Traffic Shaping for End-to-End Delay Guarantees with EDF Scheduling

Vijay Sivaraman , Fabio M. Chiussi and Mario Gerla

Abstract— **The provision of Quality of Service (QoS) in terms of end-to-end delay guarantees to real-time applications is an important issue in emerging broadband packet networks. Of the various packet scheduling schemes that have been proposed in the literature, Earliest Deadline First (EDF) scheduling in conjunction with per-hop traffic shaping (jointly referred to as Rate-Controlled EDF or RC-EDF) has been recognized as an effective means of end-to-end deterministic delay provisioning. An important aspect that has not been addressed satisfactorily in the literature, however, concerns the choice of RC-EDF shaping parameters that realize maximal network utilizations.**

In this paper, we first establish that except in trivial cases, it is *infeasible* to identify "optimal" shapers that realize maximal RC-EDF schedulable regions. Ascertaining the optimal flow shaper requires the state of the entire network to be considered, making it computationally impractical. We then propose a *heuristic* choice of shaper derived from the number of hops traversed by the flow. The resulting shaper is easy to compute, and varies gracefully between the known optimal shapers for limiting values of the hop-length. We show via simulations that for a realistic traffic mix, our choice of shaper allows RC-EDF to outperform the GPS (Generalized Processor Sharing) scheduling discipline as well as RC-EDF disciplines that use shapers chosen independent of the flow hop-length.

I. INTRODUCTION

The provision of Quality of Service (QoS) to real-time communication streams is a key requirement in emerging broadband packet-switched networks. Applications such as voice and video typically demand QoS guarantees in terms of end-to-end transfer delays. Supporting the heterogeneous delay constraints of these applications with widely varying characteristics requires packet scheduling schemes more sophisticated than First-In-First-Out (FIFO) at each switch in the network (for a survey of such scheduling schemes see [13], [28]). Of these, Generalized Processor Sharing (GPS) [17], [18] (also known as Weighted Fair Queueing (WFQ) [6]) and Earliest Deadline First (EDF) [8], [26] are among the most popular.

GPS guarantees a maximum queueing delay by reserving a certain amount of the link bandwidth at each hop for the given flow. Its main attraction is its simplicity, both in the associated Call Admission Control (CAC) framework [18] as well as in the implementation of the scheduler (recent techniques [22], [3], [1], [2] have made the cost of GPS-related schedulers very affordable). The simplicity, however, comes at a price - GPS is suboptimal in its performance, and yields reduced network admissible regions.

EDF associates a per-hop deadline with each packet and schedules packets in order of deadlines. In the case of a single node, EDF is known to be the optimal scheduling policy [10], [15] in terms of the *schedulable region* for a set of flows with given traffic envelopes and deterministic delay requirements (details in section II-A). In the multi-node setting, however, traffic interactions could severely distort the traffic, and the absence of knowledge of the traffic envelopes at nodes internal to the network makes the use of EDF for end-to-end guarantees problematic. To overcome this problem, the authors in [29] propose the reshaping of traffic at each node in the network. The use of per-node traffic shaping in conjunction with EDF scheduling (we refer to this combination as *Rate-Controlled EDF* or *RC-EDF*) has been studied in detail in [11]. The authors derive expressions for the end-to-end delay in terms of the traffic shaper parameters, and show that the schedulable region under RC-EDF depends critically upon the choice of shaping parameters.

A crucial issue regarding RC-EDF that has not been addressed satisfactorily in the literature concerns the selection of appropriate shaper parameters that realize the largest schedulable regions under RC-EDF. Ad-hoc proposals [11] and results for restricted settings [19], [27] have been presented, but the underlying fundamental issue of identifying the "optimal" shaper, if one such exists, has not been tackled. In this paper, we first establish that except in trivial cases, identifying an "optimal" shaper for a flow under the RC-EDF scheduling discipline requires the state of the entire network to be considered, making it computationally *infeasible* in practice. We then show how our result is consistent with some seemingly contradictory

Vijay Sivaraman is with the Dept. of Computer Science, UCLA, Los Angeles, CA 90095, USA. (email:vijay@cs.ucla.edu)

Fabio M. Chiussi is with Data Networking Research at Bell Labs, Lucent Technologies, Holmdel, NJ 07733, USA. (email:fabio@bell-labs.com)

Mario Gerla is with the Dept. of Computer Science, UCLA, Los Angeles, CA 90095, USA. (email:gerla@cs.ucla.edu)

results in the literature (for example theorem 4.3 in [19], which, for the restricted setting of *single*-leaky-buckets, allows the optimal shaper to be identified under certain conditions). Finally, we propose a heuristic shaper choice that has very desirable properties – it is very simple to compute, and by being dependent on the flow hop-length, varies gracefully between the known optimal shapers for the limiting values of the hop-length. Using simulations of a realistic traffic mix consisting of numerous video flows, we show that our "hop-length dependent" heuristic shaper choice allows RC-EDF to realize significantly larger admissible regions than both GPS as well as RC-EDF disciplines that employ shapers derived independent of flow hop-lengths.

The rest of this paper is organized as follows: section II provides the requisite background on EDF scheduling and traffic shaping. Section III establishes the infeasibility of optimal shaping, and presents a heuristic choice having desirable properties. Performance results from simulation are presented in IV, and concluding remarks are offered in section V.

II. Background

A. EDF

We review some basic concepts related to EDF scheduling, and briefly describe the framework for deterministic end-to-end delay guarantees developed in [11]. The EDF scheduling discipline [8], [26] works as follows: each flow i at switch m is associated with a *local* delay bound d_i^m; then, an incoming packet of flow i arriving to the scheduler at time t is stamped with a deadline $t + d_i^m$, and packets in the scheduler are served by increasing order of their deadline.

In the deterministic setting, EDF is known to be the optimal scheduling policy at a single switch [10]. Optimality is defined in terms of the *schedulable region* associated with the scheduling policy. Given N flows with traffic envelopes $A_i(t)$[1] ($i = 1, 2, \ldots, N$) sharing an output link, and given a vector of delay bounds $\vec{d} = (d_1, d_2, \ldots d_N)$, where d_i is an upper bound on the local scheduling delay that packets of flow i can tolerate, the schedulable region of a scheduling discipline π is defined as the set of all vectors $\vec{d}$ that are schedulable under π. The authors in [10], [15] show that EDF has the largest schedulable region of all scheduling disciplines, given by the vectors that satisfy the following constraint:

$$\sum_{i=1}^{N} A_i(t - d_i) \leq Ct, \quad \forall t > 0 \tag{1}$$

where C denotes the link rate, $A_i(t) = 0$ for $t < 0$, and it is assumed that either the packet transmission time is negligible (as in ATM networks) or the scheduler is preemptive. In the case of non-preemptive scheduling with non-negligible packet sizes, the above constraint guarantees delay bound $d_i + L/C$ to every flow i, where L denotes the maximum packet size at the switch. Given the traffic envelopes and the delay requirements of each flow, inequality (1) can be used directly to devise a single-node CAC mechanism. In the multi-node setting, however, the traffic envelopes are no longer known at the inputs of the nodes inside the network, and the interactions that distort the traffic are not easily characterizable. To overcome this problem, Zhang and Ferrari in [29] propose a class of schemes called *Rate-Controlled Service* (RCS) disciplines which reshape the traffic at each hop within the network (EDF with per-hop reshaping is referred to as *Rate Controlled EDF* or RC-EDF). Georgiadis *et al.* in [11] build upon this model and derive expressions for the end-to-end delay bounds in terms of the shaper envelope $E_i(t)$ and the scheduling delay at each node; they show that no advantages are gained by having non-identical shapers for a flow at each switch it traverses. The end-to-end delay bound d_i for flow i is given by

$$d_i = d_i^{sh} + \sum_{m=1}^{M} d_i^m \tag{2}$$

where $d_i^{sh} = D(A_i \| E_i)$ denotes the maximum shaper delay and d_i^m is the local scheduler delay bound at the m-th switch for flow i. The maximum shaper delay is incurred only *once*, and is independent of the number of nodes on the path. Equation (2), in conjunction with the single-node CAC derived from inequality (1) above readily leads to an end-to-end CAC framework [11], [4]: once an appropriate shaper E_i has been chosen, the delay incurred in the shaper is computed; the remaining delay is split among the schedulers on the path of the flow, and the flow is admitted only if the single-node CAC at each switch along the path admits the flow. The schedulable region achieved with RC-EDF depends critically on the choice of shaper E_i; the authors in [11] show that the use of shaper parameters induced by GPS allows RC-EDF to outperform GPS. The design of shapers which achieve even larger (if not the largest) schedulable regions under RC-EDF, however, has not been addressed in the literature.

[1] Flow i has envelope $A_i(t)$ if the amount of flow i traffic entering the network in any interval of length t is bounded by $A_i(t)$ [5]. A typical example is the *multiple-leaky-bucket* descriptor $(\sigma_k, \rho_k)_{k=1,\ldots,K_i}$ denoting the envelope $A_i(t) = \min_{1 \leq k \leq K_i} \{\sigma_k + \rho_k t\}$.

B. Traffic Shaping / Smoothing

Smoothing traffic at the ingress to the network to make it less bursty has in many contexts been recognized as a means of increasing the schedulable region of the network. For example, numerous authors have proposed off-line work-ahead [16], [7], [30], [24] as well as on-line [20], [21], [12] smoothing techniques for the transmission of stored and interactive real-time video traffic. The focus in many of these frameworks has been the reduction of the traffic stream's peak rate, rate variance, or some related cost metric, typically under buffering constraints. Moreover, the delay bounds in many of these frameworks are not strict and the resulting QoS is often statistical. By contrast, the focus in this work is on frameworks which provide *deterministic* delay guarantees to flows with prespecified traffic envelopes in the context of RC-EDF scheduling.

References [11] and [19] establish the framework for shaping in the context of deterministic delay guarantees under RC-EDF, and help identify the family of "good" shapers that are most effective in providing these delay guarantees. We recall three lemmas from these references which shall be useful for our work in this paper. Let flow i be characterized by the multiple-leaky-bucket arrival envelope $A_i(t) = \min_{1 \le k \le K_i}\{\sigma_{i,k} + \rho_{i,k} t\}$ and end-to-end deterministic delay requirement d_i. Further, let $E_i^m(t)$ denote the shaper envelope for flow i at the m-th node on its path, and the symbol $\wedge$ represent concatenation (i.e., series placement) of shapers. (Throughout this work all traffic and shaper envelopes are assumed to be concave, increasing, piecewise-linear functions with a finite number of slopes; thus they can be described by the multiple-leaky-bucket form $(\sigma_k, \rho_k)_{k=1,\ldots,K}$ where $\sigma_1 < \sigma_2 < \cdots < \sigma_K$ and $\rho_1 > \rho_2 > \cdots > \rho_K$.) The first result shows that it suffices to consider RC-EDF disciplines that, for a flow i, employ an identical shaper E_i at each of the nodes traversed by the flow.

Lemma 1: [11, Proposition 2] Consider flow i that traverses nodes $1, \ldots, M$. Given the RC-EDF discipline that uses shaper E_i^m for the flow at node m ($m = 1, \ldots, M$), the RC-EDF discipline that uses shaper $E_i' = \bigwedge_{m=1}^{M} E_i^m$ for the flow at each node m ($m = 1, \ldots, M$) can provide the same end-to-end delay guarantees.

The second lemma shows that a "good" shaper for a given flow i is characterized by a single parameter, namely the worst-case shaping delay. Such a shaper is moreover easy to construct given the shaping delay.

Lemma 2: [11, Proposition 3] For flow i with multiple-leaky-bucket arrival envelope $A_i(t) = \min_{1 \le k \le K_i}\{\sigma_{i,k} + \rho_{i,k} t\}$, the envelope of the *smallest* shaper $E_i(d)(t)$ which guarantees that $D(A_i \| E_i(d)) \le d$, where $0 \le d \le \sigma_{i,K_i}/\rho_{i,K_i}$, is unique and given by

$$E_i(d)(t) = \begin{cases} \frac{A_i(\tau_{i,k^*})t}{(\tau_{i,k^*}+d)}, & \text{if } 0 \le t < \tau_{i,k^*} + d \\ A_i(t-d), & \text{if } t \ge \tau_{i,k^*} + d \end{cases}$$

where $\tau_{i,1} = 0$, $\tau_{i,k} = (\sigma_{i,k} - \sigma_{i,k-1})/(\rho_{i,k-1} - \rho_{i,k})$ for $2 \le k \le K_i$, and $k^* = \min_{1 \le k \le K_i}\{k \ : \ A_i(\tau_{i,k}) - \rho_{i,k}(\tau_{i,k} + d) \ge 0\}$.

The third lemma shows that the family of "good" shapers can be further restricted to ones with peak rate no larger than the link rate at any of the switches on the path of the flow. (The peak rate p of the multiple-leaky-bucket envelope $E(t) = \min_{1 \le k \le K}\{\sigma_k + \rho_k t\}$, where $\sigma_1 < \cdots < \sigma_K$ and $\rho_1 > \cdots > \rho_K$, is given by ρ_1 if $\sigma_1 = 0$ and ∞ otherwise.)

Lemma 3: [11, Proposition 4] Consider flow i traffic with arrival envelope $A_i(t)$ that traverses nodes $1, \ldots, M$ with corresponding output link speeds C^m. Then given an RC-EDF discipline that uses shaper envelope $E_i(d)(t)$, there is an RC-EDF discipline using shaper envelope $E_i(d')(t)$, $d' \ge d$, which guarantees the same end-to-end delays to all flows and whose peak rate $p' \le \min_{1 \le m \le M}\{C^m\}$.

The above three lemmas identify the family of "good" shapers to which we can restrict our attention in order to design efficient RC-EDF disciplines. However, the family of "good" shapers for a flow i comprises of the set $\{E_i(d)(t) \ : \ 0 \le d \le \sigma_{i,K_i}/\rho_{i,K_i}\}$, and could be very large. The general problem of identifying an "optimal" shaper among them has not been addressed. Results for some restricted settings, however, have been presented. For example, a result in [27] establishes that in the restricted case of *homogeneous* traffic flows, smoothing is beneficial if and only if the hop lengths are larger than a critical value. Another result [19, Theorem 4.3] which we recall in the following theorem, shows that in the special case where shaper envelopes are restricted to the *single*-leaky-bucket form $E(t) = \sigma + \rho t$, the largest schedulable regions under RC-EDF can be realized by smoothing entirely the traffic from flows with a "sufficiently" large hop-length. (A confirmation of this result will emerge in the course of our discussion in section III-A.)

Theorem 1: [19, Theorem 4.3] Consider an RC-EDF discipline using traffic shapers restricted to the single-leaky-bucket form, and carrying a flow i with arrival envelope $A_i(t) = \sigma_i + \rho_i t$ traversing M hops with link rates $C^1, \ldots, C^M$. Then, if $\sum_{m=1}^{M} \rho_i / C^m \ge 1$, the schedulable region of the RC-EDF discipline is not reduced if the shaper with envelope $E_i(t) = \sigma_i' + \rho_i t$, where $0 \le \sigma_i' \le \sigma$, is used for flow i at every switch.

In the general setting, however, where (a) *multiple* (rather than just *single*) leaky-bucket shapers are permitted, and (b) flows can have arbitrary hop-lengths, it is not known if "optimal" shapers can be identified. A few adhoc shaping strategies have been suggested in the literature, for example [11], which proposes the use of shaping parameters induced by the rate-based GPS scheduling discipline. Such a choice is shown to give reasonably good performance, i.e., at least as good as GPS, and still allow the RC-EDF discipline to accept some additional calls. Nevertheless, this choice corresponds to a significant smoothing of the traffic at the ingress. This may be reasonable for flows with large hop-lengths (as suggested by theorem 1 above), since it avoids the fragmentation of the end-to-end delay budget among the hops and instead employs it towards traffic smoothing. However, for short hop-length flows, it yields poor performance. For instance, when the hop length is 1, smoothing is known to be detrimental to network performance [14]. The GPS-induced shaping parameters, therefore, do not yield good performance under all conditions.

III. CHOOSING THE RC-EDF SHAPER

In this section, we address the issue of identifying "optimal" shapers for use with RC-EDF disciplines that guarantee end-to-end deterministic delay bounds to flows. We define the "optimal" shaper as follows:

Definition 1: An *optimal* shaper E_i for flow i is such that the RC-EDF discipline that uses shaper envelope $E_i(t)$ for flow i guarantees end-to-end delays to all flows no smaller than the RC-EDF discipline that uses any envelope $E_i'(t)$ for flow i.

The shaper envelope for the flow is typically chosen at call-setup, and not modified during the lifetime of the flow. Moreover, the choice is made independent of the other flows in the network, since 1) the number of flows in the network is typically too large, and 2) the set of flows in the network varies dynamically as flows enter and leave the network. Therefore, it is reasonable to restrict our focus to *network-state-independent shapers*, i.e., shapers that are constructed independent of the other traffic in the network. The question of interest that needs to be addressed is whether there exist shapers that are *both* optimal and network-state-independent.

In the first part of this section, we show that except in the trivial case where a flow is either constant bit rate or has hop-length one, shapers that are both optimal and network-state-independent cannot exist. We then show how this result reconciles with theorem 1 above, which suggests that in the restricted setting of single-leaky-bucket shapers, it is optimal to smooth flows with sufficiently large hop-lengths. We show that the result of theorem 1 is due to the "bad" shaper description by virtue of the single-leaky-bucket restriction, rather than an inherent advantage of shaping. In fact even a naive network-state-independent multiple-leaky-bucket shaper is shown to be capable of outperforming the best network-state-independent single-leaky-bucket shaper. Having established that the design of an optimal shaper is infeasible, we propose a heuristic choice that has desirable properties and yields very good performance for reasonably realistic traffic mixes.

We assume here that either the EDF scheduling discipline is preemptive or packet transmission times are negligible. For non-preemptive EDF with non-negligible packet sizes, the end-to-end delay bound for a flow i traversing M nodes can be adjusted by the quantity $\sum_{m=1}^{M} L^m/C^m$, where C^m denotes the link rate and L^m the maximum packet size at node m, in order to account for the effects of packetization.

A. Infeasibility of Optimal Shaping

Consider a link of rate C employing EDF scheduling, and let the workload $\mathcal{W}$ consist of N flows, where flow i is characterized by the concave piecewise linear envelope $A_i(t)$ and has a maximum delay requirement d_i. Then we define the following:

Definition 2: The *service demand* $D_{\mathcal{W}}(t)$, the *residual capacity* $F_{\mathcal{W}}(t)$ and the *effective residual capacity* $R_{\mathcal{W}}(t)$ corresponding to the workload $\mathcal{W}$ are defined by

$$D_{\mathcal{W}}(t) = \sum_{i=1}^{N} A_i(t-d_i), \quad t \geq 0 \tag{3}$$

$$F_{\mathcal{W}}(t) = Ct - D_{\mathcal{W}}(t), \quad t \geq 0 \tag{4}$$

$$R_{\mathcal{W}}(t) = \min_{t' \geq t} F_{\mathcal{W}}(t'), \quad t \geq 0 \tag{5}$$

Further define $\{R_{\mathcal{A}} \succeq R_{\mathcal{B}}\} \equiv \{\forall t \geq 0: \; R_{\mathcal{A}}(t) \geq R_{\mathcal{B}}(t)\}$, and $\{R_{\mathcal{A}} \succ R_{\mathcal{B}}\} \equiv \{R_{\mathcal{A}} \succeq R_{\mathcal{B}} \text{ and } \exists t \geq 0: \; R_{\mathcal{A}}(t) > R_{\mathcal{B}}(t)\}$.

Using the above definitions, we establish the following two lemmas (proved in appendix A and B respectively):

Lemma 4: Let $\mathcal{W}$ denote the workload at the EDF scheduler. Now consider a disjoint workload $\mathcal{U}$ with service demand $D_{\mathcal{U}}(t)$. Then the admissibility condition for workload $\mathcal{U}$ is given by $\forall t \geq 0: \; D_{\mathcal{U}}(t) \leq R_{\mathcal{W}}(t)$.

Lemma 5: Let workloads $\mathcal{A}$ and $\mathcal{B}$ yield effective residual capacities $R_{\mathcal{A}}(t)$ and $R_{\mathcal{B}}(t)$ respectively, and let $\lim_{t\to\infty} R_{\mathcal{A}}(t)/t > 0$ and $\lim_{t\to\infty} R_{\mathcal{B}}(t)/t > 0$. Then $R_{\mathcal{A}} \succeq R_{\mathcal{B}}$ if and only if, for every workload $\mathcal{U}$, $\mathcal{B} \cup \mathcal{U}$ is feasible implies that $\mathcal{A} \cup \mathcal{U}$ is also feasible.

We are now ready to prove our result that given a flow that is neither constant bit-rate nor has hop-length of one, a shaper that is optimal for the flow and network-state-independent at the same time cannot exist. Recall that it suffices to focus on shapers which, for any flow, are identical at each node (lemma 1) traversed by the flow, have the smallest envelope for a given shaping delay (as per the construction in lemma 2), and have peak rate no larger than the link rate (lemma 3) at each of the nodes traversed by the flow.

Theorem 2: Consider flow i with multiple-leaky-bucket input traffic arrival envelope $A_i(t) = \min_{1 \le k \le K}\{\sigma_{i,k} + \rho_{i,k} t\}$ ($K \ge 2$) that traverses nodes $1, \ldots, M$ ($2 \le M < \infty$) with corresponding output link speeds C^m. Further, let the peak rate of flow i be no more than the minimum link speed along the path of the flow (i.e., $\sigma_{i,1} = 0$ and $\rho_{i,1} \le \min_{1 \le m \le M}\{C^m\}$). Then, there does not exist a network-state-independent shaper $E_i(d_i^{sh})$ that is optimal, in the sense of guaranteeing that the RC-EDF discipline employing shaper envelope $E_i(d_i^{sh})(t)$ provides end-to-end delays to all flows no worse than the RC-EDF discipline that uses shaper envelope $E_i(d_i^{'sh})(t)$ for arbitrary $0 \le d_i^{'sh} \le \delta_{i,K}/\rho_{i,K}$.

Proof: A direct proof could be derived by computing the effective residual bandwidths for two arbitrary but distinct choices of the shaping delay and showing that the $\succeq$ relation cannot hold between them; however, here we present a proof by explicitly constructing traffic examples that contradict the optimality property. Let $d_i > 0$ denote the end-to-end delay requirement of flow i, and assume that there exists a value of d_i^{sh} such that the shaper envelope $E_i(d_i^{sh})(t)$ is network-state-independent and optimal. Thus, irrespective of the cross-traffic at the various switches, the RC-EDF discipline that uses shaper envelope $E_i(d_i^{sh})(t)$ guarantees end-to-end delays to all flows no worse than the RC-EDF discipline that uses any shaper $E_i(d_i^{'sh})(t)$ where $0 \le d_i^{'sh} \le \delta_{i,K}/\rho_{i,K}$. Consider the two cases:

Case I – $d_i^{sh} > 0$: Choose the cross-traffic at each switch m to be a single flow with dual-leaky-bucket envelope $A^m(t) = \min\{C^m t, \frac{d_i \rho_{i,1}}{M} + (C^m - \rho_{i,1})t\}$ and hop length 1. It is easily verified using (1) that the envelope $E_i(d^{'sh})(t)$ where $d^{'sh} = 0$ can guarantee an end-to-end delay bound of d_i to flow i (by guaranteeing delay bound d_i/M at each node) while simultaneously providing a delay bound of 0 to each of the other flows. The shaper envelope $E_i(d^{sh})(t)$, however, cannot guarantee these delay bounds, since $d^{sh} > 0$ implies that at least one node m on the flow's path has to guarantee a delay bound lower than d_i/M to the flow i envelope $E_i(d^{sh})(t)$, but simultaneously providing a delay bound of 0 to the cross-traffic at node m is not feasible.

Case II – $d_i^{sh} = 0$: Select the cross-traffic at each switch m to be a single flow with dual-leaky-bucket envelope $A^m(t) = \min\{C^m t, \frac{d_i - \epsilon}{M} \frac{\rho_{i,1}}{1+\epsilon(\rho_{i,1}-\rho_{i,2})/\sigma_{i,2}} + (C^m - \frac{\rho_{i,1}}{1+\epsilon(\rho_{i,1}-\rho_{i,2})/\sigma_{i,2}})t\}$ (where $0 < \epsilon < \min\{d_i, \frac{\sigma_{i,2}}{\rho_{i,2}}\}$) and hop length of 1. It can be verified that the envelope $E_i(d^{'sh})$ where $d^{'sh} = \epsilon$ can guarantee an end-to-end delay bound of d_i to flow i (by guaranteeing a delay bound of $\frac{d_i - \epsilon}{M}$ at each hop) as also delay bound 0 to the cross-traffic. However, $E_i(d^{sh})(t)$ where $d^{sh} = 0$ cannot simultaneously provide a delay bound of 0 to the cross-traffic while providing an end-to-end delay bound of d_i to flow i, as that would require at least one of the nodes m on the path to provide a delay bound no larger than d_i/M, and this is not feasible for $M > 1$. △

The above theorem establishes that the results for restricted settings considered in [27] and [19] do not extend to the general setting. Reference [27] shows that in the presence of *homogeneous* traffic flows, smoothing is beneficial if and only if the hop-lengths are larger than a critical value. In the presence of *heterogeneous* traffic, however, such an argument does not hold. The *single*-leaky-bucket restriction in theorem 1 allows optimal shapers to be identified for flows with sufficiently large hop-lengths; for general *multiple*-leaky-bucket envelopes, however, optimal shapers cannot be identified. In fact, the apparent advantage of shaping in the single-leaky-bucket case is due to *poor* shaper description (by virtue of the single-leaky-bucket restriction) rather than an inherent advantage of shaping. To demonstrate this, we show that a naive *dual*-leaky-bucket shaper that performs *peak-rate regulation* at the link rate outperforms *any* network-state-independent *single*-leaky-bucket shaper. We first establish the following two lemmas (proofs in appendix C and D respectively):

Lemma 6: Consider an arbitrary workload at an EDF scheduler operating at rate C. A flow f with envelope $A(t) = \sigma + \rho t$ and delay bound d is admissible if and only if flow f' with envelope $A'(t) = \min\{Ct, \sigma(1 - \frac{\rho}{C}) + \rho t\}$ and delay bound $d' = d - \sigma/C$ is admissible.

Lemma 7: For an arbitrary workload at an EDF scheduler operating at rate C, if the flow f with envelope $A(t) = \sigma + \rho t$ can be guaranteed delay bound d, then the flow f' with envelope $A'(t) = \sigma' + \rho t$, where $\sigma' \le \sigma$, can be guaranteed delay bound $d' = d - \frac{\sigma - \sigma'}{C}$.

In general, it is not always possible to provide a delay bound tighter than d' to flow f' in lemma 7 above. For example, consider cross traffic with envelope $E(t) = Cd - \sigma - \frac{\epsilon\sigma}{C} + \epsilon t$ where ϵ is very small. Flow f with delay bound d is admissible, while flow f' is not admissible for any delay bound lower than $d - \frac{\sigma - \sigma'}{C - \epsilon}$, which, by choosing ϵ small enough, can be made as close to d' as desired.

These lemmas help us establish the following theorem that the naive dual-leaky-bucket shaper can out-

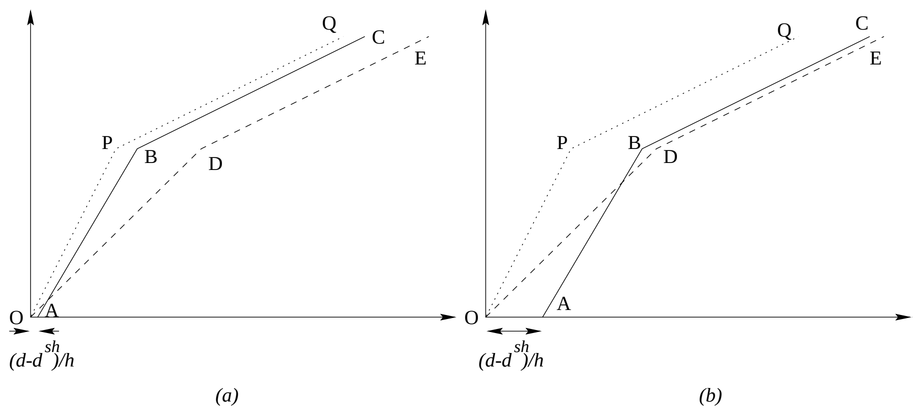

Fig. 1. Service demands for high and low choices of d^{sh} when hop-length h is (a) large, (b) small

perform the best network-state-independent single-leaky-bucket shaper.

Theorem 3: Consider a flow i with single leaky bucket input traffic envelope $A(t) = \sigma_i + \rho_i t$ that traverses nodes $1, \ldots, M$ with corresponding output link speeds C^m. Then the RC-EDF discipline that employs for flow i the network-state-independent shaper $E_i^m(t) = \min\{C^m t, \sigma_i(1 - \frac{\rho_i}{C^m}) + \rho_i t\}$ at each node $m = 1, \ldots, M$ guarantees end-to-end delay bounds no worse than *any* RC-EDF discipline that employs for flow i *only* single-leaky-bucket network-state-independent shapers.

Proof: Let d_i^m denote the delay bound at node m provided to flow i when the arrival envelope $A(t)$ is used, and let $d_i = d_i^1 + d_i^2 + \ldots + d_i^M$ denote the associated tightest end-to-end delay guarantee.

Consider first the RC-EDF discipline that is restricted to single-leaky-bucket shaping. Denote by $E_i(t) = \sigma_i' + \rho_i t$ the shaper envelope (where $\sigma_i' \geq 0$ is picked independent of the cross-traffic in the network), and by δ_i^m the delay bound at node m. Then the shaping delay $\delta_i^{sh} = \frac{\sigma_i - \sigma_i'}{\rho_i}$, and the end-to-end delay bound for flow i is $\delta_i = \delta_i^{sh} + \delta_i^1 + \ldots + \delta_i^M$. From lemma 7, $\delta_i^m = d_i^m - \frac{\sigma_i - \sigma_i'}{C}$ (note that by virtue of the network-state-independent property tighter delay bounds cannot be guaranteed). Thus $\delta_i = \frac{\sigma_i - \sigma_i'}{\rho_i} + \sum_{m=1}^{M}(d_i^m - \frac{\sigma_i - \sigma_i'}{C^m})$, i.e.,

$$\delta_i = d_i + \frac{\sigma_i - \sigma_i'}{\rho_i}\left[1 - \sum_{m=1}^{M}\frac{\rho_i}{C^m}\right] \qquad (6)$$

The above equation incidentally provides a proof for theorem 1 by showing that when $\sum_{m=1}^{M}\frac{\rho_i}{C^m} > 1$, $\delta_i < d_i$ holds and hence smoothing is beneficial irrespective of other traffic in the network.

Now consider the RC-EDF discipline that enforces the peak rate at each node, i.e. uses shaper envelope $E_i^m(t) = \min\{C^m t, \sigma_i(1 - \frac{\rho_i}{C^m}) + \rho_i t\}$ at node m. The total shaping delay is $\theta_i^{sh} = \sigma_i / C^{\min}$ where $C^{\min} = \min_{1 \leq m \leq M}\{C^m\}$. From lemma 6, the delay bound at node m is $\theta_i^m = d_i^m - \sigma_i / C^m$. The end-to-end delay bound is thus $\theta_i = \sigma_i / C^{\min} + \sum_{m=1}^{M}(d_i^m - \sigma_i / C^m)$, i.e.,

$$\theta_i = d_i + \frac{\sigma_i}{C^{\min}} - \sum_{m=1}^{M}\frac{\sigma_i}{C^m} \qquad (7)$$

Using the fact that $0 \leq \sigma_i' \leq \sigma_i$ and $\rho_i \leq C^{\min}$, it can be shown that $\theta_i \leq \delta_i$ for arbitrary choice of σ_i'. $\triangle$

Thus even if input traffic is described by single leaky buckets, RC-EDF disciplines that use the above naive dual-leaky-bucket shapers within the network can realize better end-to-end delays than any RC-EDF discipline that uses only single-leaky-bucket shapers. This shows that the advantages of smoothing in theorem 1 arise due to the poor shaper description rather than an inherent advantage of smoothing.

B. Heuristic Shaper Choice

We have established (in theorem 2) that we cannot realize network-state-independent shapers that are *optimal*. Yet, we can identify *heuristic* choices that can be expected to perform well for reasonably realistic traffic mixes.

Recall from (2) that for a given end-to-end delay budget, the worst-case shaping delay is incurred only *once*, while the remaining delay is *subdivided* among the hops. Therefore, one can expect in general that smoothing is advantageous for flows with large hop-lengths, and detrimental when the hop-lengths are small. To illustrate this with an example, consider a flow f with (p, σ, ρ) dual-leaky-bucket ingress traffic. The envelope $A(t) = \min\{pt, \sigma + \rho t\}$ is depicted by OPQ in figure 1. Let d (where $d \leq \sigma/\rho$) denote the end-to-end delay requirement and h the hop-length of the flow. Further, assume that once a shaping delay d^{sh} has been selected for the flow, the remaining scheduling delay $d - d^{sh}$ is split equally among the hops.

TABLE I
FOUR-SEGMENT CHARACTERIZATION FOR SIX MPEG-CODED MOVIE TRACES

Movie	σ_1	ρ_1	σ_2	ρ_2	σ_3	ρ_3	σ_4	ρ_4
Advertisements	0	1.6	800.0	0.8000	1333.0	0.6000	1600.0	0.5330
Jurassic	0	4.0	133.3	1.0540	400.0	0.8533	1066.0	0.7619
Mtv	0	6.0	266.6	2.3565	933.3	1.9730	1866.6	1.8666
Silence	0	4.0	266.6	0.6665	533.3	0.6000	1133.0	0.5000
Soccer	0	5.0	266.6	2.5000	1000.0	1.2380	2133.3	1.0666
Terminator	0	3.4	133.3	0.7878	266.6	0.5866	800.0	0.3666

Consider first the case when h is reasonably large. Figure 1(a) shows, at a switch, the service demand ABC when d^{sh} is very small (i.e., very little smoothing) as also the service demand ODE when $d^{sh} = d$ (complete smoothing). Though ABC $\not\geq$ ODE, the service demand ODE lies below the service demand ABC for the *most* part. Therefore it seems reasonable to expect that smoothing will yield considerable benefits when h is high. On the other hand, when h is low, the service demand ABC corresponding to small d^{sh} as shown in figure 1(b) is preferable in general over the smoothed case ODE. Thus a small value for d^{sh} can be expected to yield better performance when h is low.

The above observations are valid even in the presence of multiple-leaky-bucket envelopes, and lead us to propose the following heuristic choice of the shaping delay:

$$d^{sh} = \min\left\{d\left(1 - \frac{1}{h}\right), \frac{\sigma_K}{\rho_K}\right\} \tag{8}$$

The shaper envelope corresponding to this choice of shaping delay can be computed using lemma 2. The shaper envelope thus obtained has very desirable properties. It is *optimal* for the limiting values of the hop-length. Indeed, when $h = 1$, d^{sh} computes to zero; this corresponds to no smoothing at all, and is in conformance with the result in [14] showing that smoothing is always detrimental for flows traversing a single hop. When $h \to \infty$, (8) yields $d^{sh} = \min\{d, \sigma_K/\rho_K\}$, in accordance with the observation that the traffic should be smoothed entirely at the ingress to the network. Our proposed shaper choice varies gracefully between these optimal limiting cases of the hop-length. Moreover, it is very easy to compute since it is independent of exogenous traffic in the network, and yields very good performance, as demonstrated in the next section for a reasonably realistic traffic scenario.

IV. PERFORMANCE

To quantify the performance of our "hop-length dependent" choice of shaper, we compare via simulations the call blocking probabilities yielded by RC-EDF disciplines that use shaper choices corresponding to 1) no smoothing, i.e., $d^{sh} = 0$, 2) complete smoothing, i.e., $d^{sh} = d$ (as recommended in [11]), and 3) partial smoothing, as per our proposal in (8). For comparison, we also simulate the behavior of the GPS scheduling discipline.

For our simulations, we focus on one switch within the network, and assume that the chosen switch is the bottleneck for all the flows passing through it; the chosen switch therefore determines if an incoming flow can be accepted into the network or not. Further, the chosen switch operates at 155 Mbps (corresponding to an OC-3 ATM link), and multiplexes a traffic mix consisting of six types of video flows. The various flow types have traffic characteristics as shown in table I. Each row represents a four-segment multiple-leaky-bucket characterization $(\sigma_k, \rho_k)_{k=1,\cdots,4}$ of a movie trace, where the σ's are in Kbits and the ρ's in Mbits/s. These characterizations are borrowed from [9], and have been derived as four-segment covers of the empirical envelopes of traces of MPEG-1 coded movies in [23].

Flow arrivals are generated according to a Poisson process with parameter α and their durations are exponentially distributed with mean $1/\beta$. The ratio α/β characterizes the load offered to the link, i.e., the average number of flows that would exist at any time at a link with no capacity limitation. Each flow has traffic characteristics chosen randomly from the characteristics of the six types shown in table I. The end-to-end delay requirement d (excluding propagation delays) of the flow is uniformly distributed in [100ms, 1.5s], and its hop-length uniformly chosen in [1,5]. After a flow is generated with the above parameters, shaper envelopes for the flow are selected as per the three shaping strategies: 1) no smoothing ($d^{sh} = 0$), 2) complete smoothing ($d^{sh} = \min\{d, \sigma_4/\rho_4\}$), and 3) hop-length dependent smoothing as given in (8). The remaining delay $d - d^{sh}$ under each of the shaping strategies is then split equally among the hops, and the EDF call acceptance test is performed at the switch to determine if

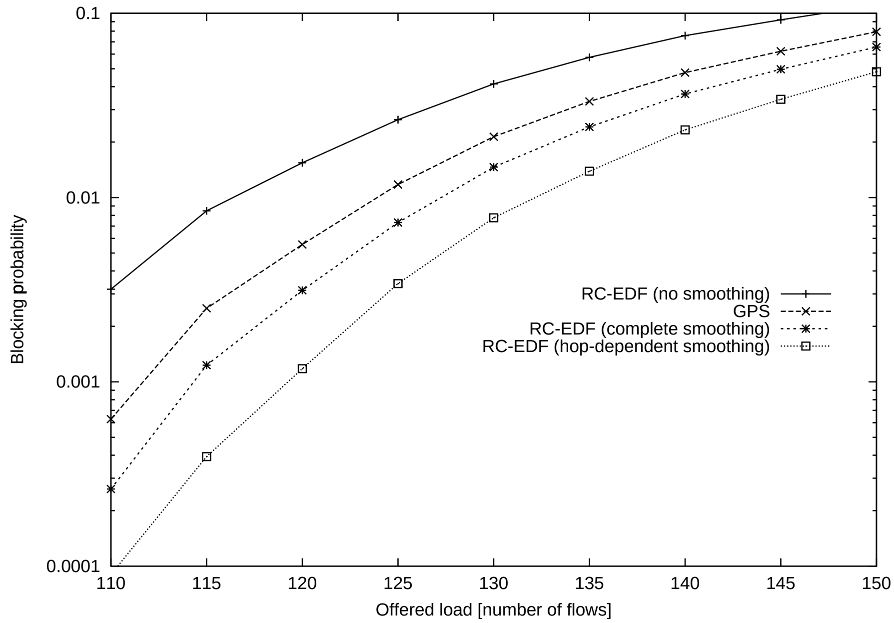

Fig. 2. Call blocking probabilities under RC-EDF (no smoothing), GPS, RC-EDF (complete smoothing) and RC-EDF (hop-length dependent smoothing). Hop-length chosen uniformly in [1,5].

the flow can be accepted into the network (as stated before, the switch considered is assumed to be the bottleneck, and hence determines if the flow can be accepted into the network or not). We employ the exact schedulability test of (1); alternatively, the approximate but simpler call admission method proposed in [9] could be used ([9] shows that for the very same traffic mix as considered here, the degradations introduced by the approximations in the call admission method are very small). We generate a million flows in each simulation run, and are interested in the link blocking probability, i.e., the ratio between the number of rejected flows and the total number of generated flows. We take the call blocking probability under each shaping strategy as a measure of its performance. For comparison, the call blocking probability under the GPS scheduling scheme is also measured.

Figure 2 plots the call blocking probabilities under 1) GPS, 2) RC-EDF with no smoothing, 3) RC-EDF with complete smoothing, and 4) RC-EDF with our hop-length dependent smoothing method of (8), as the offered load is varied from 110 to 150 calls. The confidence intervals are quite small and not shown in the figure. We first observe that the RC-EDF discipline employing complete smoothing outperforms GPS. This is in accordance with the results of [11] showing that RC-EDF disciplines employing the GPS-induced shaping parameters outperform GPS. More importantly, we observe from figure 2 that *both* GPS as well as RC-EDF that uses complete smoothing are outperformed significantly by the RC-EDF discipline employing the "hop-length dependent" shaper derived from (8). For example, if the system is designed to operate at a call rejection rate of 1%, our shaper choice allows RC-EDF to achieve a utilization around 4% larger than the complete smoothing shaper choice and around 6.5% larger than GPS. Thus our proposed shaper choice allows RC-EDF to realize better performance for realistic traffic scenarios than both GPS and RC-EDF disciplines that employ shapers chosen independent of flow hop-lengths.

V. Conclusions

RC-EDF has been proposed as a more efficient way of end-to-end delay provisioning in networks supporting deterministic QoS than GPS [11], [4]. The performance of RC-EDF, however, depends crucially upon the choice of shaping parameters. In this paper, we have addressed the question of identifying traffic shapers that realize maximal RC-EDF schedulable regions. We have shown that identifying the "optimal" shaper is in general infeasible, as it requires the entire network state to be known. We then proposed a heuristic choice that depends upon the hop-length of the flow. Such a choice is simple to compute and varies gracefully between the known optimal choices for the limiting values of the hop-length. For a realistic traffic scenario where flows have varying hop-lengths, our proposed shaper choice is shown via simulation to yield significantly larger network utilizations than GPS as well as RC-EDF disciplines that employ shaper parameters chosen independent of the flow hop-length.

Deterministic frameworks are in general excessively conservative, and result in reduced network utilizations. The performance of RC-EDF in the statistical setting, where the end-to-end delay guarantees are statistical instead of worst-case, is addressed in [25].

REFERENCES

[1] J. C. R. Bennett, D. Stephens, and H. Zhang. High speed, scalable, and accurate implementation of fair queueing algorithms in atm networks. In *Proceedings of ICNP '97*, pages 7–14, October 1997.

[2] F. M. Chiussi and A. Francini. Implementing fair queueing in atm switches: The discrete-rate approach. In *Proceedings of INFOCOM '98*, San Francisco, CA, March 1998.

[3] F. M. Chiussi, A. Francini, and J. Kneur. Implementing fair queueing in atm switches – part 2: The logarithmic calendar queue. In *Proceedings of GLOBECOM '97*, pages 519–525, November 1997.

[4] F. M. Chiussi and V. Sivaraman. Achieving high utilization in guaranteed services networks using early-deadline-first scheduling. In *International Workshop on Quality of Service*, pages 209–217, Napa, California, May 1998.

[5] R. Cruz. A calculus for network delay, part i: Network elements in isolation. *IEEE Transactions on Information Theory*, 37:114–121, January 1991.

[6] A. Demers, S. Keshav, and S. Shenker. Analysis and simulation of a fair queueing algorithm. In *Proceedings of ACM SIGCOMM*, pages 3–12, 1989.

[7] W. Feng, F. Jahanian, and S. Sechrest. An optimal bandwidth allocation strategy for the delivery of compressed prerecorded video. *ACM/Springer-Verlag Multimedia Systems Journal*, 5(5):297–309, September 1997.

[8] D. Ferrari and D. Verma. A scheme for real-time channel establishment in wide-area networks. *IEEE Journal on Selected Areas in Communications*, 8(3):368–379, April 1990.

[9] V. Firoiu, J. Kurose, and D. Towsley. Efficient admission control of piecewise linear traffic envelopes at edf schedulers. *IEEE/ACM Transactions on Networking*, 6(5):558–570, October 1998.

[10] L. Georgiadis, R. Guérin, and A. Parekh. Optimal multiplexing on a single link: Delay and buffer requirements. *IEEE Transactions on Information Theory*, 43(5):1518–1535, September 1997.

[11] L. Georgiadis, R. Guérin, V. Peris, and K. Sivarajan. Efficient network qos provisioning based on per node traffic shaping. *IEEE/ACM Transactions on Networking*, 4(4):482–501, August 1996.

[12] M. Grossglauser, S. Keshav, and D. N. C. Tse. Rcbr: A simple and efficient service for multiple time-scale traffic. *IEEE/ACM Transactions on Networking*, 5(6):741–755, December 1997.

[13] S. Keshav. *An Engineering Approach to Computer Networking: ATM Networks, the Internet, and the Telephone Network.* Addison-Wesley, Ithaca, New York, 1996.

[14] E. Knightly and P. Rossaro. On the effects of smoothing for deterministic qos. *Distributed Systems Engineering Journal: Special Issue on Quality of Service*, 4(1):3–15, March 1997.

[15] J. Liebeherr, D. Wrege, and D. Ferrari. Exact admission control for networks with a bounded delay service. *IEEE/ACM Transactions on Networking*, 4(6):885–901, December 1996.

[16] J. M. McManus and K. W. Ross. Video-on-demand over atm: Constant-rate transmission and transport. *IEEE Journal on Selected Areas in Communications*, 14(6):1087–1098, August 1996.

[17] 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*, 1(3):344–357, June 1993.

[18] A. Parekh and R. Gallager. A generalized processor sharing approach to flow control in integrated services networks: The multiple node case. *IEEE/ACM Transactions on Networking*, 2(2):137–150, April 1994.

[19] V. Peris. *Architecture for Guaranteed Delay Service in High Speed Networks.* PhD thesis, University of Maryland, College Park, 1997.

[20] A. R. Reibman and A. W. Berger. On vbr video teleconferencing over atm networks. In *Proceedings of GLOBECOM '92*, pages 314–319, Orlando, FL, USA, Decemeber 1992.

[21] A. R. Reibman and A. W. Berger. Traffic descriptors for vbr video teleconferencing over atm networks. *IEEE/ACM Transactions on Networking*, 3(3):329–339, June 1995.

[22] J. L. Rexford, A. G. Greenberg, and F. G. Bonomi. Hardware-efficient fair queueing architectures for high-speed networks. In *Proceedings of INFOCOM '96*, pages 638–646, San Francisco, CA, March 1996.

[23] O. Rose. Statistical properties of mpeg video traffic and their impact on traffic modeling in atm systems. Technical Report TR 101, University of Würzburg, Würzburg, Germany, February 1995. The MPEG coded video traces are available at ftp://ftp-info3.informatik.uni-wuerzburg.de/pub/MPEG/.

[24] J. D. Salehi, Z.-L. Zhang, J. Kurose, and D. Towsley. Supporting stored video: Reducing rate variability and end-to-end resource requirements through optimal smoothing. *IEEE/ACM Transactions on Networking*, 6(4):397–410, August 1998.

[25] V. Sivaraman and F. M. Chiussi. Providing end-to-end statistical delay guarantees with earliest deadline first scheduling and per-hop traffic shaping. In *Proceedings of INFOCOM 2000*, Tel Aviv, Israel, March 2000.

[26] D. Verma, H. Zhang, and D. Ferrari. Guaranteeing delay jitter bounds in packet switching networks. In *Proceedings of TRICOMM*, pages 35–46, Chapel Hill, NC, April 1991.

[27] T. Wu and E. W. Knightly. Buffering vs. smoothing for end-to-end qos: Fundamental issues and comparison. In *Performance '99*, Istanbul, Turkey, October 1999.

[28] H. Zhang. Service disciplines for guaranteed performance service in packet-switching networks. *Proceedings of the IEEE*, 83(10):1374–1396, October 1995.

[29] H. Zhang and D. Ferrari. Rate-controlled service disciplines. *Journal of High Speed Networks*, 3(4):389–412, 1994.

[30] J. Zhang and J. Hui. Applying traffic smoothing techniques for quality of service control in vbr video transmissions. *Computer Communications*, 21(4):375–389, April 1998.

APPENDICES

I. PROOF OF LEMMA 4

From (1), the admissibility criteria for workload $\mathcal{U}$ is seen to be $\forall t \geq 0: \; D_{\mathcal{U}}(t) \leq F_{\mathcal{W}}(t)$. Since all traffic envelopes $A_i(t)$ are non-decreasing functions of t, so is the service demand $D_{\mathcal{U}}$ of workload $\mathcal{U}$. Thus the above condition is equivalent to $\forall t \geq 0: \; D_{\mathcal{U}}(t) \leq \min_{t' \geq t} F_{\mathcal{W}}(t')$. The quantity on the right is nothing but $R_{\mathcal{W}}(t)$; this proves the result. $\triangle$

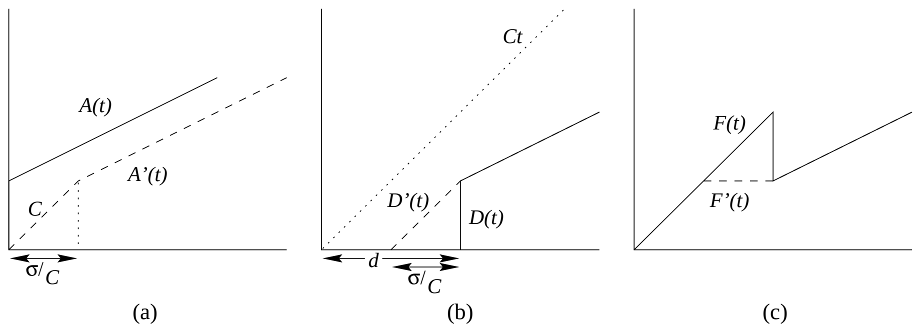

Fig. 3. The (a) envelopes, (b) service demands and (c) residual capacities for flows f (solid lines) and f' (dashed lines) in the proof of lemma 6.

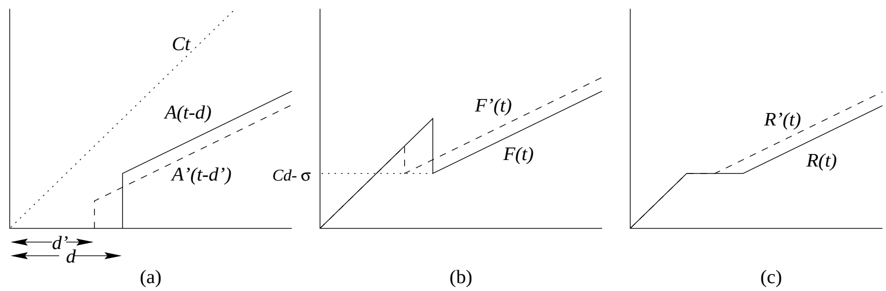

Fig. 4. The (a) service demands, (b) residual capacities and (c) effective residual capacities for flows f (solid lines) and f' (dashed lines) in the proof of lemma 7.

II. Proof of Lemma 5

If part: Say $R_{\mathcal{A}} \succeq R_{\mathcal{B}}$. Then any workload $\mathcal{U}$ satisfying $\forall t \geq 0: \ D_{\mathcal{U}}(t) \leq R_{\mathcal{B}}(t)$ (i.e., the workload $\mathcal{B} \cup \mathcal{U}$ is feasible) also satisfies $\forall t \geq 0: \ D_{\mathcal{U}}(t) \leq R_{\mathcal{A}}(t)$, and so the workload $\mathcal{A} \cup \mathcal{U}$ is also feasible.

Only if part: Say $R_{\mathcal{A}} \not\succeq R_{\mathcal{B}}$. Then $\exists \tau: \ R_{\mathcal{B}}(\tau) - R_{\mathcal{A}}(\tau) = 2\delta > 0$. Consider workload $\mathcal{U}$ consisting of a single flow with envelope $A(t) = R_{\mathcal{B}}(\tau) - \delta + \epsilon t$ and delay requirement τ. Since $R_{\mathcal{B}}(t)$ is monotonically non-decreasing and $\lim_{t\to\infty} R_{\mathcal{B}}(t)/t > 0$, ϵ can be chosen small enough such that $\forall t \geq 0: \ D_{\mathcal{U}}(t) \leq R_{\mathcal{B}}(t)$, thus making the workload $\mathcal{B} \cup \mathcal{U}$ feasible. But the workload $\mathcal{A} \cup \mathcal{U}$ is not feasible, since $D_{\mathcal{U}}(\tau) = R_{\mathcal{B}} - \delta \not\leq R_{\mathcal{A}}(\tau)$. The existence of a workload $\mathcal{U}$, such that $\mathcal{B} \cup \mathcal{U}$ is admissible while $\mathcal{A} \cup \mathcal{U}$ is not, completes the proof. $\triangle$

III. Proof of Lemma 6

The envelope $A(t)$, service demand $D(t)$ and residual capacity $F(t)$ for the workload consisting of flow f with delay bound d are shown in figure 3, as are the corresponding quantities $A'(t)$, $D'(t)$ and $F'(t)$ for the workload consisting of flow f' with delay bound d'. It is easy to see that the residual effective capacities are equivalent in both cases, and therefore by lemma 5 the result follows. $\triangle$

IV. Proof of Lemma 7

Consider the effective residual capacities $R(t)$ when the workload consists of the flow f with delay bound d, and $R'(t)$ when the workload consists of the flow f' with delay bound $d' = d - \frac{\sigma - \sigma'}{C}$. The envelopes, residual capacities and effective residual capacities are depicted graphically in figure 4, and it is easily verified that $R' \succeq R$. From lemma 5, it follows that for any workload, if flow f is feasible, so if flow f'. Thus f' is admissible if f is. $\triangle$