Inter-Class Resource Sharing using Statistical Service Envelopes

Jing-yu Qiu and Edward W. Knightly
Department of Electrical and Computer Engineering
Rice University

Abstract— **Networks that support multiple services through "link-sharing" must address the fundamental conflicting requirement between *isolation* among service classes to satisfy each class' quality of service requirements, and statistical *sharing* of resources for efficient network utilization. While a number of service disciplines have been devised which provide mechanisms to both isolate flows and fairly share excess capacity, admission control algorithms are needed which exploit the effects of inter-class resource sharing. In this paper, we develop a framework of using *statistical service envelopes* to study inter-class statistical resource sharing. We show how this service envelope enables a class to over-book resources beyond its deterministically guaranteed capacity by statistically characterizing the excess service available due to fluctuating demands of other service classes. We apply our techniques to several multi-class schedulers, including Generalized Processor Sharing, and design new admission control algorithms for multi-class link-sharing environments. We quantify the utilization gains of our approach with a set of experiments using long traces of compressed video.**

I. INTRODUCTION

Future integrated services networks will support heterogeneous Quality of Service (QoS) specifications and traffic demands. For example, a deterministic service [1] uses worst-case resource allocation to support applications requiring packet delivery without losses or delay bound violations; a statistical service [2] achieves a statistical multiplexing gain and provides statistical QoS guarantees with controlled "over-booking" of resources; a measurement-based service [3] supports QoS by basing admission control decisions on empirical observations of aggregate traffic behavior; best-effort services support applications with less stringent QoS requirements such as bulk data transfer. With appropriate admission control and traffic scheduling, these services and others can co-exist in a single network, as admission control limits the number of admitted traffic flows to ensure that each class' QoS requirements are met, and packet schedulers ensure that packets are assigned the priority levels needed to meet their QoS objectives.

In a link sharing environment as outlined in [4], traffic class k is allocated capacity c_k such that whenever packets from class k are backlogged, the class receives service at a rate of at least c_k. If class k is not backlogged, then class k's unused capacity is distributed fairly among backlogged sessions. Consequently, classes can be assured to meet their respective QoS requirements, regardless of the behavior of other traffic classes, allowing any number of services to co-exist in the network.

In the literature, a number of service disciplines have been designed to support such link sharing objectives [4], [5], [6]. For example, [5] develops a class of Hierarchical Packet Fair Queueing algorithms focusing on an algorithm's fairness, complexity, and ability to provide low end-to-end deterministic delay bounds. While scheduling algorithms for efficiently and fairly allocating excess capacity to backlogged classes are an important aspect of a link-sharing network, an admission control policy that enables one class of traffic to quantify the improved QoS it will receive due to capacity unused by other classes has not been addressed.

This work was supported by NSF CAREER Award ANI-9733610, NSF Grant ANI-9730104, and Nokia Corporation. The authors may be reached via http://www.ece.rice.edu/networks.

In addition to service disciplines, a number of admission control algorithms have also been designed both for deterministic services which do not exploit statistical resource sharing [7], [8], as well as statistical [2], [9], [10], [11], [12], [13] and measurement-based services [3] which do. However, such admission control algorithms consider traffic classes in isolation, and while a statistical multiplexing gain is achieved *within* a particular traffic class, *inter-class* resource sharing is not addressed. In particular, [10], [13] study statistical service for Generalized Processor Sharing (GPS) [8], and while the "isolation" property of GPS is exploited, inter-class statistical resource sharing is not addressed. Moreover, while [12] allows video on demand systems to exploit statistical gains from real time traffic flows, it does not address general link sharing environments.

In this paper, we address the problem of inter-class statistical resource sharing. Our key technique is to develop a framework of *statistical service envelopes* to study the problem. Inspired by [7], [14], we define a statistical service envelope as a probabilistic description of the service available to a traffic class as a function of interval length. We use this service envelope to characterize the additional capacity available to a traffic class beyond the minimum deterministically guaranteed capacity set aside by the link sharing rules. In this way, we statistically capture the fluctuating excess capacity left unused by one traffic class so that another class may exploit an inter-class statistical multiplexing gain and potentially admit additional traffic flows that would not otherwise have been deemed admissible. Thus, we use the statistical service envelope as a tool for overbooking inter-class resources in a controlled manner, so that a class can probabilistically quantify the additional resources available in a link sharing environment.

We apply this framework of statistical service envelopes to two multi-class service disciplines, namely, Static Priority (SP) and link-sharing GPS [4], [5]. We show that while the concept of a statistical service envelope was implicitly used in previous studies of SP [11], explicitly computing the service envelope of other traffic classes provides a simpler analysis and allows us to uniformly treat deterministic and statistical service classes.

For GPS, we conceptually partition traffic classes into *isolation* classes and *sharing* classes depending on whether or not the traffic class will exploit the effects of inter-class resource sharing in making admission control decisions. For example, a deterministic service is an isolation class as excess capacity from other traffic classes is not guaranteed in the worst case and hence a statistical envelope of excess capacity cannot improve this class' admissible region. We then bound the service received by a traffic class with an *arbitrary* partition of the classes, and show that the above partition into isolation and sharing classes can tightly approximate the statistical service envelope obtained by the sharing classes. In this way, each *sharing* class can characterize the capacity available beyond its guaranteed rate, incorporating the relative weights and traffic demands of all other traffic classes, and improving the class' admissible region.

We illustrate the potential utilization gains of our inter-class resource sharing scheme with a set of trace-driven simulation experiments using long traces of MPEG-compressed video. As an illustrative example with a 45 Mbps link supporting equally weighted deterministic and statistical service classes with the GPS service discipline, we find that the average utilization of the link can be improved from 47.7% to 84.6% by using the statistical service envelope to characterize the excess capacity of the deterministic class.

II. Statistical Service Envelopes: Theory and Applications

In this section, we define *statistical service envelopes* and develop their applications to inter-class resource sharing. In particular, we first study the delay distribution for a single class using statistical traffic envelopes and deterministic service envelopes. Next, we extend this analysis to include statistical traffic envelopes and *statistical* service envelopes. Finally, we illustrate the application of statistical service envelopes by deriving admission control tests for SP schedulers using this theory.

A. Single Class Queueing Model

Throughout this paper, we model a multiplexer by a discrete-time infinite buffer queue in which fluid flows into and out of the buffer only at discrete intervals. For traffic class i, let X_k^i denote its aggregate arrivals in time slot k, and let $X_{j,k}^i$ denote the total arrivals between slots j and k, such that $X_{j,k}^i = \sum_{t=j}^k X_t^i$. Let Y_k^i represent the amount of fluid served for traffic class i in time slot k, and denote $Y_{j,k}^i$ as the total fluid served between time slots j and k, such that $Y_{j,k}^i = \sum_{t=j}^k Y_t^i$.

Denoting Q_k^i as the backlog of traffic class i at the end of time slot k, Q_k^i is obtained from the Lindley recursion as,

$$Q_k^i = \max_{j \le k} \{X_{j,k}^i - Y_{j,k}^i\} \tag{1}$$

where the maximum will be reached at j if $Q_{j-1}^i = 0$.

B. Deterministic Service Envelopes

Deterministic service is studied in [7] using *deterministic* service envelopes and *deterministic* traffic envelopes. Here, we first study statistical service with *statistical* traffic envelopes and *deterministic* service envelopes, and later focus on *statistical* service envelopes. First, we formally define both deterministic and statistical traffic envelopes and service envelopes. We refer to an interval $[j,k]$ as class i's *backlogged interval* if $Q_m^i > 0$, for $m = j, \cdots, k$.

Definition 1 (Available Service) For given input process $X_{j,k}^m$ of all traffic classes except i, we define the available service $\widetilde{Y}_{j,k}^i$ as the output of the ith class in the interval $[j,k]$ given a minimally backlogging input process $\widetilde{X}_{j,k}^i$, which is defined as the minimal class i input such that class i is continuously backlogged throughout interval $[j,k]$.

Note that available service $\widetilde{Y}_{j,k}^i$ is a function of the scheduling mechanism and input process $X_{j,k}^m$, $m \neq i$, and it is independent to the input process in class i; whereas the *actual* output process $Y_{j,k}^i$ is decided by *all* classes' inputs. By using this notation of available service, we decouple the impact of class i's input on $Y_{j,k}^i$, and make $\widetilde{Y}_{j,k}^i$ a pure description of available network resources, separate from the traffic that is actually sent.

Definition 2 (Deterministic Service Envelope) [1] A non-decreasing non-negative function $s^i(t)$ is a deterministic service envelope of traffic class i, if for any interval $[j+1, j+t]$, the available service satisfies[2]

$$\widetilde{Y}_{j+1,j+t}^i \ge s^i(t).$$

To illustrate the concept of a deterministic service envelope, note that for a FCFS server with capacity C, $\widetilde{Y}_{j+1,j+t} = s(t) = Ct$. In a GPS server, a service class with guaranteed rate g^i, satisfies $\widetilde{Y}_{j+1,j+t}^i \ge s^i(t) = g^i t$.

Definition 3 (Deterministic Traffic Envelope) [15] A non-decreasing non-negative function $b^i(t)$ is a deterministic traffic envelope of class i, if for any interval $[j+1, j+t]$, the input traffic satisfies

$$X_{j+1,j+t}^i \le b^i(t).$$

Definition 4 (Statistical Traffic Envelope) [2] A sequence of random variables $B^i(t)$ is a statistical traffic envelope of class i, if for any interval $[j+1, j+t]$, the input traffic satisfies

$$X_{j+1,j+t}^i \le_{st} B^i(t).$$

where $X_{j+1,j+t}^i \le_{st} B^i(t)$ (stochastic inequality) denotes $P[X_{j+1,j+t}^i > z] \le P[B^i(t) > z]$ for all z.

Denoting D_k^i as the virtual delay experienced by a bit arriving at time slot k, the key QoS metric that we consider is the probability of delay bound violation, $P[D^i > d_0]$. As long as

$$\lim_{t\to\infty} \frac{EX_{1,t}^i}{t} < \lim_{t\to\infty} \frac{s^i(t)}{t}$$

[1] This definition is a slight generalization of the one in [7].

[2] With abuse of notation, $Y \ge c$ for a constant c denotes $P(Y \ge c) = 1$.

(the stability condition), and X_k^i is stationary and ergodic, $P[D_k^i > d_0]$ converges to a steady state tail probability $P[D^i > d_0]$.

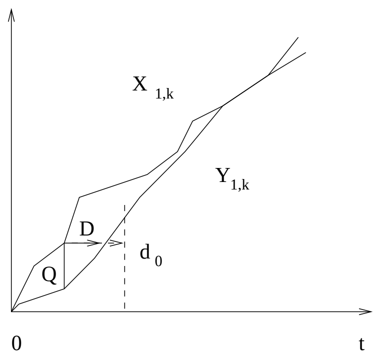

Fig. 1. Delay and Buffer Occupancy

Figure 1 shows the delay and buffer occupancy in terms of $X_{1,k}$ and $Y_{1,k}$ if the buffer is initially empty. The virtual delay D_k is defined as [7]

$$D_k = \min\{\Delta : \Delta \geq 0 \text{ and } X_{1,k} \leq Y_{1,k+\Delta}\}. \tag{2}$$

Lemma 1: For a delay bound d_0, the event of delay bound violation in class i at time slot k satisfies

$$\{D_k^i > d_0\} \quad \subseteq \quad \{\max_{j \leq k}\{X_{j,k}^i - \widetilde{Y}_{j,k+d_0}^i\} > 0\}. \tag{3}$$

Proof. By definition

$$\begin{aligned}\{D_k^i > d_0\} \quad &\equiv \quad \{X_{1,k}^i - Y_{1,k+d_0}^i > 0\} \\ &= \quad \{\max_{j \leq k}\{X_{j,k}^i - Y_{j,k+d_0}^i\} > 0\}.\end{aligned}$$

Observe that if $\max_{j \leq k}\{X_{j,k}^i - Y_{j,k+d_0}^i\} > 0$, then $\max_{j \leq k}\{X_{j,k}^i - \widetilde{Y}_{j,k+d_0}^i\} > 0$. This is because if $\max_{j \leq k}\{X_{j,k}^i - Y_{j,k+d_0}^i\} > 0$, there must exist an

$$s = \max\{j : j < k \text{ and } Q_j^i = 0\}$$

such that

$$\max_{j \leq k}\{X_{j,k}^i - Y_{j,k+d_0}^i\} = X_{s+1,k}^i - Y_{s+1,k+d_0}^i,$$

$[s+1, k+d_0]$ is a backlogged interval of class i, and

$$\widetilde{Y}_{s+1,k+d_0}^i \leq Y_{s+1,k+d_0}^i,$$

since $\widetilde{Y}_{s+1,k+d_0}^i$ is the minimum backlogged service. Thus

$$\{D_k^i > d_0\} \subseteq \{\max_{j \leq k}\{X_{j,k}^i - \widetilde{Y}_{j,k+d_0}^i\} > 0\}. \qquad \square$$

Theorem 1: For a service class i, with deterministic service envelope $s^i(t)$ and statistical traffic envelope $B^i(t)$, the tail probability of $P[D^i > d_0]$ is given by

$$P[D^i > d_0] \leq P[\max_{t \geq 0}\{B^i(t) - s^i(t+d_0)\} > 0]. \tag{4}$$

Proof. $P[D_k^i > d_0]$ converges to $P[D^i > d_0]$. From Equation (3),

$$P[D_k^i > d_0] = P[\max_{j \leq k}\{X_{j,k}^i - \widetilde{Y}_{j,k+d_0}^i\} > 0]. \tag{5}$$

From Definition 4 and Definition 2,

$$\begin{aligned}&\max_{j \leq k}\{X_{j,k}^i - \widetilde{Y}_{j,k+d_0}^i\} \\ &\qquad \leq_{st} \max_{j \leq k}\{B^i(k-j+1) - s^i(k+d_0-j+1)\}\end{aligned}$$

such that

$$\begin{aligned}&P[\max_{j \leq k}\{X_{j,k}^i - \widetilde{Y}_{j,k+d_0}^i\} > 0] \\ &\qquad \leq P[\max_{t}\{B^i(t) - s^i(t+d_0)\} > 0]. \qquad \square\end{aligned}$$

C. Statistical Service Envelopes

Theorem 1 enables us to exploit the statistical multiplexing gain of flows within a service class. While the deterministic service envelope $s^i(t)$ provides isolation among service classes and simplifies admission control, it precludes statistical inter-class resource sharing. In multi-class schedulers such as SP and GPS, the utilization gains available from exploiting inter-class resource sharing can be significant. Next we introduce a statistical service envelope to study the inter-class resource sharing problem, and develop new theory to calculate the delay bound violation probability using statistical service envelopes.

In a multi-class server, the available service for class i, $\widetilde{Y}_{j,k}^i$, is a function of the input traffic in other classes, and of the particular service discipline which specifies how to schedule services among competing classes. The interference among classes is reflected in $\widetilde{Y}_{j,k}^i$, and in some cases, it is possible that the available service is far greater than the minimally guaranteed service, i.e., $\widetilde{Y}_{j,k}^i \gg s^i(k-j+1)$. Thus we define a statistical service envelope to describe the available service beyond the deterministically guaranteed $s^i(t)$.

Definition 5 (Statistical Service Envelope) A sequence of random variables $S^i(t)$ is a statistical service envelope of traffic in class i, if for any interval $[j+1, j+t]$, the available service satisfies

$$\widetilde{Y}_{j+1,j+t}^i \geq_{st} S^i(t).$$

Notice that while a deterministic service envelope $s^i(t)$ describes the service of a class in isolation, the statistical service envelope $S^i(t)$ describes inter-class resource sharing. We employ $S^i(t)$ in the delay distribution calculation with the following theorem.

Theorem 2: For a service class i, with statistical service envelope $S^i(t)$ and statistical traffic envelope $B^i(t)$, the tail probability of $P[D^i > d_0]$ is given by

$$P[D^i > d_0] \leq P[\max_{t \geq 0}\{B^i(t) - S^i(t+d_0)\} > 0]. \tag{6}$$

Proof. From Equation (5),

$$P[D_k^i > d_0] = P[\max_{j\le k}\{X_{j,k}^i - \widetilde{Y}_{j,k+d_0}^i\} > 0]. \quad (7)$$

From Definition 4 and Definition 5,

$$\begin{aligned}&\max_{j\le k}\{X_{j,k}^i - \widetilde{Y}_{j,k+d_0}^i\}\\ &\quad \le_{st} \max_{j\le k}\{B^i(k-j+1) - S^i(k+d_0-j+1)\}\end{aligned}$$

so that

$$\begin{aligned}&P[\max_{j\le k}\{X_{j,k}^i - \widetilde{Y}_{j,k+d_0}^i\} > 0]\\ &\quad \le P[\max_t\{B^i(t) - S^i(t+d_0)\} > 0]. \qquad \square\end{aligned}$$

Below we employ Theorem 2 to devise admission control algorithms for multi-class servers that exploit inter-class statistical resource sharing.

D. Static Priority

Admission control for static priority schedulers was studied in [11], [16], here we approach the problem using service envelopes.

Consider an SP scheduler with N priority queues, link speed C, and the aggregate traffic in class i bounded by $B^i(t)$ and $b^i(t)$, with $i = 1, \ldots, N$ denoting the priority level from higher priority to lower priority. The statistical service envelope for class i is

$$S^i(t) = (Ct - \sum_{j=1}^{i-1} B^j(t))^+ \quad (8)$$

The deterministic service envelope for class i is

$$s^i(t) = (Ct - \sum_{j=1}^{i-1} b^j(t))^+ \quad (9)$$

where $b^i(t) = \sum_{j\in C_i} b_j(t), B^i(t) = \sum_{j\in C_i} B_j(t)$, and $b_j(t)$ and $B_j(t)$ are the statistical and deterministic envelopes of the jth flow in class i.

Lemma 2: Consider an SP scheduler with N priority queues and link speed C. For each service class, traffic is bounded by $B^i(t)$ and $b^i(t)$, with QoS parameters (d^i, P^i), where d^i is the delay bound, and P^i is the delay bound violation probability. The QoS for all service classes in this multi-service SP scheduler is satisfied if for all deterministic service classes with $P^i = 0$,

$$\max_t\{b^i(t) + \sum_{k=1}^{i-1} b^k(t+d^i) - C(t+d^i)\} \le 0$$

and for all statistical service classes with $P^i > 0$,

$$P[\max_t\{B^i(t) + \sum_{k=1}^{i-1} B^k(t+d^i) - C(t+d^i)\} > 0] \le P^i.$$

Proof. For statistical service classes, Equation (8) gives

$$\begin{aligned}&B^i(t) - S^i(t+d^i)\\ &\quad \le_{st} B^i(t) - C(t+d^i) + \sum_{k=1}^{i-1} B^k(t+d^i),\end{aligned}$$

and applying Theorem 2 requires $P[\max_t\{B^i(t) - S^i(t+d^i)\} > 0] < P^i$. Thus, if $P[\max_t\{B^i(t) + \sum_{k=1}^{i-1} B^k(t+d^i) - C(t+d^i)\} > 0] \le P^i$, then the statistical service in the ith service class is satisfied. For deterministic service classes, the proof is similar. $\square$

Note that inter-class interference in an SP scheduler is in a single direction, only from higher priority classes to lower priority ones. For GPS, we will see that every class affects every other class such that the statistical service envelope for one class becomes a function of the traffic envelopes and relative weights of all other classes.

III. Inter-Services Resource Sharing in Link-Sharing GPS

In Section II, we developed tools for managing multi-class services using statistical service envelopes, considering SP as a specific example. Here we study a link-sharing GPS server, again using the framework of statistical service envelopes, with a goal of increasing the total utilization of the multi-class GPS server by exploiting inter-class resource sharing.

A. Generalized Processor Sharing

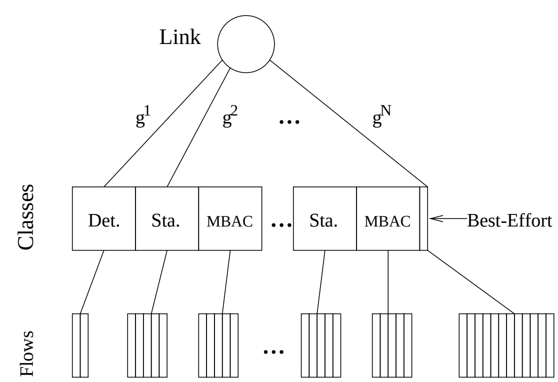

Fig. 2. System Model for Admission Control

Figure 2 shows the system model for admission control in a multi-class GPS scheduler (see [4] for example). There are N service classes in the system, each allocated a weight ϕ^i. Each service class provides either deterministic, statistical, measurement-based, or best-effort services.[3] The admission control algorithm should admit a new flow only if the QoS of all classes can be satisfied. This multi-class service model can also support flow-based services, in which some service classes serve only one flow. Without considering inter-class resource sharing,

[3]Here, we study multiple deterministic and statistical service classes and leave study of measurement-based service to future work.

one could view each service class as a FCFS server with capacity g^i, which is the guaranteed service rate $g^i = \frac{\phi^i}{\sum_m \phi^m} C$, as defined by the GPS service discipline. However, while exploiting this isolation property of GPS simplifies admission control, it does not corporate potential utilization gains due to inter-class statistical sharing.

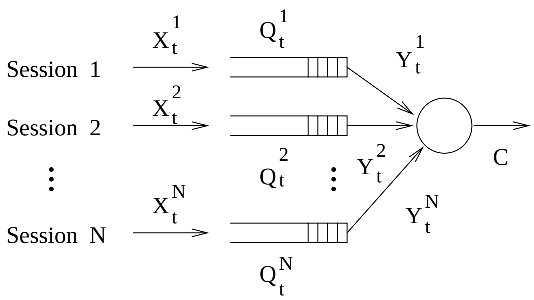

Fig. 3. GPS System

Figure 3 illustrates the GPS system in the view of inputs, outputs and buffers. The aggregate traffic in each class is viewed as a session, and the notation for inputs, outputs and queues are as defined in Section II.

For $1 \le i \le N$, let $Y^i_{j,k}$ be the amount of class i traffic served during $[j, k]$. By definition of GPS,

$$\frac{Y^i_{j,k}}{Y^m_{j,k}} \ge \frac{\phi^i}{\phi^m}, m = 1, 2, \ldots, N \tag{10}$$

for any class i backlogged during $[j, k]$. Since each class has a guaranteed rate g^i whenever it is backlogged, the deterministic service envelope of class i is $s^i(t) = g^i t$.

B. Statistical Service Envelopes in GPS

Our goal is to calculate the statistical service envelope for class i which is a lower bound for class i's available service. First, we lower bound class i's service in a backlogged interval as follows.

We define

$$\Theta^i_{j,k}(A_1) = \frac{\phi^i}{\sum_{m \in A_1} \phi^m}[C \cdot (k - j + 1) - \sum_{n \in A_2} Y^n_{j,k}], \tag{11}$$

with the N classes separated into two arbitrary subsets, A_1 and A_2, such that $A_1 \cup A_2 = \{1, \cdots, N\}$, $A_1 \cap A_2$ is an empty set, and $i \in A_1$.

From Equation (10), we have

$$Y^i_{j,k} \ge \Theta^i_{j,k}(A_1) \tag{12}$$

if class i is backlogged throughout $[j, k]$. This property enables us to estimate the backlog service for class i, using $\Theta^i_{j,k}(A_1)$ with an arbitrary partition of A_1.

For each interval $[j, k]$ with at least one backlogged class, one could in principle dynamically partition the N classes into two subsets: subset $\mathcal{B}$ containing all classes that are continuously backlogged throughout $[j, k]$, and subset $\mathcal{U}$ containing classes that are *not* continuously backlogged throughout $[j, k]$ (although they may be backlogged for a sub-interval in $[j, k]$). For any $i \in \mathcal{B}$, by definition,

$$Y^i_{j,k} = \Theta^i_{j,k}(\mathcal{B}). \tag{13}$$

We also claim that if $A_1 \cap \mathcal{U}$ is empty set, then $Y^i_{j,k} = \Theta^i_{j,k}(A_1)$; otherwise $Y^i_{j,k} \ge \Theta^i_{j,k}(A_1)$.

Since the exact distribution of class i's backlogged service can be very difficult to compute due to the dynamics of the sets $\mathcal{B}$ and $\mathcal{U}$, we next lower bound the available service $\widetilde{Y}^i_{j,k}$ for any partition.

Lemma 3: The available service for class i in interval $[j, k]$, $\widetilde{Y}^i_{j,k}$, always satisfies

$$\widetilde{Y}^i_{j,k} \ge \Theta^i_{j,k}(A_1) \tag{14}$$

for an arbitrary partition A_1.

Proof. If class i is backlogged throughout $[j, k]$, then from Equation (12), $\widetilde{Y}^i_{j,k} = Y^i_{j,k} \ge \Theta^i_{j,k}(A_1)$. If class i is not continuously backlogged throughout $[j, k]$ with input traffic $X^i_{j,k}$, consider sufficient class i traffic $\widetilde{X}^i_{j,k}$, such that class i is backlogged throughout $[j, k]$, while all Q^m_{j-1} and $X^m_{j,k}$, for $m \ne i$, remain the same. The outputs of the N classes will be rearranged according to the GPS service discipline and the new inputs, and will change from $Y^m_{j,k}$ to $Y'^{,m}_{j,k}$, for all $m = 1, \cdots, N$. For the new outputs, we can again construct two subsets, a backlogged subset $\mathcal{B}'$ and an unbacklogged subset $\mathcal{U}'$, such that

$$\begin{aligned} \widetilde{Y}^i_{j,k} &= Y'^{,i}_{j,k} = \Theta'^{,i}_{j,k}(\mathcal{B}') \\ &= \frac{\phi^i}{\sum_{m \in \mathcal{B}'} \phi^m}[C \cdot (k - j + 1) - \sum_{n \in \mathcal{U}'} Y'^{,n}_{j,k}]. \end{aligned}$$

For any other partition of A_1, $\widetilde{Y}^i_{j,k} \ge \Theta'^{,i}_{j,k}(A_1)$. Since $Y'^{,m}_{j,k} \le Y^m_{j,k}$, for $m \ne i$, $\Theta'^{,i}_{j,k}(A_1) \ge \Theta^i_{j,k}(A_1)$, thus, we have shown that $\widetilde{Y}^i_{j,k} \ge \Theta^i_{j,k}$. □

Equation (14) enables us to statistically lower bound $\widetilde{Y}^i_{j,k}$ using the distribution of $\Theta^i_{j,k}(A_1)$ with an arbitrary partition of $A_1 \ni i$,

$$S^i(t) = \frac{\phi^i}{\sum_{m \in A_1} \phi^m}[Ct - \sum_{n \in A_2} B^n_{out}(t)], \tag{15}$$

where $B^n_{out}(t)$ is the statistical traffic envelope for the output traffic $Y^n_{j,k}$. By deliberately setting A_1 and A_2, we can obtain a tight statistical lower bound for $\widetilde{Y}^i_{j,k}$. The technique of choosing A_1 is explored in detail below.

C. Multi-Class Admission Control in GPS

Equation (15) establishes a way to calculate statistical service envelopes with an arbitrary partition of classes, yet a tight lower bound is required to fully exploit inter-class resource sharing. We devise a technique for this purpose as follows.

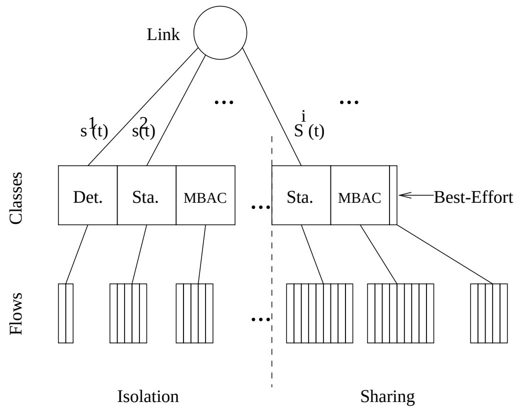

Fig. 4. An Isolation/Sharing Model for Admission Control

First, we illustrate an isolation/sharing model for admission control in Figure 4. In this model, some service classes will use their deterministic service envelope $s^i(t)$ in admission control. These service classes may support deterministic services, in which deterministic traffic envelopes $b^i(t)$ are used. Or they may support less aggressive statistical services which do not wish to exploit spare capacity from other classes. In view of service envelopes, we refer to these service classes as *isolation* classes. Apart from these isolation classes, other service classes will exploit inter-class resource sharing using their statistical service envelope $S^i(t)$ to admit an increased number of flows into the traffic class. We refer to these service classes as *sharing* classes. Sharing classes cannot support deterministic services, but can support statistical, measurement-based, and best-effort services.

Returning to the problem of constructing a tight statistical service envelope, from Equation (13), we know that when class i is backlogged throughout $[j,k]$, if $A_1 \cap \mathcal{B}$ is an empty set, then $\Theta^i_{j,k}(A_1) = \widetilde{Y}^i_{j,k}$. If we move any unbacklogged class into A_1, then $\Theta^i_{j,k}(A_1) < \widetilde{Y}^i_{j,k}$. In this sense, we should ensure that all unbacklogged classes are in A_2. When class i is not backlogged throughout $[j,k]$, from Equation (14), we know that if too many classes are in A_2, $\Theta^i_{j,k}(A_1)$ will be small again. For example, a greedy best-effort class should never be put into A_2.

In order for $\Theta^i_{j,k}$ to closely approximate $\widetilde{Y}^i_{j,k}$, we propose partitioning all sharing classes into A_1, and all isolation classes into A_2. We refer to this partition of A_1 and A_2 as sets $\mathcal{S}$ (sharing) and $\mathcal{I}$ (isolation).

Another issue with Equation (15) is that the statistics of the output traffic, $B^n_{out}(t)$ for $n \in A_2$, are difficult to compute, and the bound in [2], $B^n_{out}(t) \leq_{st} B^n_{in}(t+d^n)$, can be quite loose in practice. Consequently, for $n \in \mathcal{I}$, we approximate $B^n_{out}(t)$ by $B^n_{in}(t)$, because for these isolation classes, admission control is based on the worst case service $s^n(t)$, while the actual service received is typically higher than the worst case scenario. Consequently, these classes are not backlogged most of the time, and the distortions of the outputs to inputs are relatively small and can be neglected, such that we can approximate $B^n_{out}(t)$ with the input traffic envelopes $B^n(t)$. Thus, we propose the following statistical service envelope for any class $i \in \mathcal{S}$, [4]

$$S^i(t) = \frac{\phi^i}{\sum_{m\in\mathcal{S}}\phi^m}[Ct - \sum_{n\in\mathcal{I}} B^n(t)]. \tag{16}$$

We conclude by describing the complete admission control algorithm for a multi-class GPS server. Each class provides traffic parameters $b^i(t)$ and $B^i(t)$, and QoS parameters d^i and P^i. Each class has a weight ϕ^i and guaranteed rate g^i, with guaranteed service envelope $s^i(t) = g^i t$. For deterministic service classes, if $\max_t\{b^i(t) - s^i(t+d^i)\} \leq 0$, then the deterministic QoS for flows inside class i is guaranteed. For *i*solation statistical service classes, if $P[\max_t\{B^i(t) - s^i(t+d^i)\} \leq 0] \leq P^i$, then the statistical QoS of class i is satisfied. For *s*haring statistical service classes, the statistical QoS is satisfied if $P[\max_t\{B^i(t) - s^i(t+d^i)\} \leq 0] \leq P^i$, or if $P[\max_t\{B^i(t) - S^i(t+d^i)\} \leq 0] \leq P^i$ exists for a statistical service envelope $S^i(t)$ obtained via any partition A_1, A_2 using Equation (15). For simplicity, we use Equation (16) instead of testing all partitions of A_1 and A_2.

IV. Computational and Experimental Investigation

In Sections II and III, we studied the delay bound violation probability using statistical traffic and service envelopes for SP and GPS schedulers. In this section, we address the computational aspects of these admission control algorithms and perform trace-driven simulations to quantify the ability of our approach to exploit inter-class resource sharing. The workload consists of a set of 30-minute traces of MPEG compressed video from [17].

A. Computing the Delay Tail Probability

To approximate each flow's traffic descriptor $B_j(t)$, we use the rate variance envelopes in [11], where

$$RV(t) = \text{var}\left\{\frac{X_{s,s+t-1}}{tT}\right\},$$

$m = EX/T$, and T is the length of the time slot, such that $E\{B_j(t)\} = m_j t$ and $\text{var}\{B_j(t)\} = t^2 RV_j(t)$. When flows are multiplexed, the aggregate traffic envelope for the ith class approaches a Gaussian envelope with $B^i(t)$ having mean $\sum_{j\in C_i} tm_j$, and variance $\sum_{j\in C_i} t^2 RV_j(t)$ [11]. In practice, traffic flows can specify policing parameters, and use [18] to compute such statistical traffic envelopes from the deterministic parameters.

To calculate $P[\max_t\{B(t) - S(t+d_0)\} > 0]$ in Equation (6) we utilize the "maximum variance" approximation of [9]. Let

$$\begin{aligned}\sigma_t^2 &= \text{var}\{B(t) - S(t+d_0)\},\\ \alpha_t &= \frac{0 - E\{B(t) - S(t+d_0)\}}{\sigma_t},\\ \alpha &:= \inf_t \alpha_t.\end{aligned}$$

[4]For the special case in which all classes are sharing classes, a tighter bound with alternative partition of A_1, A_2 may exist. We leave study of this issue to future work.

Approximating $\{B(t) - S(t+d_0)\}$ as Gaussian, under conditions (C1)-(C2) in [9],

$$P[\max_t \{B(t) - S(t+d_0)\} > 0] \geq \max_t P[B(t) - S(t+d_0) > 0] = \phi(\alpha) \quad (17)$$

and

$$P[\max_t \{B(t) - S(t+d_0)\} > 0] \leq e^{-\frac{\alpha^2}{2}} \quad (18)$$

where $\phi(\alpha) = \frac{1}{\sqrt{2\pi}} \int_\alpha^\infty e^{-\frac{x^2}{2}} dx$. Proof of these two bounds is given in [9], and we utilize both in the experiments below.

B. Admissible Regions in Multi-Class GPS

The scenario we consider is a link sharing GPS server with a total capacity of 45 Mbps. Different weights ϕ^i are given to different classes, which require either deterministic or statistical services. In the experiments, some classes will exploit inter-class resource sharing, while others will not.

In each experiment, we calculate the admissible region for each class according to the flows' traffic characterizations and QoS requirements using the admission control algorithm described in Section III. We then perform trace-driven simulations using a GPS scheduler with each flow having a randomly shifted initial phase. After the simulation, we measure the utilization and the experimental delay bound violation probability, and compare the simulation results with the required QoS.

In the first experiment, we consider a GPS server with two service classes. Class 1 requires deterministic service, with $d^1 = 10$ msec, class 2 requires statistical service, with $d^2 = 20$ msec and $P^2 = 10^{-4}$. In the admission control tests, we use both the lower bound of Equation (17) and the upper bound of Equation (18) to approximate $P[D > d]$.

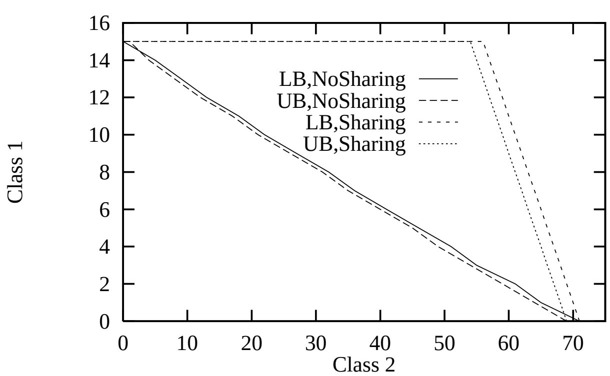

Fig. 5. Admissible Regions for Deterministic and Statistical Services

Figure 5 shows the admissible regions for class 1 and 2 under four different conditions: with and without inter-class sharing for class 2, and upper and lower bounds for $P[D > d]$. Notice the significant increase in the admissible region due to exploiting inter-class resource sharing using our framework of statistical service envelopes. For example, using the lower bound for $P[D > d]$ and setting $g^1 = g^2 = C/2$, without inter-class sharing, the admissible region is $(7, 31)$ flows and the total utilization is 45.3%. In contrast, with inter-class sharing, the admissible region is $(7, 62)$ flows and the total utilization is 82.2%, an increase of 81%. We also observe that the differences in the admissible regions using the lower and upper bounds are merely 1 or 2 flows. We next perform trace-driven simulations and measure the experimental delay bound violation rates using the admissible region calculated from the "sharing" tests. For the lower bound, the mean delay bound violation rate for class 2 is 5×10^{-4}, while for the upper bound, the mean violation rate for class 2 is approximately 5×10^{-5}. Since the QoS parameter is $P^2 = 10^{-4}$, we observe that the actual admissible region must be between the LB and UB sharing curves, and that the admissible regions calculated using both bounds are very close to the true ones.

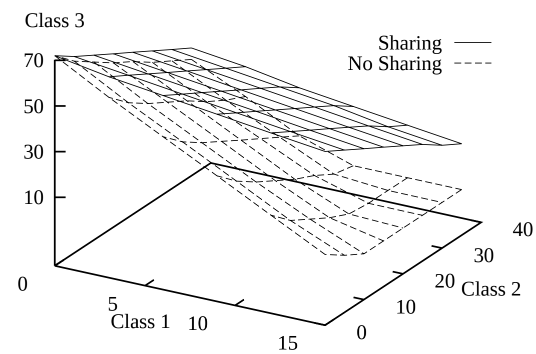

Fig. 6. Admissible Regions for a Three-Class GPS Server

In the next experiment, we consider a three-class GPS scheduler. Class 1 requires deterministic service with $d^1 = 20$ msec, class 2 requires statistical service with $d^2 = 20$ msec and $P^2 = 10^{-4}$, and class 3 requires statistical service with $d^3 = 30$ msec and $P^3 = 10^{-4}$. Class 1 and 2 are isolation classes. We perform admission tests with and without class 3 exploiting inter-class sharing, and use the lower bound of Equation (17) to approximate $P[D > d]$. The admissible region is shown in Figure 6, which also illustrates the significant utilization gain of the approach.

In the above two experiments, the deterministic service class is exploited by the statistical service class to allow inter-class sharing. In the next experiment, we show that our approach is also able to exploit inter-class sharing among statistical service classes. We consider a three class GPS server with each class providing statistical services with the same QoS: $d = 20$ msec and $P = 10^{-4}$. Class 1 and 2 are set to isolation classes. In Figure 7, we show the difference in the admissible regions by allowing class 3 to exploit inter-class sharing.

From Figure 7, observe that ignoring inter-class sharing leads to as many as 8 fewer flows admitted in class 3, for a loss of approximately 10% of the resource utilization. In this scenario,

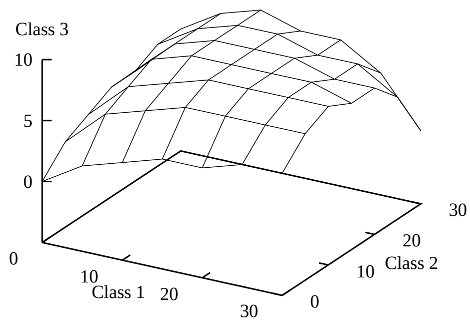

Fig. 7. Increase in Admissible Regions for 3 Statistical Classes

the intra-class statistics are fully exploited, and the gain comes solely from the inter-class statistics. In a high-speed GPS server, even if each class provides statistical service, when the number of service classes is large, the inter-class resource sharing gain can be signficant.

V. Conclusions

In this paper, we developed multi-class admission control algorithms that exploit inter-class statistical resource sharing. We developed a framework of statistical service envelopes to study the problem and showed how such envelopes characterize the excess capacity available to a traffic class due to varying resource demands of other classes. We applied the approach to Static Priority and Generalized Processor Sharing schedulers and experimentally demonstrated that our admission control algorithms are able to extract a significant utilization gain from inter-class resource sharing.

VI. Acknowledgments

The authors are grateful to Tao Wu and Zhi-Li Zhang for their insightful comments.

References

[1] D. Wrege, E. Knightly, H. Zhang, and J. Liebeherr, "Deterministic delay bounds for VBR video in packet-switching networks: Fundamental limits and practical tradeoffs," *IEEE/ACM Transactions on Networking*, vol. 4, no. 3, pp. 352–362, June 1996.

[2] J. Kurose, "On computing per-session performance bounds in high-speed multi-hop computer networks," in *Proceedings of ACM SIGMETRICS '92*, Newport, RI, June 1992, pp. 128–139.

[3] S. Jamin, P. Danzig, S. Shenker, and L. Zhang, "A measurement-based admission control algorithm for integrated services packet networks," *IEEE/ACM Transactions on Networking*, vol. 5, no. 1, pp. 56–70, Feb. 1997.

[4] S. Floyd and V. Jacobson, "Link-sharing and resource management models for packet network," *IEEE/ACM Transactions on Networking*, vol. 3, no. 4, pp. 365–386, Aug. 1995.

[5] J. Bennett and H. Zhang, "Hierarchical packet fair queueing algorithms," *IEEE/ACM Transactions on Networking*, vol. 5, no. 5, pp. 675–689, Oct. 1997.

[6] I. Stoica, H. Zhang, and T. Ng, "A hierarchical fair service curve algorithm for link-sharing," in *Proceedings of ACM SIGCOMM'97*, Cannes, France, Sept. 1997.

[7] R. Cruz, "Quality of service guarantees in virtual circuit switched networks," *IEEE Journal on Selected Areas in Communications*, vol. 13, no. 6, pp. 1048–1056, Aug. 1995.

[8] A. Parekh and R. Gallager, "A generalized processor sharing approach to flow control in integrated services networks: the single-node case," *IEEE/ACM Transactions on Networking*, vol. 1, no. 3, pp. 344–357, June 1993.

[9] J. Choe and N. Shroff, "A central limit theorem based approach to analyze queue behavior in ATM networks," *IEEE/ACM Transactions on Networking*, vol. 6, no. 5, pp. 659–671, Oct. 1998.

[10] G. de Veciana and G. Kesidis, "Bandwidth allocation for multiple qualities of service using generalized processor sharing," *IEEE Transactions on Information Theory*, vol. 42, no. 1, pp. 268–272, Jan. 1996.

[11] E. Knightly, "Second moment resource allocation in multi-service networks," in *Proceedings of ACM SIGMETRICS '97*, Seattle, WA, June 1997, pp. 181–191.

[12] S.-K. Kweon and K. Shin, "Video-on-demand service on packet-switched networks using a statistical traffic envelope," Preprint, 1998.

[13] Z. Zhang, D. Towsley, and J. Kurose, "Statistical analysis of generalized processor sharing scheduling discipline," *IEEE Journal on Selected Areas in Communications*, vol. 13, no. 6, pp. 368–379, Aug. 1995.

[14] R. Cruz, "Quality of service management in integrated services networks," Technical Report, Proc. 1st Semi-Annual Research Review, CWC, University of California at San Diego, June 1996.

[15] R. Cruz, "A calculus for network delay, part I : Network elements in isolation," *IEEE Transactions on Information Theory*, vol. 37, no. 1, pp. 114–121, Jan. 1991.

[16] J. Liebeherr, D. Wrege, and D. Ferrari, "Exact admission control for networks with bounded delay services," *IEEE/ACM Transactions on Networking*, vol. 4, no. 6, pp. 885–901, Dec. 1996.

[17] O. Rose, "Statistical properties of MPEG video traffic and their impact on traffic modeling in ATM systems," in *Proceedings of IEEE Conference on Local Computer Networks*, Minneapolis, MN, Oct. 1995, pp. 397–406.

[18] E. Knightly, "Enforceable quality of service guarantees for bursty traffic streams," in *Proceedings of IEEE INFOCOM '98*, San Francisco, CA, Mar. 1998.