The proposed model consists of three modules. These are: the generation, the buffer, and the scheduling module. Generation is the major standards of cell generation was adhered to without any throttling mechanism to increase or slowing down the generating rate of cells, distorted signals/ cells was assumed to have been discarded prior to entrance into the buffer. However, the generation focuses on two different services: Constant Bit Rate (CBR) and the Variable Bit Rate (VBR). The former traffic pattern is deterministic and continuous, while the latter however, is timed, synchronous generation, which makes it poisson in arrival.

Buffer is the major contribution of this paper work and is aimed at showcasing the effectiveness of its combination with a choice-scheduling algorithm to properly man the traffic shaper of an ATM network. The presented buffering is therefore tagged CS/ GR a dynamic type.

To this end, the buffer was presented as a virtual layered segment. Cells of type 1 belonging to CBR are stored in segment 1, and cells of type 2 belonging to VBR are stored in segment 2. The above is subject to the proviso that: (1) if segment 1 is already filled-up and segment 2 is not; cells of type 1 can begin to occupy segment 2 gradually. (2) If cell type 2 arrives at full buffer as a result of type 1 occupying segment 2, then cell type 1 must release the space gradually to cell type 2, and (3) either cells arrival to a full buffer, i.e., segment 1 + segment 2 = total capacity lead to a loss.

Scheduling is based on the WRR/ AI. It is however prescribed that cells manipulation/ treatment within the buffer is based on FIFO. It is should however be noted that it will be very difficult if not totally impossible to clearly distinguished the above three entities as the operation is parallel to one another.

3.2. Numerical Model Development

The configuration of the traffic shaper consisted of two queues. Queue 1 is reserved for CBR traffic class, and queue 2 is reserved for the VBR traffic class. With a finite buffer capacity of size (X entries), made available to the traffic shaper, the implementation of the buffer space allocation was such that each queue is provided with a guaranteed number of entries ($^{X}/_{2}$), however the whole buffer space can be consumed by a single queue if the other party (queue) is idle, but with an understanding that: a minimum amount of the buffer space will be relinquished gradually by the queue utilizing more than its guaranteed allotment. It is assumed that the acceptance of either of the queue into the traffic shaper's buffer prior to transmission has also undergone policing during which cell discard could have occurs. Cell loss under these circumstances is not considered, as shown in Figure 4.

From the above, the role of the traffic shaper is to assign the available bandwidth to the cells that are ready and wanting to be sent. Knowing fully well that the bandwidth is neither created nor destroyed/ eliminated, the incremental pattern of scheduling and the buffer management employed will keep cell-loss and delay at the barest minimum. However, delay introduced by a specific traffic-shaping algorithm must be considered in the actual implementation. One unit time is assumed to be the time required to send an ATM cell. Counter reset and determination of which class is allowed to send a cell is prejudged just before each cell output. The CBR source is transmitting data at a constant rate (which is less than or equal to the output link capacity). The VBR sources were transmitting "burst" traffic at a similar rate but with an ON and OFF period. This "burst" traffic is actually poisson. The poisson distribution has facilitated the easy determination of the mean value to be the time during which the queue is active (ON). To this end, the input rate say λ is denoted by λ $(O<\lambda<1)$, and the probability $P(x)$ where X is the cell arrival at a unit time is given by:

$$P(x) = e^{-\lambda}\ \lambda/X! \quad (2)$$

weighted round robin cell scheduling with absolute increment according to [13] has the following notations. Suffice it to say that *Wn* signifies the weight of the classes present in the network; *Cn* is the initial counter values of the classes as well.

In this paper, since two important classes were considered, we have the weights to be (*W1*, *W2*) and the initial counter values to be (*C1*, *C2*), however, a more generalized format will be presented to allow for scalability as would be deemed fit. To this end, during the first cycle we have:

W1 = Weight of Class 1
W2 = Weight of Class 2
. :
Wn = Weight of Class *n*

and

C1 = Initial counter values of Class 1
C2 = Initial counter values of Class 2
:
Cn - Initial counter value of Class *n*

The effect of the absolute increment begins from the next counter-reset. Therefore, at the next cycle:

C1 = *ABS {min (C1, C2-----Cn)} + W1*
C2 = *ABS {min (C1, C2-----Cn)} + W2*
:
Cn = *ABS {min (Cs, C2-----Cn)} + Wn*

From the above, a generalized model/ formula is thus given as:

Cn + ABS {(C1, C2-----Cn)} + Wn (3)

It is to be noted that there exist one-to-one correspondence between each classes and their weights, i.e., *Class 1* (*C1*) has *weight 1* (*W1*) etc. Dynamically, managing buffer space means that all shared buffer space is flexibly allocated to Virtual Connections (VCs) and as at needed basis. To carefully allot proportionate sizes to the network needs we mathematically proceed as follows: Suppose δn represents the rate for all the classes, and the time taken is given by t total, then: for CBR transmission, the mean rate is given as:

$$\bar{\delta}_{CBR} = \delta_{CBR} \quad (4)$$

The (VBR) transmission, the mean rate is therefore given as:

$$\bar{\delta}_{VBR} = \frac{\delta_{VBR} * t_{ON}}{t\ total} \quad (5)$$

The total time taken in a cycle is therefore T total = $t_{ON} + t_{OFF}$. Since we are considering these two classes in this paper, the output link will be contending with the maximum of: sum ((4) and (5))

$$\text{i.e.,} \quad = \frac{\bar{\delta}_{VBR} * t_{ON}}{t\ total} + \delta_{CBR}$$

$$\Rightarrow \bar{\delta}_{CBR} + \bar{\delta}_{VBR\ ont} = \delta_{CBR} + \delta_{VBR} \quad (6)$$

Equation 5 shows the total amount of data to be buffered prior to scheduling. The equation is readily scalable so long the rate of transmission can be determined. Graphically, it can be represented by Figure 2, showing the analysis and representation of the traffics i.e., CBR and VBR.

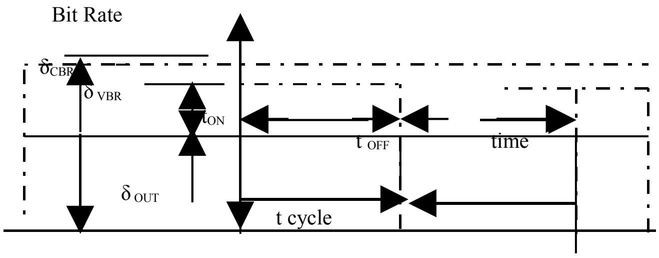
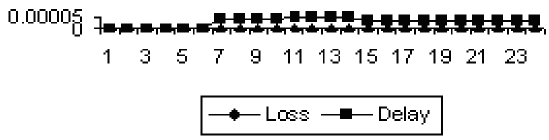


Figure 2. Traffic representation of CBR and VBR.

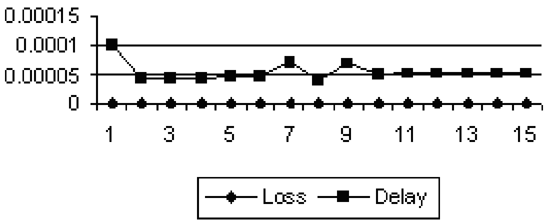
4. Model Simulation and Results Analysis

The model was simulated and performance evaluation carried out using the sun Microsystems Java 2 software development kit. This is because Java offers unique capabilities that are fueling the development of mobile agent systems and also very good in poking computer communication ports. The simulation environment was abstracted as a two-dimensional space divided into two queues. During the simulation, the following generic parameter is considered: the service rate is 366792 cells per sec. This is with a view of working with the approved standard for service rate based on 155.52 Mbps for optic fiber links.

Performance evaluation of CS/ GR: -in the simulation, the delay and loss probability of the real-time cells and non-real time cells as a function of the transmission rate adopted for the class type were measured. Figure 4.1-a shows the effect of transmission rate of cells as to the delay incurred and the possibility of cell loss. It would be recalled that the class 1 type is very sensitive to delay CBR and thus more weighted than the other classes. However, there is no discrimination in allotment and occupation of the buffer. From Figure 3-a, at a transmission rate of 2000 cell/sec, the delay associated with class 1 type is almost out the zero level, there is a gradual increase which is still less than 1 x 10^{-4} (timing) until the transmission rate gets to 3000 cells/sec. The sharp increase witnessed at this point could be understood by recalling that the transmission of this cell type is not in isolation of the other classes. Therefore, the burst arrival of the class 2 cells at this generating rate cause the unprecedented increase which is however, still manageable (almost insignificant) with values less than 0.00004. The delay witnessed continues to hang around the value of 0.00003 and 0.00004 because of the presence of more of class type 2 in the buffer.



a. Delay and loss probabilities of real time cells (CBR).



b. Delay and loss probabilities of non-real time cells (VBR).

Figure 3. Delay and loss probabilities of CBR and VBR.

However, the output mean rate (*δout*) hitherto calculated in determining the amount of data to be buffered has greatly reduced the incidence of cell loss to a constant value of zero even up to a transmission rate of 6400 cell/ sec. This is achievable due to the incremental nature of the cells scheduling and the dynamic buffer management adopted in this paper.

The efficacy of the buffering pattern is further shown in Figure 3-b. The essence of CS/ GR is clearly vindicated here. It is clearly seen here that the delay associated with class 2 type cells is almost four-times that of the class 1. This is attributed to the fact that, class 1 has begun transmission and buffer occupying prior to class 2 commencement of transmission. The implication of this is that, class 1 has already taking over part of the buffer space belonging to class 2, but has to release it gradually as it is requested of it. Another factor is that, since class 2 is transmitting for half of the time in a total cycle, its arrival intrudes an already established process. At this point, it can be deduced that the dynamic allotment of the buffer space is important to safeguard the loss of cells of other types of traffic classes. After the initial problem, the system was able to adjust and accommodate subsequent arrival, due to the δ_{ot} determination and the scheduling method employed. Since provision was made for more cells to be transmitted at the end of each reset, the delay for this class dropped considerably and began taking a steady pattern, until at 3000 cells/ sec. rate when there comes another surge. The delay again became epileptic, moving between 0.00004 and 0.00008 presenting a zigzagging chart resulting from the increase in transmittable number of cells.

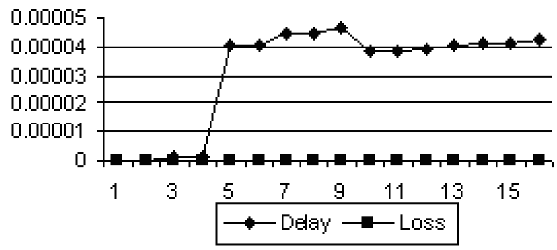
Despite the burst nature of the class 2 type traffic, CS/ GR buffering pattern is able to successfully accommodate and manage the increasing transmittable cells with a zero cell-loss recorded as high as 5,000 cells/ sec.

4.1. Comparative Analysis of CS/ GR on other Scheduling Schemes

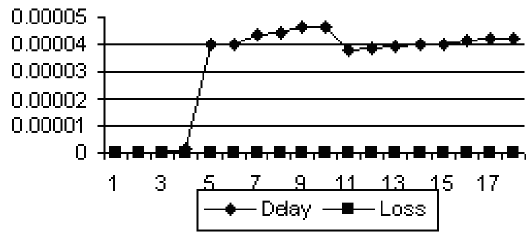
CS/ GR method of buffering partition has been demonstrated to combine successfully with WRR/ AI method of scheduling as shown in Figure 3-a and 3-b respectively. However, it has become expedient to further depict its dynamic method of operation by combining it with other known scheduling schemes.

WRR is very popular for its simplicity and ease of implementation. However, a big question has risen as to its support for throughput, fairness and strict adhering to the negotiated QoS during the process of transmission. The scheduling scheme is tested with the two classes of service under the study. The results of the simulation present the effect of the delay and loss probabilities of both classes as a function of the transmission rate to allow for a balanced scene in comparison.

Figure 4-a represents the result of the real-time cells CBR, similarly Figure 4-b represent the result of non-real time cells VBR. These two classes were considered together, to allow for pointing out the similarities and the effect of buffering.



a. Delay and loss probabilities of real time cells using WRR.



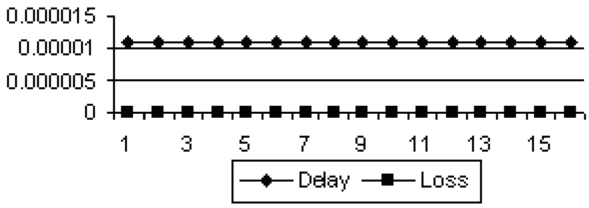
b. Delay and loss probabilities of non-real time cells using WRR.

Figure 4. Delay and loss probabilities of real time and non-real time using WRR.

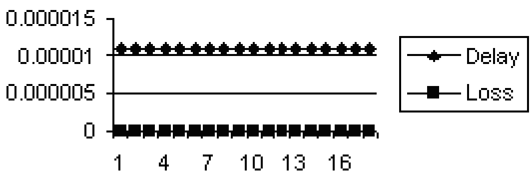
The first thing to be noted is the close pattern of delay witnessed by both classes (CBR and VBR). Although, the cell service rate is fair, there is no distinction in the treatment of class 1 types and class 2 types, as shown in Figures 4-a. This has thus rubbished the QoS negotiated for the classes prior to transmission. In real life, class 1 type is not suppose to be treated the same way as class 2 type when considering associated delay and number of transmittable cells.

However, loss of cells of class 1 type which could have resulted from the continuous streaming -in of cells was contained even during burst traffic states due to the dynamic buffering management employed.

First-In-First-Out (FIFO) is one of the oldest methods of scheduling with almost a 100% of simplicity of implementation. Its effectiveness in treating cells is however questionable, as there is no weight/ priority associated to any traffic type, but all were treated as “who comes in first-gets-out first”. The effect of the combination of FIFO with CS/ GR is also simulated. The result of the simulation is as shown in Figures 5-a and 5-b respectively.



a. Delay and loss probabilities for real time cells using FIFO.



b. Delay and loss probabilities of non-real time cells using FIFO.

Figure 5. Delay and loss probabilities of real time and non-real time using FIFO.

It is clearly shown that this method (combination) recorded the worst delay for the two classes of service. The reason is not far fetched, but for the “wait for your turn” method of operation not minding the urgency required by a particular class. To this end, there is no guaranteed QoS from this scheduling scheme given in Figure 5-a and 5-b, respectively. There is no distinction in cell type transmission even for the real-time class CBR. The delay is high and constant for the two classes. The effect of the buffering method is however noticeable in handling the number of cell loss recorded in the experiment. This again can be attributed to the accommodating nature of the buffer, striking a balance (fairness) in the treatment of both classes, as shown in Figure 5-a and 5-b representing once more a constant factor. Comparative analysis of CS/ GR with Dynamic Level of Threshold (DLT) technique.

Supporting a set of QoS parameters and classes for all ATM services and minimizing network and end-system complexity cum maximizing the utilization of available resources is the major objectives of traffic control. The dynamic level of threshold falls under the priority scheme with bias for cell-loss. This has necessitated the treatment of high priority cells mostly at the expense of the lower priority ells. The provision for cell loss has therefore made it another platform to benchmark with CS/ GR.

The following parameters were used in the performance evaluation of CS/ GR and DLT scheme:

- Fairness: this factor marks the throughput for a source based on the demands by other sources. Under this factor, we looked at fairness and index criteria. Fairness Criteria: - Sharing - the available bandwidth is equally shared among connections. Other fairness criteria are: allocation proportional to Minimum Cell Rate (MCR) and weighted allocation. Fairness index: - this shows the share of bandwidth for each source to be equal to or converge to the optional value based on some optimality criterion. From the above, Dynamic Level of Threshold technique (DLT) is not fair enough as the fairness criteria and fairness index were not followed. Since high priority cells are always given attendance, even at the expense of the low priority cells, it is not therefore supported as being fair.
- Implementability: as shown again in Table 1, buffer requirement is enormous in the DLT schemes making it more expensive, and the possibility of wide range of implementation is not feasible since buffering is synonymous to “memories”. Therefore, DLT is more likely to be supported in a particular architecture. WRR/ AI and CS/ GR is however simple and easy to implement.

Table 1. Loss probability using numerical model of the DLT.

S/ N	Maximum Buffer Capacity	Loss probability for High priority Cells (HPC)	Loss probability for Low priority Cells (LPC)
1	500,000	0	0.0161264
2	600,000	0.00317328	0.140099
3	700,000	0.00107805	0.12311
4	800,000	0	0.108889
5	900,000	0	0.962204
6	1,000,000	0.00306701	0.0875698
7	1,100,000	0.00167837	0.0804963
8	1,200,000	0.000278299	0.0745131
9	1,300,000	0	0.0682759
10	1,400,000	0	0.0624293
11	1,500,000	0.00197023	0.0591605

- Handling of cell loss: this is another cogent factor to be considered, since the operation of the DLT is biased towards the high priority cells, low priority cells were lost and the delay associated was not properly addressed. DLT excels over the fixed level of threshold by dynamically addressing the rate of refusal of cells admittance at a particular threshold as stated during its implementation. However, strict adherence to the threshold in buffer space management does not make it very robust.

From the simulation results of CS/ GR presented earlier, Complete buffer sharing with gradual release performs better than the DLT in the following ways: fair treatment in buffer usage to both classes of service, elimination of undue starvation of low priority cells, possibility of coexistence of both class of service at every stage of transmission and maximum utilization of buffer space during the idle period of non-real time class.

Figure 6 shows the simulated result of DLT, which described the probability of finding high and low priority cells. From this figure, the true nature of the transmitting pattern of ATM networks comes to play as being deterministic and non-deterministic. The continuous streaming-in of the high priority cells representing the CBR in this paper affects the presence of the other classes tremendously. It could be seen that even the staggering amount of buffer presence could not help the gradual decline in the number of transmittable cells of low-priority cells as the presence of the high-priority does not usually reduce nor stop.

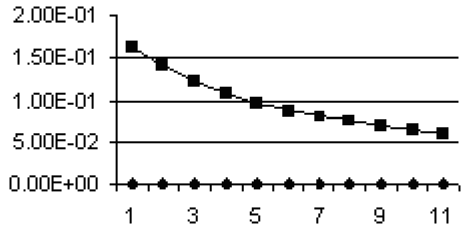


Figure 6. Probability of finding high and low priority cells using DLT.

In Figure 6, the low-priority cell graph tapering towards 0.05, which is likely going to be downward slopping in the presence of other types of cells in the buffer. The above has rubbished the concept of fairness-index and fairness-criteria as regards to bandwidth allocation and utilization in this scheme. However, CS/ GR presented a more balanced scenario with the issue of fairness index and criteria. CS/ GR guarantees that the knowledge of the transmission pattern is used in determining the buffer space allocation of the particular type of traffic.

To this end, any class is assured of its already negotiated space and probably more in the absence of other party not using its own allocation. In fact this flexibility/ workability surnamed the new method. There is an understanding that if a class is using another party's allocation, it will be gradually released as the demand comes for it. The above is presented in Figure 7, where the existence of both types of traffic under study is presented as a function of delay associated to the transmission pattern of delay associated to the transmission pattern of each class. The Figure shows that, within each transmitting cycle, each class of service is guaranteed a space and subsequently transmitted from the buffer.

The method of handling of cell-loss of DLT is sacrificial. This will thus cause loss of cells of the other class, which is a costly situation in an ATM networks. Imagine the loss of chunks of cells in transmission; this will usually be followed by possible retransmission which another cogent factor is leading to network congestion in ATM. Figure 8 presents the charts of the simulated result of DLT in handling cells loss. It is evident at this juncture that, the warping volume of available buffer could not stop the loss of cells in both types of transmission. The probability is very high for low-priority cells, which is understandable because of the favoritism displayed against it. However, the graph shows that there are little loss of cells on the case of the high-priority cells which is not a welcome development in buffer space management. Even if the low-priority cells could be sacrificed, its not the case with the high-priority class comprising of voice, and video, no matter how small, loss is not an expected situation here. CS/ GR was able to keep a constant zero-loss rate based on the buffer space allocation and management that is determined prior to their admittance into the system as calculated in equation 6.

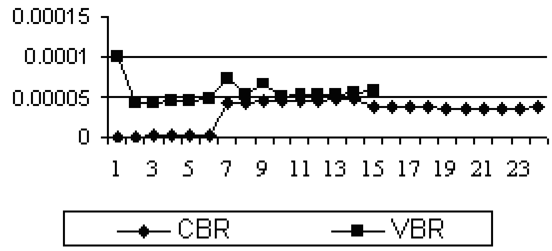


Figure 7. Existence of class 1 and 2 cells and value of their associated delay in CS/ GR.

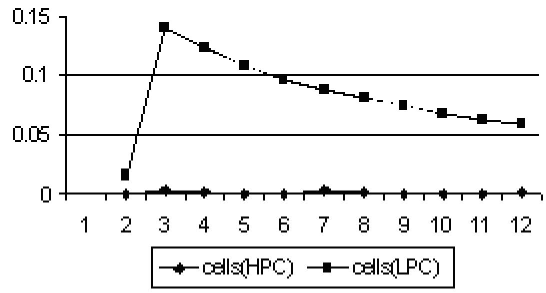


Figure 8. Loss probability using the numerical model of DLT.

5. Conclusion

The simulation result and comparative analysis has thus been presented in two modes: One is the performance evaluation of CS/ GR as a preferable buffer space management when compared with the DLT in the manner of handling of cell-loss based on the traffic parameters negotiated at the inception of transmission. The other aspect was majored on the design of a traffic control/ congestion control comprising of the duo of a scheduler and buffer space management technique. Varieties of scheduling schemes were tested based on CS/ GR in an attempt to coming-up with the best among them. The WRR/ AI is found to be the best in guaranteeing QoS as regards to delay and cell loss on the part of CS/ GR.

Finally, the paper has dealt with a major problem facing ATM network, i.e., buffer space management based traffic control/ congestion control for a scalable QoS in ATM networks. It is worth mentioning at this juncture that previous research work on buffer space management has been quite commendable, however, the newly developed technique presented a more versatile method of handling traffic congestion in ATM networks. The need for specifying the maximum amount of space a class of service can occupy vis-à-vis the assigning of a particular level of threshold was replaced by a much more dynamic pattern of buffer space management, thus enhancing optimum utilization of buffer space and guaranteeing negotiated quality of service to every concerned classes of service.

References

[1] Aderounmu G., Adagunodo E., and Akinde A., "Weighted Round-Robin Cell Scheduling in ATM Networks: A New Approach," *Journal of Computer Science and its Application*, vol. 7, no. 1, no. 66-72, 1999.

[2] Aderounmu G., Oyeniyi S., Adagunodo E., and Akinde A., "A New Buffer Management Scheme for Multi-QoS Traffic over ATM Switching Systems," *International Journal of Information and Computing Science*, vol. 3, no. 2, pp. 57-71, 2000.

[3] ATM Traffic Management Specification Version 4, http://www.atmforum.com/pages/aboutatmtech/atmhistory.html, 1996.

[4] Bac J. and Suda T., "Survey of Traffic Control Schemes and Protocols in ATM Networks," *IEEE*, vol. 79, no. 2, pp. 170-189, 1991.

[5] Chich-Chao W., Jim-Chyang J., Chorng-Horng Y., Shong C., and Wakid C., "Design of Real-Time QoS Management in ATM Networks," *in IEEE International Conference on Communications*, vol. 1, no. 1, pp. 228-232, 1998.

[6] Hung T. and Panwar S., "Optimal Buffer Control during Congestion in an ATM Network Node," *IEEE/ ACM Transactions on Networking*, vol. 2, no. 4, pp. 1059-1066, 1994.

[7] Jian R., "Congestion Control and Traffic Management in ATM Networks: Recent Advances and a Survey," *Computer Networks and ISDN Systems*, vol. 28, no. 13, pp. 1723-1738, 1996.

[8] Lizambri T., Duran F., and Wakid S., "Priority Scheduling and Buffer Management for ATM Traffic Shaping," *in Proceedings of 7th IEEE Workshop on Future Tends of Distributed Computing Systems (FTDCS'99)*, pp. 36-43, 2001.

[9] Saewoong B. and El-Zarki M., "Preventive Congestion Control Based Routing in ATM Networks," *in Proceedings of the 1994 IEEE International Conference Communications*, vol. 34, pp. 1592-1599, 1994.

[10] Shimonish H. and Suzuki H., "Analysis of Weighted Round Robin Cell Scheduling and its Improvement in ATM Networks," *IEICE Transactions on Communication*, vol. ESI-B, no. 8, pp. 910-918, 1998.

[11] Tatsuya S., "Asynchronous Transfer Mode (ATM): Networks: Fundamentals," *Encyclopedia of Electrical an Electronic Engineering*, Wiley and Sons, 1998.

[12] Vickers B., Kim B., Suda T., and Hong D., "Congestion Control and Resource Management in Diverse ATM Environments," *IECE Journal*, vol. J76-B-I, no. 11, pp. 759-774, 1993.

[13] Wu G. and Mark J., "A Buffer Allocation Scheme for ATM Networks: Complete Sharing Based on Virtual Partition," *IEEE/ ACM Transactions on Networking*, vol. 3, no. 6, pp. 660-670, 1995.

Francies Ogwu has been working as senior lecturer in the Department of Computer Science at the University of Botswana for the last seven years. He is also an adjunct associate professor at TU International, USA. He has been the head of the Department of Computer Science and Engineering at Obafemi Awolowo University Ile-Ife, Osun State, Nigeria. He has over 31 research papers in international journals and conferences. His area of interest is computer communications and networking, and artificial intelligence. He is specialized in speech processing. Presently, he is working on smellography.

Mohammad Talib is presently a senior lecturer in the Department of Computer Science of the University of Botswana and an adjunct associate professor at the TU International, USA. He was conferred upon the degree of the Doctor of Philosophy in computer science with specialization in computer vision from the prestigious University of Lucknow in

India with certificate of honor. Besides PhD, he is also flanked by an MSc in computer science, MSc in statistics and PG Diploma in Computing. He has supervised over a dozen Master and four PhD students in different areas of computer science, business and IT. His research areas include bioinformatics, computer vision, and robotics. Presently, he is working on a two way interactive video communication through the virtual screen with the essence of smellography.

Ganiyu Aderounmu holds a research degree MSc/ PhD in computer science from Obafemi Awolowo University, Ile-Ife, Nigeria in 1991 and 2001, respectively. He is a member of the Nigerian Society of Engineers (NSE) and is also a registered computer engineer with council for the regulation of engineering practice in Nigeria (COREN). He is also a member of Nigerian Computer Society (NCS) and Computer Professional registration council of Nigeria (CPN). His special interest includes education in Nigeria, curriculum development and computer communication and network. He is currently a senior lecturer and acting head of Department of Computer Science and Engineering of the same university.

Olufade Onifade is a lecturer in the Department of Computer Science at University of Ibadan in Nigeria. He has bachelor's degree in computer science which is followed by an MSc also in computer science discipline. His area of interest is computer networks and communications. He is also doing research in computer science.