The Effective Bandwidth of Markov Modulated Fluid Process Sources with a Generalized Processor Sharing Server

Shiwen Mao   Shivendra S. Panwar

Dept. of ECE, Polytechnic University,
6 MetroTech Center, Brooklyn, NY 11201

George Lapiotis

Telcordia Technologies,
445 South Street, Morristown, NJ 07960

Abstract— **Generalized Processor Sharing (GPS) is an important scheduling discipline because it enables bandwidth sharing with work conservation and traffic isolation properties. While Markov Modulated Fluid Processes (MMFP) capture the dynamics of the sources, the analysis of such sources with a GPS server is difficult because of the large state space. In this paper, we study a multi-queue GPS system with MMFP classes and propose a scalable, low complexity algorithm for the tail distributions of the logical queues. The effective bandwidth of the classes and a simple Connection Admission Control (CAC) scheme are derived. Numerical results illustrate the efficiency and accuracy of the technique. The application to an example system of classes consisting of voice and variable-bit-rate (VBR) video traffic is included.**

Keywords— **Generalized Processor Sharing, Markov Modulated Fluid Process, Effective Bandwidth.**

I. Introduction

The main objective for next generation networks is the accommodation of a variety of services with different traffic characterizations and Quality of Service (QoS) requirements. Scheduling is used in switches and routers to enforce service differentiation. Among the scheduling disciplines, Generalized Processor Sharing (GPS), along with its variants such as Weighted Round Robin, has such desirable properties as bandwidth sharing and fairness provisioning. It is an efficient and controllable mechanism to support different service classes.

GPS has been widely studied [1]-[6]. Most previous work take bounding approaches and focuses on general arrival processes, with deterministic or stochastic settings. GPS systems are studied with various source characterizations, such as Poisson with symmetric service sharing [1], leaky bucket regulated sources [2], exponential bounded burstiness source [3], and long-tailed sources [4]. By the notion of *feasible ordering/partitioning*, a GPS system can be transformed to a priority system and hence performance bounds are obtained [2][3]. A queue decomposition technique is proposed in [3] to decouple the service. These results are generally expected to be loose since the finer dynamics of the sources are not exploited [5]. In this paper we study a multi-queue GPS system with Markov Modulated Fluid Process (MMFP) classes. It is well known that data traffic is *self-similar*, which poses a great challenge to network control and QoS provisioning. On the other hand, Markovian processes have been widely used to model network traffic and are efficient in modeling Voice over IP traffic. Previous work shows that *Long Range Dependent* traffic, such as VBR video, can be adequately approximated by *Short Range Dependent* traffic models for traffic engineering purposes [7][8][9]. This makes it possible to investigate certain aspects of the impact on the performance of the long-range correlation structure within the confines of traditional Markovian analysis [10].

In order to analyze a GPS system with MMFP sources, two basic questions should be answered:
1. How to de-couple the service, since the service a class receives depends on not only its own queue, but also all other logical queues?
2. How to design a scalable algorithm to handle a large number of classes, since with MMFP sources, the state space scales up exponentially with the number of classes?

In this paper, we introduce a scalable technique for deriving the tail distribution of a MMFP class in GPS systems. First we adopt the queue decoupling technique used in [5]. We get the departure processes of all other queues using their guaranteed service [11], and use them to modulate the service of the queue under study. This is equivalent to a *MMFP/MMFP/1* system. Instead of solving this system directly, we apply a technique to transform the *MMFP/MMFP/1* to the more tractable *MMFP/D/1* queue. Then the *Chernoff-Dominant eigenvalue* approximation in [11] can be applied and the tail of the class is found. Large systems, which are very hard to handle with exact analysis, can be solved asymptotically with manageable computational effort using this technique. The tails can be calculated with reasonable accuracy and far less computational complexity. We also derive the Effective Bandwidth of the MMFP classes based on the approximation and a simple Connection Admission Control (CAC) scheme is designed.

The rest of the paper is organized as follows: in Section II, the system model and background, including a brief review of the key results in [5] and [13] which we will use, are presented. In Section III we introduce the scalable algorithm, and derive the effective bandwidth of the MMFP classes with a GPS server. Numerical studies are presented in Section IV. Section

This work was supported in part by the New York State Center for Advanced Technology in Telecommunications (CATT) at Polytechnic University, and also in part by the National Science Foundation under Grants 0081527 and 0081357.

V presents our future work and conclusions.

II. Preliminaries and Background

A. The System Model

GPS is a work conserving scheduling discipline in which N traffic classes share a deterministic server with rate c [2]-[6]. There is a set of parameters ω_i, $i = 1, \ldots, N$, called GPS *weights*. Each class is guaranteed a service rate $g_i = \omega_i c$, and the residual service of the non-backlogged classes is distributed to the backlogged classes in proportion to their weights.

The system is shown in Fig.1. Each class i, with instant rate $r_i(t)$, is modeled as a MMFP with state space $\mathcal{S}_i$, rate matrix $\mathbf{\Lambda}_i$, and infinitesimal generator $\mathbf{M}_i$. $\mathbf{M}_i$ governs the transitions between the states and $r_i(t) = \mathbf{\Lambda}_i(s,s)$ when the source is in state s. Assume $\sum_i \overline{\lambda}_i \leq c$, which guarantees the ergodicity of the system; and $\overline{\lambda}_i < g_i$, $i = 1, \ldots, N$, where $\overline{\lambda}_i$ is the average rate of class i. The buffer is infinite, with each class having its own logical queue with occupancy $X_i(t)$.

We are interested in the tail distribution of class i's queue occupancy, which upper bounds the loss class i experiences in a finite buffer system and can also be used to bound its delay distribution.

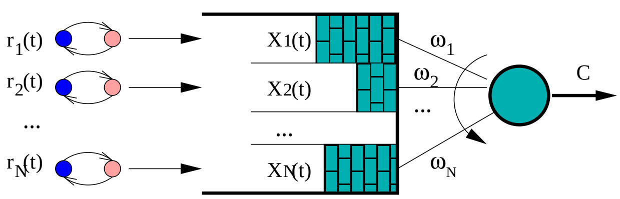

Fig. 1. The system model

B. A Service Bound

In [5] Lo Presti, Z. Zhang and D. Towsley proposed an approximate lower bound for the service that a class receives in a multi-queue GPS system with MMFP classes (LZT bound), which de-couples the correlated GPS service.

It takes two steps to obtain the LZT bound.

Step 1 : The departure process of each class, $r'_j(t)$, is obtained while assuming the service rate is g_j. Note that the technique in [11] is used to characterize the departure processes.

Step 2 : For each class i, the queue length distribution can be derived from a *MMFP/MMFP/1* queue as shown in Fig.2, with input r_i and modulated service:

$$s'_i \equiv g_i + \sum_{j \neq i} \frac{\omega_i}{\sum_{k \neq j} \omega_k} (g_j - r'_j(t)) \qquad (1)$$

The modulated service s'_i, which is also a MMFP, consists of the guaranteed service rate g_i and the residual service rate seen by class i, while assuming all other classes are busy. By solving the de-coupled system, an upper bound on the distribution of the buffer occupancy of class i can be found [5].

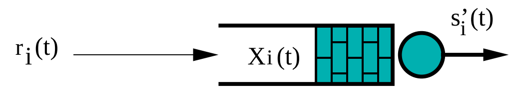

Fig. 2. The LZT bound model

It should be noted that the LZT bound could be loose in some cases [6]. Its tightness is determined by:

1. how accurately the departure processes are modeled;
2. the number of classes in the system.

Note that g_i, which is a lower bound of the service class i receives, is used to characterize the departure process. The output process is also an approximation where the underload states are modeled exactly but all busy states are aggregated into *one* busy state [11]. Furthermore, (1) is an approximation of the GPS dynamics which in general becomes looser when the number of classes increases. We gave a tighter bound in [6], which models the finer dynamics of GPS service sharing and generally will not get looser when the number of classes is large. Unfortunately, the technique we discuss here cannot apply to that bound directly.

In [11], it is argued that the output characterization is accurate when the non-empty buffer probability is lower than 10^{-3}, which is not atypical for real-time traffic with a stringent delay requirement. Also in the context of *Diffserv* [15], the flows are aggregated and the number of classes is usually small. For such systems, we believe the LZT bound can be reasonably tight for the cases of interest.

C. The Chernoff-Dominant Eigenvalue Approximation for FIFO Systems

Elwalid and Mitra derived the effective bandwidth of MMFP sources in a *MMFP/D/1* queue in [13]. Let the system have N MMFP input sources, and each source i have rate matrix $\mathbf{\Lambda}_i$ and generator $\boldsymbol{M}_i$, $i = 1, \ldots, N$; and the aggregated traffic be characterized by rate matrix $\mathbf{\Lambda}$ and generator $\boldsymbol{M}$. To obtain the system's dominant eigenvalue, the following inverse eigenvalue problem is solved:

$$c\phi = (\mathbf{\Lambda} - \frac{1}{z}\boldsymbol{M})\phi \qquad (2)$$

Since both $\mathbf{\Lambda}$ and $\boldsymbol{M}$ can be decomposed into *Kronecker* sums of the rate matrices and generators of the individual classes, (2) has *Kronecker property*. The dominant eigenvalue z_d of the system is derived by solving:

$$\sum_{i=1}^{N} g_1^i(z_d) = c \qquad (3)$$

where $g_1^i(z)$ is the maximum real eigenvalue (MRE) of the i^{th} inverse eigenvalue sub-problem, i.e., $g_1^i(z) = MRE(\mathbf{\Lambda}_i - \frac{1}{z}\boldsymbol{M}_i)$. With large buffer sizes, the tail distribution can be approximated by:

$$G(x) = Pr(X \geq x) \cong L_o e^{z_d x} \qquad (4)$$

where L_o is the loss in a bufferless multiplexing system, which compensates the statistical gain of multiplexing a large number of sources [11][12].

III. The Scalable Algorithm and the Effective Bandwidth of MMFP Sources

A. The Scalable Algorithm

The exact analysis of the system in Fig.2 is difficult since the dimension of the state space increases exponentially with number of sources. For example, the generator of the aggregated traffic of N heterogeneous on-off sources is of the order $(2^N, 2^N)$. In the case that there is a large number of states, the huge matrices of the system are too unwieldy to handle and numerical problems encountered often make it unsolvable. Hence it is desirable to further decompose the de-coupled system. In this section, we will derive an equivalent system based on the LZT bound that is much easier to solve.

Without loss of generality, assume: $\sum_{i=1}^{N} \omega_i = 1$. Define $\omega_i / \sum_{k \neq j} \omega_k \equiv \sigma_j$, then (1) becomes:

$$s_i^{'} = g_i + \sum_{j \neq i} \sigma_j g_j - \sum_{j \neq i} \sigma_j r_j^{'}(t)$$

Note that the first two terms on the right hand side are state independent. The drift of logical queue i is:

$$\begin{aligned} \frac{d}{dt} X_i(t) &= r_i(t) - s_i^{'}(t) \\ &= r_i(t) + \sum_{j \neq i} \sigma_j r_j^{'}(t) - (g_i + \sum_{j \neq i} \sigma_j g_j) \end{aligned}$$

Define sources: $r_i^* = r_i(t)$, $r_j^*(t) = \sigma_i r_j^{'}(t)$ for all $j \neq i$, and the deterministic service rate $c_i^* = g_i + \sum_{j \neq i} \sigma_j g_j$, then:

$$\frac{d}{dt} X_i(t) = \sum_{j=1}^{N} r_j^*(t) - c_i^* \qquad (5)$$

This is identical to the system equation of a *MMFP/D/1* queue with service rate c_i^*, fed by the scaled sources $r_j^*(t)$, $j = 1, \ldots, N$, as shown in Fig.3.

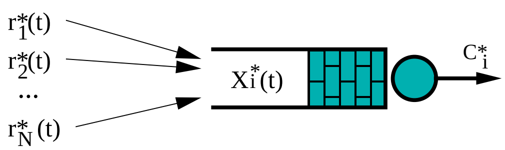

Fig. 3. The equivalent model

B. The Chernoff-Bound Prefactor L_o

The system in Fig.3 has *Kronecker property* and can be decomposed as described in Section II-C. Hence the tail distribution of class i's queue occupancy is of the form Eq.(4). It still remains to determine the prefector L_o.

One straightforward scheme for computing L_o then would be apply the technique in [13] to the scaled equivalent system of Eq.(5). The "hybrid scheme" proposed in [12] can also be applied. Assume the system in Fig.3 have aggregated average rate $\overline{\lambda}$, aggregated rate λ_k and equilibrium probability π_k when the system is in state k, L_o is:

$$L_o^i = \frac{1}{\overline{\lambda}} \sum_{k:\lambda_k > c_i^*} \lambda_k \pi_k (1 - \frac{c_i^*}{\lambda_k}) \qquad (6)$$

C. The Effective Bandwidth Of MMFP Sources

With the results in the previous subsections, and as an extension of [13], we get the effective bandwidth of class i as follows:

Proposition I: Let class i have buffer allocation *B*, and acceptable loss rate p. The *Chernoff Bound* prefactor is L_o and define $\eta = log(p/L_o)/B$. If the equivalent system in Fig.3 has rate matrix Λ^* and generator M^*, the effective bandwidth of class i, c_i^e, is:

$$c_i^e = MRE(\Lambda^* - \frac{1}{\eta} M^*) \qquad (7)$$

The effective bandwidth of class i defined here is more complicated than that proposed in [13]. In the latter case, c_i^e is the function only of the source parameters and its QoS requirements, i.e., it is independent on other classes within the system. GPS provides the ability to differentiate the classes, but also introduces correlation among the classes. In GPS systems, c_i^e not only depends on class i itself, but also depends on all other classes in the system. However, c_i^e still sheds some light on the property of the GPS system, and may provide heuristic clues to the GPS weight design and admission control.

Equation (5) requires much less computational effort than solving the system in Fig.2 directly. In both cases the output processes for the classes should be obtained first. The saving is from the dominant eigenvalue and *Chernoff bound* computation, because we process much smaller matrices. For example, if all classes are *on-off*, the modulated service in (1) has 2^{N-1} states. So the complexity of getting the dominant eigenvalue is of $O(2^N)$ for Fig.2. For the equivalent system, the results in Section II-C can be applied straightforwardly. The dominant eigenvalue z_d is found by numerically solving (3) with iterations, resulting in a complexity of $O(N)$.

A straightforward CAC policy can be:

```
for i:=1 to N do
  if c_i^e > c_i^*
  then output NO and stop;
endfor;
output YES and stop.
```

IV. Numerical Investigations

In this section we present numerical studies with our technique. The results shown here illustrate the quantitative trends of the technique.

Consider a system with 3 classes where each class consists of K *on-off* sources. The service rate is c. The source parameters are given in Table I, where α and β are the transition rates from off to on, and on to off, respectively; λ is the input rate when the source is on. Table II gives the service rate c and the GPS weights of the classes used in Fig.4-7, where ω_i is the GPS weight of class i and load $= (\overline{\lambda}_1 + \overline{\lambda}_2 + \overline{\lambda}_3)/c$. Note that while the examples considered here consists of classes with homogeneous sources, the technique also applies to the case of a class consisting of heterogeneous sources. These bursty sources are also used in [11], and they are used in modeling Voice over IP for traffic engineering purposes [6][14].

TABLE I
Source parameters used in Fig.4 through Fig. 7

-	α	β	λ
class 1	0.4	1.0	1.00
class 2	0.4	1.0	1.20
class 3	1.0	1.0	0.61

TABLE II
GPS Weights of the Classes in Figures Fig.4 through Fig. 7

-	**c**	**load**	ω_1	ω_2	ω_3	K
Fig. 4	15.1	61.8%	0.33	0.40	0.27	10
Fig. 5	41.4	67.7%	0.32	0.38	0.30	30
Fig. 6	120.1	77.7%	0.33	0.40	0.27	100
Fig. 7	130.1	66.0%	0.33	0.40	0.27	100

Fig.4-6 give the analytical results of the tail distributions of the classes, compared with fluid simulations. We increase the number of sources in each class from 10 to 100, and increase the load from 61.8% to 77.7%. Observe that the analytical results are conservative except for the very small buffer region. The analysis becomes more conservative as load increases. Table III gives the numerical values of the decay rates of the tails obtained by simulation and analysis in Fig.4. They are very close to each other, implying that this technique is asymptotically accurate. We also give the slopes of class 3 in Table IV. It is shown that the analysis decay rates are very close to the simulations in all the cases.

Although the analytical decay rates match the simulations very well, there is still some gap between the analytical and simulated tail curves. This is caused by the approximation error in output characterization and the LZT service bound. An accurate output characterization requires low buffer non-empty probability, i.e., low loads. Even for load as high as 77.7%, the tail distributions are accurate enough for engineering purposes. Better techniques, which can improve the pre-factor L_o or use multiple-term approximation of the tails [16], can make the result more accurate. In Fig.7, we calculate L_o using (6) while keep all other settings the same as Fig.6. It can be seem that the analysis is much closer to the simulation results.

For the tail distribution of class 1 in Fig.4, both the rate and generator matrix used for solving the system in Fig.2 directly are of size 616×616. It takes several minutes to compute the dominant eigenvalue using a computer with a *Pentium*© III 450*MHz* CPU. With our technique, we process three much smaller matrices, which are of sizes 11×11, 7×7, and 8×8, respectively. Consequently, it takes far less time to get the tails. Table IV gives the time in seconds used to compute the tail distributions of class 3 for larger systems. Although the system consists as many as 300 sources, it takes the algorithm reasonable time to finish the computations. In some cases, the sources within a class are homogeneous, the state space can be reduced by aggregate equivalent states into one state and exact analysis can be applied. However, in the cases that the classes consist of heterogeneous sources, which is not atypical in *Diffserv* where different sources are aggregated into a small number of classes [15], exact analysis is almost impossible. For example, it is impossible to define a matrix with size $(2^{30}, 2^{30})$ needed to analyze a system with 30 heterogeneous sources using *MATLAB V5.2* because the product of dimensions is greater than the maximum integer value.

TABLE III
Slopes of $Log_{10}G(x)$ in Fig.4

-	**Simulation**	**Analysis**
Class 1	-0.8936	-0.7900
Class 2	-0.6839	-0.6479
Class 3	-1.9886	-1.9183

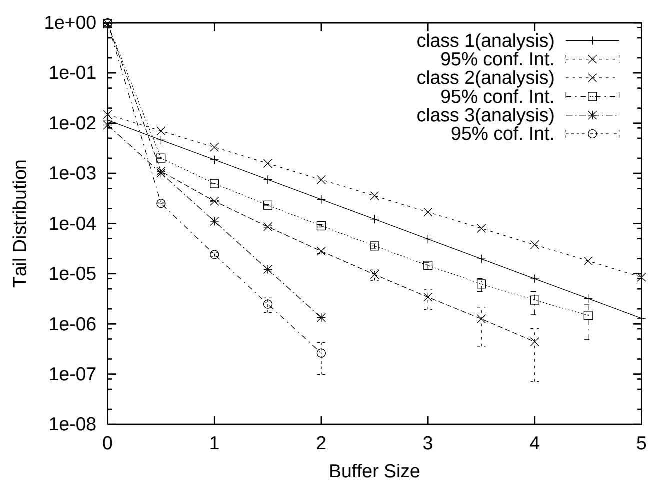

Fig. 4. Tail distributions of three classes, while each class has 10 on-off sources and $c = 15.1$.

Next we study a 3-queue GPS system with two video classes and a voice class. The video model is from [9], in which four-state Discrete-time Markov Modulated Poisson Process (DMMPP) is used for video traffic modeling. A video source

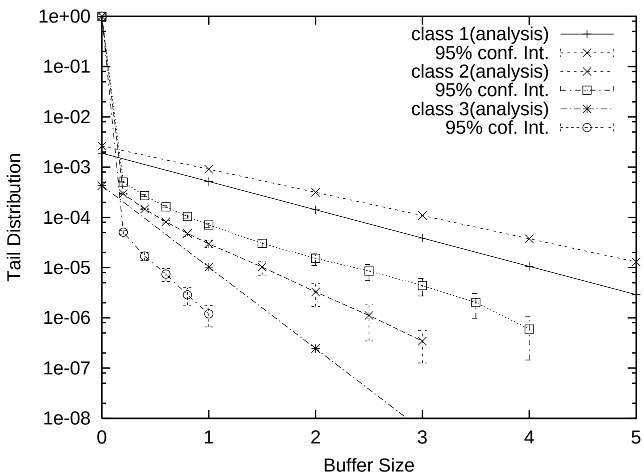

Fig. 5. Tail distributions of three classes, while each class has 30 on-off sources and $c = 41.4$.

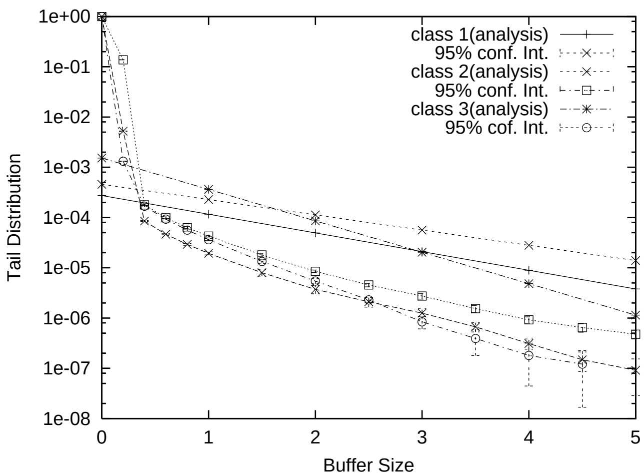

Fig. 6. Tail distributions of three classes, while each class has 100 on-off sources and $c = 120.1$.

has transition matrix P and rate vector V as given below. Note $\widehat{X} = (1 - X)$ in P. The parameters for the video sources are given in Table V. We use their fluid equivalent for the analysis and simulations. Class 1 consists of 3 type 1 video sources, while class 2 consists of 3 type 2 video sources. Class 3 consists of 20 voice sources, where each is an on-off source with λ =8.0, α =0.56, and β =0.83. The GPS weights for the classes are 0.41, 0.46, and 0.13, respectively. The service rate is 518.3. Fig.8-10 give the tail distributions of the classes. Again, the analysis results match the simulation results closely.

$$P = \begin{bmatrix} \widehat{\alpha_1}\widehat{\alpha_2} & \widehat{\alpha_1}\alpha_2 & \alpha_1\widehat{\alpha_2} & \alpha_1\alpha_2 \\ \widehat{\alpha_1}\beta_2 & \widehat{\alpha_1}\widehat{\beta_2} & \alpha_1\beta_2 & \alpha_1\widehat{\beta_2} \\ \beta_1\widehat{\alpha_2} & \beta_1\alpha_2 & \widehat{\beta_1}\widehat{\alpha_2} & \widehat{\beta_1}\alpha_2 \\ \beta_1\beta_2 & \beta_1\widehat{\beta_2} & \widehat{\beta_1}\beta_2 & \widehat{\beta_1}\widehat{\beta_2} \end{bmatrix}$$

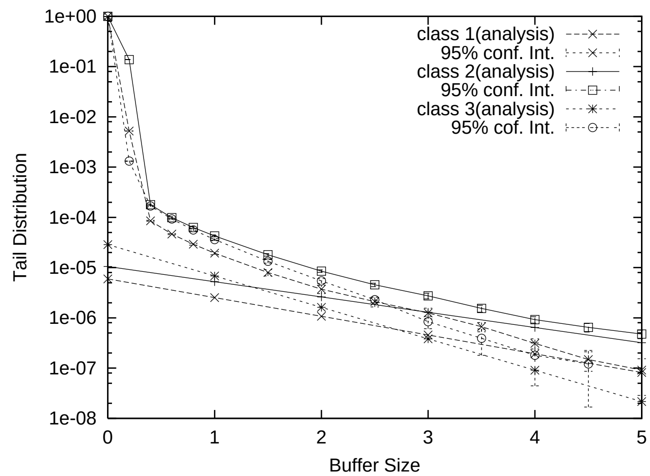

Fig. 7. Tail distributions of three classes, while each class has 100 on-off sources and $c = 120.1$. L_o is obtained using (6).

TABLE IV

SLOPE OF $Log_{10}G(x)$ FOR CLASS 3 AND THE TIME USED TO COMPUTED THE TAILS IN THE FIGURES

-	**Fig. 4**	**Fig. 5**	**Fig. 6**
Analysis	-1.918	-1.622	-0.625
Simulation	-1.988	-1.916	-0.681
Time(seconds)	0.691	2.093	8.157

$$V = [\ \lambda_1 \quad \lambda_2 \quad \lambda_3 \quad \lambda_4\ ]$$

TABLE V

PARAMETERS OF THE VIDEO SOURCES

Parameters	**Type 1**	**Type 2**
α_1	0.02	0.0018
β_1	0.08	0.00064
α_2	0.002	0.1568
β_2	0.008	0.0234
λ_1	30	95.24
λ_2	40	58.9
λ_3	80	73.92
λ_4	90	37.58

V. CONCLUSION

In this paper we present a simple and scalable analytical technique for determining the tail distributions of MMFP sources in a GPS system. The effective bandwidth of the MMFP class is derived and a CAC scheme is designed. Numerical results illustrate the accuracy and efficiency of the technique. More experiments to validate the CAC scheme based on the technique will be presented in our future work.

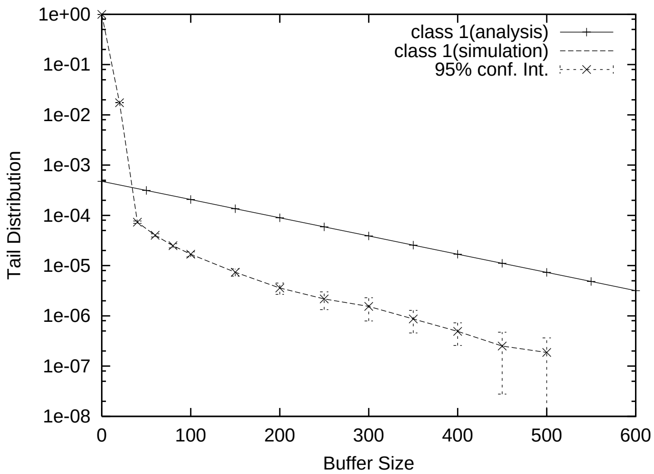

Fig. 8. Tail distributions of class 1 (3 Type 1 video sources) in the three classes system with $c = 518.3$.

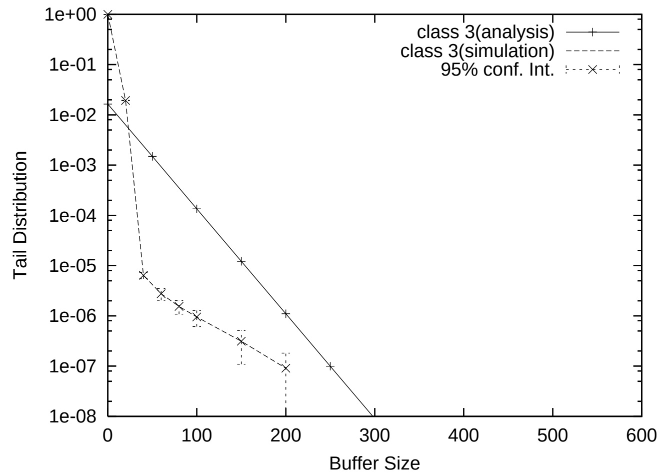

Fig. 10. Tail distributions of class 3 (20 voice sources) in the three classes system with $c = 518.3$.

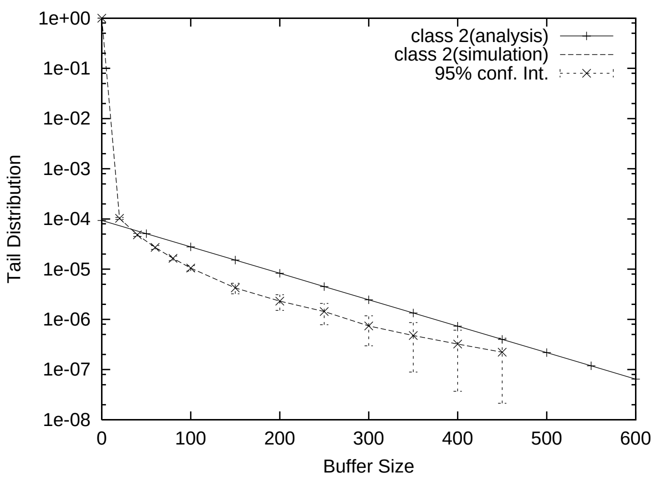

Fig. 9. Tail distributions of class 2 (3 Type 2 video sources) in the three classes system with $c = 518.3$.

References

[1] A. G. Konheim, I. Meilijson, and A. Melkman, "Processor-sharing of two parallel lines," *J. Appl. Prob.*, 18, pp:952-956, 1981.

[2] D. Nandita, J. Kuri, and H. S. Jamadagni, "Optimal call admission control in generalized processor sharing (GPS) schedulers," *Proc. of IEEE INFOCOM 2001*, pp:468-477, 2001.

[3] Z. Zhang, D. Towsley, and J. Kurose, "Statistical analysis of generalized processor sharing scheduling discipline," *IEEE Journal on Selected Area in Communications*, vol. 13, No. 6, pp. 1071-1080, August 1995.

[4] S. Borst, O. Boxma, and P. Jelenkovie, "Generalized Processor Sharing with Long-Tailed traffic sources," *Proc. of ITC-16*, pp.345-354, June, 1999.

[5] F. Lo Presti, Z. Zhang, and D. Towsley, "Bounds, approximations and applications for a two-queue GPS system," *Proc. of IEEE INFOCOM'96*, March 1996. An extended version can be found at: http://www.cs.umass.edu/csinfo/techrep.html.

[6] G. Lapiotis, S. Mao, and S. S. Panwar, "GPS Analysis of Multiple Sessions with Applications to Admission Control," *Proc. IEEE ICC 2001*, June 2001.

[7] D. Heyman and T.V. Lakshman, "Source models for VBR broadcast video traffic," *IEEE/ACM Trans. Networking*, vol. 5, No. 1, pp.40-48, February 1996.

[8] D. Heyman and T. V. Lakshman, "Statistical analysis and simulation study of video teleconference traffic in ATM networks," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 2, No. 1, pp:49-59, March 1992.

[9] T. C. Yu and J. A. Silvester, "A four-state DMMPP for characterization multimedia traffic with short-term and long-term correlation," *Proc. Of IEEE ICC'99*, vol. 2, pp.880-885, September 1999.

[10] K. Park, W. Willinger, *Self-similar network traffic and performance evaluation*, John Wiley & Sons, Inc., 2000.

[11] A. Elwalid and D. Mitra, "Analysis, approximation and admission control of a multi-service multiplexing system with priorities," *Proc. of IEEE INFOCOM'95*, pp.463-472, 1995.

[12] N. B. Shroff and M. Schwartz, "Improved Loss Calculations at an ATM Multiplexer," *IEEE/ACM Trans. Netwroking*, vol. 6, No. 4, August 1998.

[13] A. Elwalid and D. Mitra, "Effective bandwidth of general markovian traffic sources and admission control of high speed networks," *IEEE/ACM Trans. Networking*, vol. 1, No. 3, pp.329-343, June 1993.

[14] J. N. Daigle and J. D. Langford, "Models for analysis of packet voice communications systems," *IEEE Journal on Selected Areas in Communications*, vol.SAC 4, No.6, pp:847-855, September 1986.

[15] Y. Bernet, J. Binder, S. Blake, M. Carlson, B. Carpenter, S. Kenshav, E. Davies, B. Ohlman, D. Verma, Z. Wang, and W. Weiss, "A framework for differentiated services," *Internet draft*, 1999.

[16] G. L. Choudhury, D. M. Lucantoni, and W. Whitt, "Squeezing the most out of ATM," *IEEE Trans. on Communications*, vol.44, No.2, pp.203 217, February 1996.