

# A DYNAMIC BANDWIDTH ALLOCATION SCHEME FOR MULTIMEDIA DATA OVER ATM NETWORKS

Han Zhou  
Fidelity Investments, VSI group  
82 Devonshire Street A2A  
Boston, MA 02109 USA

C. H. Chang  
Electrical Engineering and Computer Science Department  
Tufts University  
Medford, MA 02155 USA

## ABSTRACT

In this paper, a new dynamic bandwidth allocation scheme called *Minimum Overflow Traffic Algorithm* (MOTA) is proposed to assign the bandwidth for each traffic class in the hierarchical admission control structure in ATM network. The traffic types can be video, image, voice and so on. An *Overflow traffic* function  $Q$  is used to adjust the bandwidth assignment each time when a new connection is required. This paper uses mean virtual cell loss probability (MCLP) as the quality of service (QoS) for the performance measurement of the system. This novel approach can minimize the mean virtual cell loss probability of the system by optimizing  $Q$  when the delay constraint is satisfied and provide high throughput in multi-class traffic environment.

**Keywords:** ATM Admission Control, Bandwidth Allocation, Cell Loss Probability, Cell Delay, Mean virtual cell loss probability, Overflow Traffic.

## 1. Introduction

The Asynchronous Transfer Mode (ATM) is widely recognized as the most promising technology for developing the broadband ISDN. ATM is capable of supporting the multimedia data streams, such as data, voice, image and video, which require wide range of bandwidth by using statistically multiplexing techniques. Due to the statistical nature of resource allocation in ATM, the certain performance requirements may not be satisfied unless resource allocation is applied by means of dynamic real time control. In order to efficiently manage the network resources, an architecture for a structured logical partition of network resources among traffic classes with different grade of service has been proposed in [2]. The data streams which

have the same traffic parameters belong to one traffic class. The bandwidth assigned to each class is based on the minimum capacity necessary to support  $L$  connections, which is under the hypothesis that the load generated by the  $L$  connections could be approximated by a variable with a Gaussian distribution. This strategy does not consider the cell loss probability as a grad of service. The capacity of ATM link is not fully utilized since it only gives the lower bound of the bandwidth to each class. The number of connections in a traffic class has to be large enough for Gaussian process approximation, however, it is not true in many cases.

In this paper, a new dynamic bandwidth allocation scheme called *Minimum Overflow Traffic Algorithm* (MOTA) is proposed to assign the bandwidth  $b_i$  to each traffic class in the hierarchical admission control structure. This algorithm sets the *overflow traffic* function  $Q$  as the optimization object function and adjusts the bandwidth assignment to each traffic class by optimizing  $Q$ . Since MCLP (Mean Cell Loss Probability) provides a good estimate of cell loss probability without respect to the burst lengths and is considered as a very promising service quality measure for ATM networks [5], we use it as the performance measure in this paper. The main advantage of this algorithm is that the mean cell loss probability of the system and  $Q$  will reach their minimum values in direct ratio under the delay constraint. The bandwidth allocation can be adjusted based on single call request, thus it can use bandwidth fairly and efficiently and it can provide high throughput in multi-class traffic environment.

## 2. Structure of the Bandwidth Allocation and Admission Control System

### 2.1 Admission Control System

The admission control functionality allows a network to decide whether to accept a new connection request or not on the basis of the current network performance, the grade of performance required for the service type and the traffic characteristics of the required connection. In order to efficiently manage the network resource and provide simple control schemes, a structured admission control system is described as below.

The overall access control and bandwidth assignment scheme is depicted in Figure 1.  $\lambda_{ji}$  is the peak traffic arrival rate and  $r_{ji}$  is the mean traffic density of connection  $i$  in class  $j$ ,  $i = 1 \dots L_j$ ,  $j = 1 \dots M$ .

The peak arrival rate of class  $i$  is  $F_i$ ,  $F_i = \sum_{n=1}^{L_i} \lambda_{in}$ . The

mean arrival rate of class  $i$  is  $S_i$ ,  $S_i = \sum_{n=1}^{L_i} r_{in} \lambda_{in}$ .  $b_i$  is

the bandwidth share for class  $i$ ,  $\sum_{i=1}^M b_i \leq C$ .

Upstream of the ATM channel is an multiplexer,  $M$  admission controllers (one for each traffic class) and a bandwidth allocation controller. When a new connection is required, bandwidth allocations are recomputed on-line by an allocation controller, whose goals reflect overall cell loss and refused traffic, as well as overall average delay. The multiplexer is assumed to operate synchronously. Each admission controller  $i$  implements a decision rule for the acceptance of incoming calls. The bandwidth allocation controller "sees" the average traffic density of each class,  $r_i$ , and the peak rate of each class,  $F_i$ ,  $i=1,2,\dots,M$ . It divides the total capacity  $C$  of ATM channel (in Mbit/s) into virtual capacities  $b_i$ ,  $i=1,2,\dots,M$ . The assignment is made on the basis of the dynamic variations in the traffic flows and the assignment should be fair to the all the admission controllers.

As shown in Figure 1, the ATM multiplexer is connected to  $M$  virtual multiplexers with length  $N_i$ . At the decision time, the local admission controller can "see" the virtual multiplexer with  $N_i$  and  $b_i$ . Each admission controller acts independently of the others. The acceptance decision is based on cell loss and delay requirements, respectively and simultaneously.

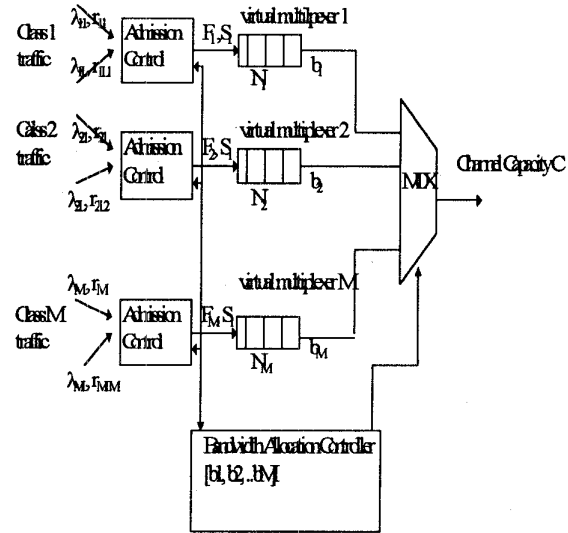


Figure 1.

## 2.2 The Individual Traffic Source Model

The virtual multiplexer depicted in Figure 1 can be viewed as a buffer receiving ATM cells with fixed length from  $L_i$  independent traffic sources which belong to the same traffic class. Each traffic source is characterized by its peak rate  $\lambda_i$  bits/second and average traffic density  $r_i$  bits/second, average active period of  $1/\alpha_i$  second, average silent period of  $1/\beta_i$  second (Figure 2). The cells are sent out from the buffer at rate  $b_i$  bits/second. It is assumed that a connection is modeled as a Bernoulli process when it is active. Let  $\Gamma_j$  be the probability that an active connection in traffic class  $j$  may generate a cell

during a slot interval,  $\Gamma_j = \frac{\bar{\lambda}_j}{b_j}$ , where  $\bar{\lambda}_j$  is the average peak rate in traffic class  $j$ . When a connection is idle it does not generate cells.

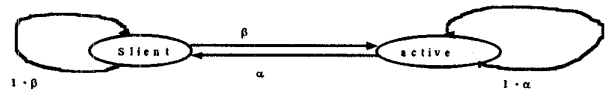


Figure 2

The steady-state probabilities  $W_i$  and  $W_a$  of a connection being idle and of a connection being active are given by the following expressions:

$$W_i = \frac{\alpha_i}{\alpha_i + \beta_i}; \quad W_a = \frac{\beta_i}{\alpha_i + \beta_i}$$

$W_a$  is also defined as mean traffic density  $r_i$  which is the average time fraction during that a connection is active. Thus, the mean arrival rate  $P_{ij} =$

$$\lambda_j r_i = \frac{\beta_i \lambda_j}{\alpha_i + \beta_i}.$$

### 3. Overflow Traffic Function

MOTA defines a overflow traffic function  $Q$  and uses  $Q$  as the bandwidth assignment optimization object function. The VCLP due to the lack of assigned bandwidth ( $b_i < R_i$ ) will trend to its minimum value with optimizing the function  $Q$ .

#### 3.1 The Optimization Object Function

The overflow traffic function is expressed in (1). It measures the expectation value of the overflow traffic to virtual link capacity ratio in the hierarchical admission control structure.

$$Q = \sum_{i=1}^M \sum_{n_i \in \{R_i - b_i > 0\}} P_i(n_i) \left( \frac{\sigma_i R_i - b_i}{b_i} \right) \quad (1)$$

$$P_i(n_i) = C_{L_i}^{n_i} (W_a)^{n_i} (W_i)^{L_i - n_i} \quad (2)$$

$$R_i = \sum_{j \in n_j} \lambda_j \quad (3)$$

Where  $\sigma_i$  is the weighting coefficient,  $\sigma_i > 0$ .  $P_i(n_i)$  is the probability of having  $n_i$  active connections out of  $L_i$  connections. For simplicity, we assume that the peak rates in one class is same,  $\lambda_{i1} = \lambda_{i2} = \lambda_{i3} = \dots = \lambda_i$ . The equation (1) and (3) can be rewritten as

$$Q = \sum_{i=1}^M \sum_{n_i \in \{R_i - b_i > 0\}}^{n_i=L_i} P_i(n_i) \left( \frac{\sigma_i R_i - b_i}{b_i} \right) \quad (4)$$

$$R_i = n_i * \lambda_i \quad (5)$$

The problem to be solved in bandwidth allocation controller is to optimize the function  $Q$ .

$$Q_{\min} = \min \sum_{i=1}^M \sum_{n_i \in \{R_i - b_i > 0\}}^{n_i=L_i} P_i(n_i) \left( \frac{\sigma_i R_i - b_i}{b_i} \right) \quad (6)$$

s.t.

$$\sum_{i=1}^M b_i \leq C \quad (6-1)$$

$$\text{CLP}(b_i) < \text{CLP}_{\max}(i), i = 1, 2, \dots, M \quad (6-2)$$

$$\text{Delay}(b_i) < \text{Delay}_{\max}(i), i = 1, 2, \dots, M \quad (6-3)$$

$$b_i \leq F_i = L_i \lambda_i \quad (6-4)$$

(6-1), (6-2), (6-3) and (6-4) are constraint conditions. Where  $\text{CLP}(b_i)$  is cell loss probability in class  $i$ .  $\text{Delay}(b_i)$  is the cell queuing delay.  $\text{CLP}_{\max}(i)$  is the maximum value of cell loss probability in class  $i$ .  $\text{Delay}_{\max}(i)$  is the maximum value of cell queuing delay in class  $i$ .

### 3.2 Discussion of Optimization Object Function

#### 3.2.1 Convexity of function $Q$

**Lamma 1:** The optimization object function  $Q$  is a strict convex function.

Proof:

The optimization object function is

$$Q = f(b_i) = \sum_{i=1}^M \sum_{n_i \in \{R_i - b_i > 0\}}^{n_i=L_i} P_i(n_i) \left( \frac{\sigma_i R_i - b_i}{b_i} \right) \quad (7)$$

The Hessian matrix of  $Q$  is expressed in equation (8)

$$H(f) = \begin{vmatrix} \frac{\partial^2 f}{\partial b_1^2} & \frac{\partial^2 f}{\partial b_1 \partial b_2} & \dots & \frac{\partial^2 f}{\partial b_1 \partial b_m} \\ \frac{\partial^2 f}{\partial b_1 \partial b_2} & \frac{\partial^2 f}{\partial b_2^2} & \dots & \frac{\partial^2 f}{\partial b_2 \partial b_m} \\ \dots & \dots & \dots & \dots \\ \frac{\partial^2 f}{\partial b_m \partial b_1} & \frac{\partial^2 f}{\partial b_m \partial b_2} & \dots & \frac{\partial^2 f}{\partial b_m^2} \end{vmatrix}$$

$$= \begin{pmatrix} \sum_{n_1 \in \{R_1 - b_1 > 0\}}^{n_1=L_1} P_1(n_1) \frac{2\sigma_1 R_1}{b_1^3} & 0 & \dots & 0 \\ 0 & \sum_{n_2 \in \{R_2 - b_2 > 0\}}^{n_2=L_2} P_2(n_2) \frac{2\sigma_2 R_2}{b_2^3} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \sum_{n_m \in \{R_m - b_m > 0\}}^{n_m=L_m} P_m(n_m) \frac{2\sigma_m R_m}{b_m^3} \end{pmatrix} \quad (8)$$

Where  $\sum_{n_i \in \{R_i - b_i > 0\}}^{n_i=L_i} P_i(n_i) \frac{2\sigma_i R_i}{b_i^3} > 0$ . We can

learn from (8) that the Hessian matrix of the optimization object function is a (m x m) diagonal matrix. According to convex function theory [8], if function  $f(x)$  has the second order derivative, its Hessian matrix is a diagonal matrix and every element in Hessian matrix is positive, then the Hessian is a positive-definite and  $f(x)$  must be a strict convex function at the convex set D,  $b_1, b_2, \dots, b_M \in D$ .

We assume that the constraints in (6-1), (6-2), (6-3 and (6-4) are on the convex sets. Condition (6-1) shows that the optimization will not cause ATM link over load, the capacity of a link will be always greater than or equal to the total traffic carried on. Condition (6-2) means that the optimization is carried out when CLP( $b_i$ ) is not greater than its upper bound  $CLP_{\max}(i)$ . Condition (6-3) means that the optimization is processed when Delay( $b_i$ ) is not greater than its limitation  $Delay_{\max}(i)$ .  $CLP_{\max}(i)$  and  $Delay_{\max}(i)$  are predetermined. Condition (6-4) limits the virtual link capacity under the peak rate of the class.

### 3.2.2 Mean virtual cell loss probability

**Definition 1:** The mean virtual cell loss probability is denoted by MCLP and defined as

MCLP = mean traffic overflow  
X/total traffic Y

where

$$X = \sum_{i=1}^M \sum_{n_i \in \{R_i - b_i > 0\}}^{L_i} P_i(n_i) (R_i - b_i) \quad (9)$$

$$Y = \sum_{i=1}^M S_i = \sum_{i=1}^M L_i \lambda_i \frac{\beta}{\alpha + \beta} \quad (10)$$

MCLP is based on the link overflow model in which cells are discarded when instantaneous total traffic load  $R_i$  exceeds the virtual link capacity  $b_i$ . It provides a good estimate of cell loss probability without respect to the burst lengths. MCLP is considered as a promising measure for the general grad of service of the system [9]. From (7) and (9) we can learn that Q and MCLP will be optimized in direct ratio. When  $F_i = b_i, i=1, \dots, M, Q=MCLP=0$  and the network has the best performance.

### 3.2.3 Performance Effect

One important characteristic of function Q is that its optimization will cause MCLP to be minimized. There is a set of  $\{b\}$  which generates the optimized  $Q_{\min}$ . If we apply this set of  $\{b\}$  to each class, MCLP of the system will trend to its minimum value and  $|MCLP - MCLP_{\min}| < \varepsilon$ . It is shown in simulation results.

Another important characteristic of function Q is that it can *fairly* allocate the network resources among these traffic classes. "Fairness" means that low and high bandwidth users should be treated equally. Function Q can avoid the high speed users "eat up" all the link capacity and assign relative more bandwidth to the classes who require low capacity and accept more connections in these classes. Therefore, the system will have high throughput. It is also shown in the simulation results.

## 4. Admission Control and Bandwidth Reassignment

An admission controller of traffic class  $i$  should decide to accept a new call in the node only if the bandwidth  $b_i$  assigned by allocation controller is sufficient to support the new connection and those already in progress to satisfy performance requirements in term of cell loss probability and cell delay. In this section, we will describe the calculation of cell delay and cell loss probability for one class and present a bandwidth assignment and admission control algorithm. The structured admission control scheme has the flexibility to measure the quality of service for different class. Each admission controller can simultaenously calculate individual CLP using different algorithms according the traffic characteristics. In this paper, we

consider the long-term time-averaged cell loss probability as CLP for every class.

#### 4.1 Calculation of Cell Delay and Cell Loss Probability

##### Cell Delay:

In this paper, we simply use maximum queuing delay as the cell delay upper bound. The maximum cell delay in class  $i$  can be expressed as

$$Delay(i) = \frac{q_{\max} CL}{b_i} \quad (11)$$

Where  $q_{\max}$  is maximum queue length in class  $i$ ,  $CL$  is the cell length (53 bytes for ATM) and  $b_i$  is the bandwidth assigned to the class. Since  $Delay(i) \leq Delay_{\max}(i)$ ,

$$b_i \geq \frac{q_{\max} L}{Delay_{\max}(i)} \quad (12)$$

##### Cell Loss Probability:

The long-term time-averaged value of cell loss probability for one traffic classes defined as

$$CLP = \sum_{n=0}^N P_{\text{loss}(n)} P_{n,L} \quad (13)$$

Where  $P_{\text{loss}(n)}$  represents the steady-state cell loss probability in a traffic class where  $n$  connections are active and  $P_{n,L}$  is the steady state probability of having  $n$  active connections out of  $N$  connections. According to the characteristics of the bandwidth assignment in MTOA, the probability that a traffic class may lose cells is

$$\zeta_i = (F_i - b_i)/F_i \quad 0 < \zeta_i < 1 \quad (14)$$

Let  $P_{(n)}$  be the steady-state value of the "instantaneous" cell loss probability, then we have

$$P_{\text{loss}(n)} = P_{(n)} \zeta_i \quad (15)$$

In order to calculate  $P_{(n)}$ , we first to calculate the probability of instantaneous cell loss for a single traffic class, adapting the derivation presented in [2] to our model.

With the traffic model and the system structure described in section 2, a queued cell in a virtual multiplexer is not necessarily served in the time interval considered. The probability that the cell relating to the traffic class  $i$  is served in the time unit

with  $\Omega_i$ .  $\Omega_i$  is equal to the ratio between the bandwidth allocated to the traffic class  $i$  being considered and the total capacity of the ATM link.

$$\Omega_i = \frac{b_i}{C} \quad (16)$$

On this basis, the probabilities of transition  $g_{kj}(n)$  from the state in which  $k$  cells of class  $i$  are queued in the  $m$ -th slot to the state in which  $j$  cells of the same class are queued in the  $(m+1)$ -th slot, given  $n$  active connections. Thus, we can have

If  $k \neq 0$

$$g_{kj}(n) = \begin{cases} 0 & j < k-1 \\ f_{j-k+1}(n)\Omega_i + f_{j-k}(n)(1-\Omega_i) & k-1 \leq j < Q_i \\ f_{Q_i-k}(n)(1-\Omega_i) + \sum_{s=Q_i-k+1}^n f_s(n) & j = Q_i \end{cases} \quad (17a)$$

If  $k = 0$  and  $\sum_k^j \equiv 0$  if  $k > j$

$$g_{kj}(n) = \begin{cases} f_0(n) + f_1(n)\Omega_i & j = 0 \\ f_{j+1}(n)\Omega_i + f_j(n)(1-\Omega_i) & 0 < j < Q_i \\ f_{Q_i}(n)(1-\Omega_i) + \sum_{s=Q_i+1}^n f_s(n) & j = Q_i \end{cases} \quad (17b)$$

where  $f_j(n)$  represents the probability that  $i$  out of  $n$  active connections in class  $i$  generate cells in the  $m$ -th slot.

$$f_j(n) = \begin{cases} 0 & j < 0 \\ \binom{n}{j} (\Gamma_i)^j (1-\Gamma_i)^{n-j} & 0 \leq j \leq n \end{cases} \quad (18)$$

We can express the probability distribution of queue length in slot  $m$  as

$$\Pi(m) = \{\pi_0(m), \pi_1(m), \dots, \pi_{Q_i}(m)\} \quad (19)$$

Where  $Q$  is the queue size. The probability distribution of queue length in slot  $m+1$  can be described by the recursive equation

$$\Pi(m+1) = \Pi(m)G \quad (20)$$

Where  $G$  is the matrix of the probabilities of transition

$$G = \begin{pmatrix} g_{00} & g_{01} & \cdots & g_{0Q} \\ g_{10} & g_{11} & \cdots & g_{1Q} \\ \cdots & \cdots & \cdots & \cdots \\ g_{Q0} & g_{Q1} & \cdots & g_{QQ} \end{pmatrix} \quad (21)$$

According to the quasi-stationary condition, for each fixed number of active connections  $n$ , the probability of instantaneous cell loss in slot  $m$   $P(m,n)$  will converge to its equilibrium value  $P(n)$ .

$$P(n) = \sum_{k=0}^Q \pi_k \left[ \frac{\sum_{j=0}^n \max(k+j-Q, 0) \alpha_j f_j(n)}{\sum_{j=0}^n j f_j(n)} + \frac{\sum_{j=0}^n \max(k+j-Q, 0) (1-\alpha_j) f_j(n)}{\sum_{j=0}^n j f_j(n)} \right] \quad (22)$$

where

$$\pi_i \equiv \lim_{m \rightarrow \infty} \pi_i(m) \quad (23)$$

is the steady state probability of having  $I$  queued cells, and is obtained by solving the following set of linear equations

$$\sum_{j=0}^{Q_i} \pi_j = 1 \quad (24)$$

So, from (2), (13), (15) and (22) we can have

$$CLP = \sum_{n=1}^N C_n (W_n)^{N-n} \sum_{i=0}^Q \pi_i \left[ \frac{\sum_{j=0}^n \max(i+j-Q, 0) \alpha_j f_j(n)}{\sum_{j=0}^n j f_j(n)} + \frac{\sum_{j=0}^n \max(i+j-Q, 0) (1-\alpha_j) f_j(n)}{\sum_{j=0}^n j f_j(n)} \right] \quad (25)$$

When  $F_i = b_i$ ,  $i = 1, \dots, M$ ,  $CLP = 0$  in every class, thus the network has the best performance. It also proves

that the network will obtain the best quality of service when  $Q \rightarrow 0$ .

#### 4.2 Bandwidth Assignment Algorithm

The bandwidth allocation controller will calculate  $\{b\}$  when there is a new connection request. This paper proposes a algorithm: Minimum  $Q$  algorithm which searches the approximated optimal  $Q$  value.

##### Minimum $Q$ Algorithm (MQ)

This algorithm is based on searching minimum  $Q$  value which expresses the mean overflow traffic to virtual link capacity ratio. This algorithm can be described as

$$Q_{\min} = \min \sum_{i=1}^M \sum_{n_i \in \{R_i - b_i > 0\}}^{n_i = L_i} P_i(n_i) \left( \frac{\sigma_i R_i - b_i}{b_i} \right)$$

Let  $m_i = \max\left[\frac{q_{\max} L}{Delay_{\max}}, \rho_i\right]$ , we can write the conditions as

$$m_i \leq b_i \leq F_i \text{ and } \sum_{i=1}^M b_i \leq C.$$

Where  $\rho_i$  is the minimum capacity which is needed to support the active connections by guaranteeing their performance requirement. For large number of connections the requirement is defined as the probability that the loading of the class is over the maximum permitted load  $G_i$  below a given threshold  $\tau$  and the traffic load can be approximated by Gaussian process.

$$\text{Prob}\{\text{loading of class } i > G_i\} < \tau$$

According to [5], let us assume that channel loading arises from a large number of sources multiplexed so that the loading of the channel can be approximated as a Gaussian process and  $\tau$  is  $10^{-8}$ , then  $\rho_i$  can be described as:

$$\rho_i = \frac{5.243(\lambda_i \sqrt{L_i r_i (1-r_i)} + L_i \lambda_i r_i)}{G_i} \quad (27)$$

where  $r_i$  is traffic density (average time fraction during which a connection is active) in class  $i$ . For relative small number of connections, the Gaussian approximation is not correct, therefore, we use average traffic load as the minimum capacity,  $\rho_i = S_i$ .

Assume that the precise of the algorithm is  $q$  (kbps).  $b_i$  can be one of these values

$$[m_i, m_i + q, m_i + 2q, \dots, m_i + Nq]$$

where  $m_i + Nq \leq F_i$ , so  $N \leq \frac{F_i - m_i}{q}$ . This

algorithm will compute every possible value  $b_i$  in  $M$

classes under the condition  $C\theta \leq \sum_{i=1}^M b_i \leq C$ ,  $0 < \theta$

$< 1$ , and find the minimum  $Q$ . If the worst case is

considered, the time complexity for searching all the  $b_i$  in class  $i$  is  $O(\frac{F_i - m_i}{q})$ , the time complexity for

searching all the  $b_i$  in  $M$  classes is  $O(\prod_{i=1}^M \frac{F_i - m_i}{q})$ .

## 5. Simulation Results

Some simulation experiments have been conducted in order to verify and test the bandwidth allocation algorithm proposed. The simulations were intended to present interaction and relationships between traffic blocked function  $Q$  and mean cell loss probability as well as the number of connections and network performance in term of mean cell loss probability and individual cell loss probability. There are two traffic classes, voice class and video conference class, in this simulation model. The total ATM link capacity is 30 Mbps and the precise of bandwidth assignment is 0.1 Mbps. The queue size for each class is 10 cells. The number of connections in voice class, class 1, is 100 and it is characterized by  $\lambda_1 = 64$  Kbps and  $W_a = 0.5$ . Its maximum utilization  $G_1 = 0.9$ . The number of connections in video conference class, class 2, is 20 and it is characterized by  $\lambda_2 = 2$  Mbps and  $W_a = 0.4$ . Its maximum utilization  $G_2 = 0.9$ .

The plot in Figure 3 shows that overflow traffic function  $Q$  and mean virtual cell loss probability MCLP are minimized in direct ratio.

Figure 4 presents the effect of  $\sigma$  on bandwidth shares in class 1. When  $\sigma$  increases, class 1 will obtain larger bandwidth shares.

Figure 5 and 6 show the relationship between number of connections in two classes and MCLP. We compared MOTA algorithm and GA algorithm. The results in Figure 5 show that MCLP decreases when number of connections in class 1 increase because GA does not optimize MCLP. In Figure 6, MOTA has much better performance than GA does.

Figure 7 and 8 describe the relationship between number of connections (L1 and L2) and individual CLP in two classes. The results also show that MOTA's performance is superior to GA's.

## 6. Conclusions

In this paper, we have introduced a new bandwidth allocation strategy for hierarchical admission control system. The goal of the strategy is to minimize the mean overflow traffic ratio over the virtual link capacity. The main advantages of the algorithm are to minimize the mean cell loss probability in the system and fairly assign bandwidth to each class. We also propose a admission control rule which is integrated with the bandwidth allocation scheme.

## References

- [1] H. Akimaru, T. Okuda and K. Nagai, "A simplified performance evaluation for bursty multiclass traffic in ATM systems." *IEEE Trans. on Comm.* Vol. 42, No. 5, may 1994.
- [2] R. Bolla, F. Danovaro, F. Davoli and M. Marchese, "An integrated dynamic resource allocation scheme for ATM networks." *INFOCOM'93*.
- [3] D. T. Han, H. Zhou and C. H. Chang, "Fairness Optimization Communication Networks with Throughput/Delay Constraints." *ICCCN'94*, San Francisco.
- [4] S. Q. Li and J. W. Mark, "Traffic characterization for integrated services networks," *IEEE Trans. on comm.* vol. 38, no. 8, Aug. 1990.
- [5] A. Lombardo, S. Palazzo and D. Panno, "Admission control over mixed traffic in ATM networks." *International Journal of digital and analog comm. systems.* Vol. 3, 155-159.
- [6] N. M. Marafih, R. L. Pickholtz and Y. Q. Zhang, "New approximation for analyzing the multiplexing of bursty sources in ATM networks." to be published.
- [7] P. Papalambros and D. J. Wilde, "Principles of optimal design - modeling and computation." Cambridge university press, 1988.
- [8] M. D. Prycker, "Asynchronous Transfer Mode - solution for broadband ISDN." Ellis Horwood, 1993.
- [9] K. Sriram and W. Whitt, "Characterizing superposition arrival process in packet multiplexer for voice and data." *IEEE JSAC*, vol. SAC-4, No. 6 September 1986.

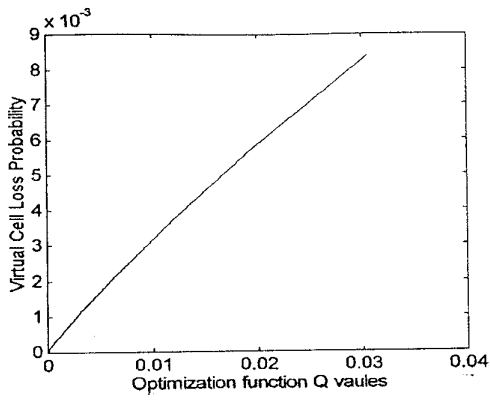


Figure 3. Overflow Traffic Function Q vs MCLP

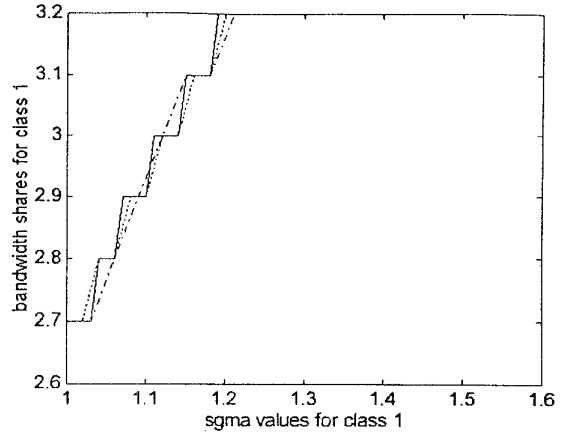


Figure 4. Coefficient  $\sigma$  vs bandwidth shares

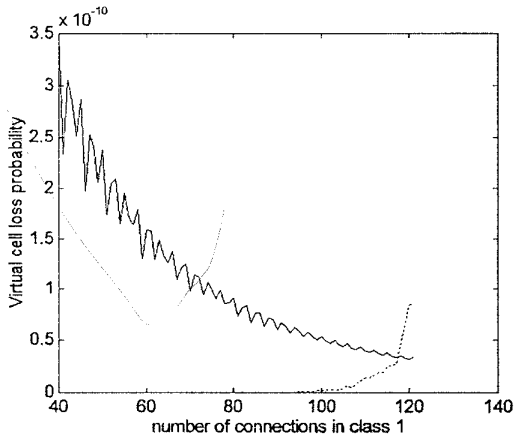


Figure 5. Connections in class 1 vs MCLP

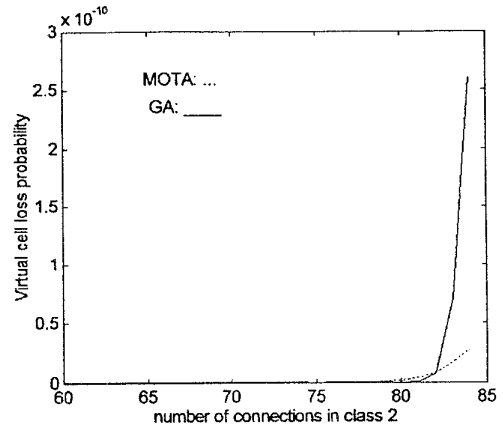


Figure 6. Connections in class 2 vs MCLP

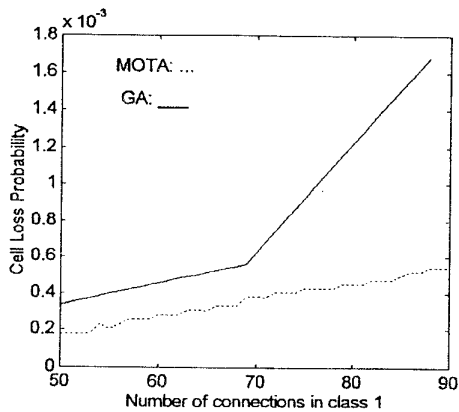


Figure 7. Connections in class 1 vs CLP

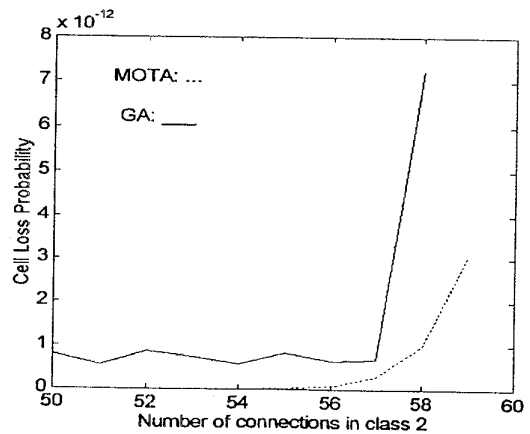


Figure 8. Connections in class 2 vs CLP