Queueing Systems 38, 471–483, 2001

The Buffer-Bandwidth Trade-off Curve is Convex

KRISHNAN KUMARAN and MICHEL MANDJES {kumaran, michel}@research.bell-labs.com
Bell Laboratories/Lucent Technologies, 600 Mountain Ave., P.O. Box 636, Murray Hill, NJ 07974-0636, USA

Received 24 October 2000; Revised 1 March 2001

Abstract. To achieve a constant overflow probability, the two queueing resources, viz. buffer and bandwidth, can be traded off. In this paper we prove that, under general circumstances, the corresponding trade-off curve is *convex* in the 'many-sources scaling'. This convexity enables optimal resource partitioning in a queueing system supporting heterogeneous traffic, with heterogeneous quality-of-service requirements.

Keywords: queueing, large deviations, buffer-bandwidth trade-off, resource partitioning, generalized processor sharing

1. Introduction

A generic model of a queue in the following consists of i.i.d. sources (say n) that feed into a buffer, which is emptied at a constant rate C (i.e., the link capacity, or bandwidth). A major performance measure is the probability that the buffer level is above some level B, commonly called the *overflow probability*. To achieve a fixed overflow probability the queueing resources, viz. buffer and link rate, can be traded off against each other. In this paper we study this trade-off in the so-called 'many sources regime' [9]. In this regime buffer and bandwidth are scaled by the number of sources: $B \equiv nb$ and $C \equiv nc$, where n is typically large.

Large deviations techniques can be used to show that, under general conditions, the overflow probability decays exponentially in n [1,2]. The exponential decay rate can be expressed explicitly in terms of b, c, and the stochastic characteristics of the sources. Our objective is to characterize the way b and c trade off, with the decay rate held fixed.

Our main result is that the corresponding trade-off curve is *convex*. This convexity has important consequences for differential Quality-of-Service (QoS), i.e., multiple classes with class-specific overflow probabilities. It enables the computation of a resource partitioning that optimizes the number of admissible sources. A common technique to implement differential QoS is generalized processor sharing (GPS). We relate the partitioning to the design of weights for GPS.

This note is structured as follows. Section 2 describes our main result: the trade-off curve between buffer and bandwidth is convex, under the many-sources scaling. Section 3 considers partitioning procedures for heterogeneous QoS, and shows that straightforward algorithms directly follow from the convexity. Section 4 briefly discusses our results and some directions for future research.

2. Convexity of tradeoff curve under many-sources scaling

2.1. Model and preliminaries

We consider traffic from n independent, statistically identical, stationary sources feeding into a buffered resource. This resource is modeled as a queue with constant depletion rate C. Define

$$A(t) := \{\text{Traffic generated by a single source, in steady state, in a time interval of length } t\}.$$

Let $A_i(t)$ be defined analogously, but now specifically for the ith source. We are interested in the steady-state probability of the buffer content exceeding level B, denoted by $p(B, C)$. As known, it holds that (under the usual stationarity conditions)

$$p(B, C) = \mathbb{P}\left(\exists t \in T\colon \sum_{i=1}^{n} A_i(t) - Ct > B\right).$$

The queue can be defined in discrete time ($T = \mathbb{N}$) and continuous time ($T = \mathbb{R}_+$).

Many-sources scaling. We rescale the resources by the number of sources: $C \equiv nc$ and $B \equiv nb$, where n is typically large. This scaling, introduced by Weiss [9], is natural since queues in real networks tend to multiplex large numbers of sources. We assume that the system is stable and nontrivial:

$$\mathbb{E}A(t) < ct.$$

In the scaled model we define

$$p_n(b, c) := \text{steady-state probability that the buffer content exceeds level } nb.$$

Exponential decay. Under non-restrictive conditions, the overflow probability $p_n(b, c)$ decays *exponentially* in the scaling parameter n. We define the corresponding exponential *decay rate*:

$$I(b, c) := -\lim_{n\to\infty} \frac{1}{n} \log p_n(b, c).$$

The key result on $I(b, c)$, based on large deviations arguments, is given below in theorems 2.2 and 2.4. A major contribution to the development of these results was given by Botvich and Duffield [1], whereas related results were derived in [2,8]. Recently, a significant improvement was made by Likhanov and Mazumdar [6]. We will distinguish between discrete and continuous time.

Discrete time. In discrete time, we will rely on the results by Likhanov and Mazumdar [6]. They impose the following assumption on the input traffic.

Assumption 2.1 (See [6]). Assume $I_t(b,c)$ is larger than $\alpha \log t$, for t large enough, and a positive α, where

$$I_t(b,c) := \sup_{\theta>0}\bigl(\theta(b+ct) - \log \mathbb{E}\, \mathrm{e}^{\theta A(t)}\bigr).$$

Theorem 2.2 (Decay rate in discrete time). Under assumption 2.1,

$$I(b,c) = \inf_{t\in\mathbb{N}} I_t(b,c) = \inf_{t\in\mathbb{N}} \sup_{\theta>0}\bigl(\theta(b+ct) - \log \mathbb{E}\, \mathrm{e}^{\theta A(t)}\bigr). \tag{1}$$

In fact, [6] proves a stronger result: a subexponential function $f(\cdot)$ (i.e., $\log f(n) = \mathrm{o}(n)$) is found such that $p_n(b,c)\exp(nI(b,c)) f(n) \to 1$ as $n \to \infty$. To do so, it is assumed that the infimum over $t \in \mathbb{N}$ is attained at a *unique* $t^\star$; we do not need that assumption here.

Continuous time. In continuous-time, an additional assumption has to be made on the regularity of the traffic. For this purpose we here use [1, hypothesis 1(iv)], which essentially implies that the decay rate for discrete time ($T = \mathbb{N}$) carries over to continuous time ($T = \mathbb{R}_+$). This is proven analogously to the proof of [1, theorem 1, p. 302].

Hypothesis 1(iv) is stated as follows. Define

$$A^n_{t,r} := \sup_{0<r'<r} \sum_{i=1}^{n} A_i(t) - A_i\bigl(t-r'\bigr).$$

Then it is required that

$$\limsup_{r\downarrow 0}\, \limsup_{n\to\infty} \frac{1}{n} \sup_{t\geqslant 0} \log \mathbb{E} \exp\bigl(\theta A^n_{t,r}\bigr) \leqslant 0.$$

It is easily verified that, due to the stationarity and i.i.d. assumptions, this requirement reduces to the following.

Assumption 2.3 (See [1]). Assume that for all $\theta \in \mathbb{R}$,

$$\limsup_{r\downarrow 0} \log \mathbb{E} \exp\Bigl(\theta \sup_{0<r'<r} A(r')\Bigr) \leqslant 0.$$

Theorem 2.4 (Decay rate in continuous time). Under assumptions 2.1 and 2.3,

$$I(b,c) = \inf_{t>0} \sup_{\theta>0}\bigl(\theta(b+ct) - \log \mathbb{E}\, \mathrm{e}^{\theta A(t)}\bigr). \tag{2}$$

Applicability. The striking feature of the above theorems is that they are applicable for almost all types of traffic sources of practical interest. Short-range dependent sources satisfy the conditions (Markov fluid sources, Markov-modulated Poisson processes, on–off sources with light-tailed on-times), see [1], but also long-range dependent processes (fractional Brownian motion, on–off sources with heavy tailed on-times), see [3,6,7].

2.2. Results

Using the above large deviations results, we are in a position to show that buffer and bandwidth trade off in a *convex* manner, for fixed decay rate δ. To that end, first introduce some notation. Let $t^\star \in T$ and $\theta^\star$ be the parameter values that optimize (1), respectively (2), for given (b, c). Define

$$f(\theta, t) := \log \mathbb{E}e^{\theta A(t)} \quad \text{and} \quad f_\theta(\theta, t) := \frac{\partial}{\partial \theta} \log \mathbb{E}e^{\theta A(t)}.$$

For given $I(b, c) = \delta$, we necessarily have

$$\theta^\star\big(b + ct^\star\big) - f\big(\theta^\star, t^\star\big) = \delta, \tag{3}$$

$$b + ct^\star = f_\theta\big(\theta^\star, t^\star\big). \tag{4}$$

Equality (3) is by definition of $t^\star$ and $\theta^\star$. Clearly, (4) is a first order condition necessary for optimality (obtained by differentiating the objective function to θ). The following lemma gives an explicit relation between b and c, for fixed δ.

Lemma 2.5. For fixed decay rate δ, resources b and c are related through

$$b = \sup_{t \in T} f_\theta\big(\theta_\delta(t), t\big) - ct, \tag{5}$$

where $\theta_\delta(t)$ is uniquely determined by

$$\theta_\delta(t) f_\theta\big(\theta_\delta(t), t\big) - f\big(\theta_\delta(t), t\big) = \delta. \tag{6}$$

Proof. Because the decay rate is δ, we have that for all $t \in T$

$$\sup_{\theta > 0} \theta(b + ct) - f(\theta, t) \geqslant \delta,$$

with equality at $t^\star$. It is not hard to see that this is equivalent to saying that for all $t \in T$

$$\inf_{\theta > 0} \left(\frac{\delta + f(\theta, t)}{\theta}\right) - ct \leqslant b, \tag{7}$$

with equality at $t^\star$, or in other words,

$$b = \sup_{t \in T} \left(\inf_{\theta > 0} \left(\frac{\delta + f(\theta, t)}{\theta}\right) - ct\right).$$

For given t, the optimum over $\theta > 0$ (say, $\theta_\delta(t)$) in (7) satisfies relation (6), as follows immediately from differentiation to θ. The convexity of f in θ immediately implies that $\theta f_\theta(\theta, t) - f(\theta, t)$ increases in θ. In addition, the function vanishes at the origin. Hence, the value of $\theta_\delta(t)$ is uniquely determined for any $\delta > 0$.

Now it is easily seen that, due to (6),

$$b = \sup_{t \in T} \frac{\delta + f(\theta_\delta(t), t)}{\theta_\delta(t)} - ct = \sup_{t \in T} f_\theta(\theta_\delta(t), t) - ct. \qquad \square$$

The above lemma immediately implies the convexity of the *b*–*c* trade-off curve. Note the striking form of equation (5) which is identical to the maximum buffer content computation for a deterministic arrival process specified by $\phi_\delta(t) := f_\theta(\theta_\delta(t), t)$. Thus, $\phi_\delta(t)$ can be interpreted as an 'effective arrival process' for the given traffic and loss exponent δ. We later enumerate some interesting properties of $\phi_\delta(t)$.

Lemma 2.6. For given δ, it holds that

$$\frac{\mathrm{d}b}{\mathrm{d}c} = -t^\star \tag{8}$$

is increasing in c. Hence, the trade-off curve is convex.

Proof. The stated convexity and relation (2.6) immediately follow by observing that b is a *Legendre transform* $\sup_t f_\theta(\theta_\delta(t), t) - ct$, see (5). Due to the optimality of $t^\star$,

$$\frac{\mathrm{d}}{\mathrm{d}t} f_\theta(\theta_\delta(t), t)\Big|_{t^\star} = c,$$

while the product rule of differentiation yields

$$\frac{\mathrm{d}b}{\mathrm{d}c} = \left(\frac{\mathrm{d}f_\theta(\theta_\delta(t), t)}{\mathrm{d}t}\Bigg|_{t^\star} - c\right)\frac{\mathrm{d}t^\star}{\mathrm{d}c} - t^\star \equiv -t^\star$$

as claimed in lemma 2.6. □

The convexity proof has the simple pictorial interpretation shown in figure 1. As c increases, figure 1(i) depicts the property that the function $\phi_\delta(t) - ct$ has the property that its local maxima shift towards smaller t, and the further local maxima fall by a greater amount. Thus, the *global* maximum can only shift leftward – either by the left-shift of a specific local maximum, or by a different local maximum to the left of the original becoming globally larger. During this process of increasing c, note that the height of the global optimum itself varies continuously, and each 'jump' in the global optimum (by a left shift to a different local maximum) represents a nondifferentiable point on the *b*–*c* trade-off curve. A similar explanation can be given for figure 1(ii), which depicts

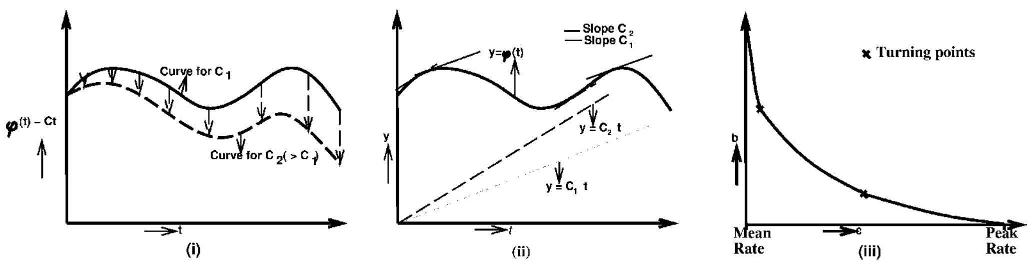

Figure 1. Pictorial proof of lemma 2.6. All local maxima (minima) move leftward (rightward) as C increases, and local maxima/minima further away from the origin diminish more than nearby ones. (iii) depicts the resulting piecewise smooth *b*–*c* trade-off curve.

the left-shift of points of slope c on the $\phi_\delta(t)$ curve with increasing c. The b–c trade-off curve itself is thus continuous and piece-wise differentiable, as shown in figure 1(iii). Note that the proof does not require continuity of $\phi_\delta(t)$, thus also applying to the discrete time case.

Notice that the above observations are just pictorial explanantions of the well-known convexity of Legendre transforms, included here to add intuition.

3. Explicit calculations for continuous time

For continuous t ($T = \mathbb{R}_+$) all calculations can be done explicitly. Assume differentiability of $f(\theta, t)$ with respect to t.

3.1. Properties of the effective arrival process $\phi_\delta(t)$

The next lemma shows that the $\phi_\delta(t)$ function corresponds to a proper arrival process, in that it is a semi-monotonically increasing function of t. Hence $\phi_\delta(t)$ can be interpreted as a true 'effective arrival process', which, if known, can be directly used for traffic management purposes (admission control, resource allocation). If not known explicitly, it can potentially be estimated through measurements.

Lemma 3.1. The following are direct consequences of equations (5) and (6):

$$\frac{\mathrm{d}\phi_\delta(t)}{\mathrm{d}t} = \frac{f_t(\theta_\delta(t), t)}{\theta_{\delta(t)}} \tag{9}$$

for fixed δ,

$$\frac{\partial\phi_\delta(t)}{\partial\delta} = \frac{1}{\theta_{\delta(t)}} \tag{10}$$

for fixed t. (9) implies semi-monotonic increase of $\phi_\delta(t)$ with respect to t for given δ, which is required of a causally valid arrival process. (10) implies semi-monotonicity and concavity of $\phi_\delta(t)$ with respect to δ for fixed t, which is also to be expected for reasonable arrival processes.

Proof. Directly differentiating (6) with respect to t for fixed δ, we get

$$\begin{aligned}\frac{\mathrm{d}\phi_\delta(t)}{\mathrm{d}t} &\equiv \frac{\mathrm{d}f_\theta(\theta_\delta(t), t)}{\mathrm{d}t} = \frac{\mathrm{d}((\delta + f(\theta_\delta(t), t))/\theta_\delta(t))}{\mathrm{d}t}\\ &= \frac{f_t(\theta_\delta(t), t)}{\theta_\delta(t)} + \frac{1}{\theta_\delta(t)}\frac{\mathrm{d}\theta_\delta(t)}{\mathrm{d}t}\left(f_\theta\big(\theta_\delta(t), t\big) - \frac{\delta + f(\theta_\delta(t), t)}{\theta_\delta(t)}\right) \equiv \frac{f_t(\theta_\delta(t), t)}{\theta_\delta(t)},\end{aligned}$$

where the last equality follows from (6) itself. Due to the non-negativity of $f(\theta, t)$ for non-negative $\theta,\ t$, it follows that $\phi_\delta(t)$ has a non-negative derivative with respect to time for fixed δ, and is hence non-decreasing as claimed. Note that the derivative has the natural interpretation of the 'effective arrival rate', and is obtainable as the expectation value of the original arrival rate with respect to a *twisted* distribution, where $\theta_\delta(t)$ denotes the loss-dependent twist parameter.

Similarly differentiating (6) with respect to δ for fixed t, we get

$$\begin{aligned}\frac{\partial\phi_\delta(t)}{\partial\delta} \equiv \frac{\partial f_\theta(\theta_\delta(t),t)}{\partial\delta} &= \frac{\partial((\delta+f(\theta_\delta(t),t))/\theta_\delta(t))}{\partial\delta}\\ &= \frac{1}{\theta_\delta(t)} + \frac{1}{\theta_\delta(t)}\frac{\partial\theta_\delta(t)}{\partial\delta}\left(f_\theta\big(\theta_\delta(t),t\big) - \frac{\delta+f(\theta_\delta(t),t)}{\theta_\delta(t)}\right) \equiv \frac{1}{\theta_\delta(t)}\end{aligned}$$

as claimed. Note that (6) directly implies that $\theta_\delta(t)$ is an increasing function of δ for fixed t due to the monotonicity of the LHS. In fact it is possible to show by direct differentiation that

$$\frac{\partial\theta_\delta(t)}{\partial\delta} = \frac{1}{f_{\theta\theta}(\theta_\delta(t),t)}$$

with the RHS clearly non-negative due to the convexity of $f(\theta,t)$ with respect to θ. Thus, we see that

$$\frac{\partial\phi_\delta(t)}{\partial\delta}$$

has a non-negative, but decreasing, value as δ increases, hence proving its semi-monotonicity and concavity, respectively. □

3.2. *The second derivative of the trade-off curve*

We now compute the second derivative of the *b*–*c* trade-off curve, and explicitly show that this is positive. Note that (3), (4) and

$$c\theta^\star = f_t(\theta^\star,t^\star), \quad \text{with } f_t(\theta,t) := \frac{\partial}{\partial t}\log\mathbb{E}\,\mathrm{e}^{\theta A(t)} \tag{11}$$

provide three equations in the four unknowns b, c, $\theta^\star$, and $t^\star$. Hence b, $\theta^\star$, and $t^\star$ can be considered as functions of c. Thus these equations hold in a *functional* sense, i.e., for all c, and therefore we may differentiate them to compute various derivatives. Our aim here is to prove that

$$\frac{\mathrm{d}^2b}{\mathrm{d}c^2} \equiv -\frac{\mathrm{d}t^\star}{\mathrm{d}c}$$

is non-negative, to prove convexity of the buffer-bandwidth trade-off curve.

This may be accomplished by combining (3) and (4) to $\delta + f(\theta^\star,t^\star) = \theta^\star f_\theta$, as before, followed by differentiation of the above and (11) as

$$\begin{aligned} f_t\frac{\mathrm{d}t^\star}{\mathrm{d}c} &= \theta^\star\left(f_{\theta\theta}\frac{\mathrm{d}\theta^\star}{\mathrm{d}c} + f_{\theta t}\frac{\mathrm{d}t^\star}{\mathrm{d}c}\right),\\ \theta^\star + c\frac{\mathrm{d}\theta^\star}{\mathrm{d}c} &= f_{\theta t}\frac{\mathrm{d}\theta^\star}{\mathrm{d}c} + f_{tt}\frac{\mathrm{d}t^\star}{\mathrm{d}c},\end{aligned}$$

defining

$$f_{\theta\theta} := \frac{\partial^2}{\partial\theta^2} f\bigl(\theta^\star, t^\star\bigr), \qquad f_{\theta t} := \frac{\partial^2}{\partial\theta\partial t} f\bigl(\theta^\star, t^\star\bigr), \qquad f_{tt} := \frac{\partial^2}{\partial t^2} f\bigl(\theta^\star, t^\star\bigr),$$

assuming the existence of all derivatives mentioned above. Solving for $\mathrm{d}t^\star/\mathrm{d}C$ in the above directly leads to the relation

$$\frac{\mathrm{d}^2 b}{\mathrm{d}c^2} = -\frac{\mathrm{d}t^\star}{\mathrm{d}c} = -\frac{\theta^\star f_{\theta\theta}}{f_{\theta\theta} f_{tt} - (c - f_{\theta t})^2}. \tag{12}$$

Now note that f_{tt} is positive at t^*, where the mimimum over t is attained. On the other hand, $f_{\theta\theta}$ is negative, because there the maximum over θ is attained. This establishes that the right-hand side of (12) is non-negative, thus proving that the *b–c* trade-off curve is convex.

4. Optimal partitioning of queueing resources

We now consider the case of multiple heterogeneous classes of sources sharing common resources B and C. The sources are homogeneous within each class, but there is heterogeneity across classes. We consider the problem of optimally partitioning (B, C) into (B_i, C_i) among the classes. 'Optimal' here refers to maximizing the size of the admissible region by maximizing n for a given *connection mix vector*

$$\boldsymbol{\eta} = \{\eta_1, \ldots, \eta_K\} := \left\{\frac{n_1}{n}, \ldots, \frac{n_K}{n}\right\}, \quad \text{where } n := \sum_{i=1}^{K} n_i.$$

Clearly, partitioning of resources can support diverse QoS requirements by protecting individual classes. In general, it loses the multiplexing advantage obtained by sharing across classes. Thus for homogeneous QoS requirements, we get a conservative solution to the FIFO system.

We begin this section by characterizing (and interpreting) the solution to the optimal partitioning problem. Then we indicate how to use this solution when QoS differentiation is offered through generalized processor sharing (GPS).

4.1. *Solution to the partitioning problem*

In this section we assume that the 'large deviations approximation' is accurate. In other words: in a system with n homogeneous inputs, and resources B and C,

$$p(B, C) \approx \exp\Bigl(-\inf_{t\in T}\sup_{\theta}\bigl(\theta(B + Ct) - n \log \mathbb{E} \exp\bigl(\theta A(t)\bigr)\bigr)\Bigr). \tag{13}$$

Notice that we neglect a prefactor that is of the order $n^{-3/2}$, see [6], which is obviously less significant than the exponential term. As concluded in the previous section, for a fixed value of $p(B, C)$, and n large (but constant), B and C trade off in a convex way.

Now consider the system with K traffic classes with heterogeneous QoS requirements specified by L_i, the maximum admissible loss ratio for a class i source. Let $\Delta_i := -\log L_i$. Suppose that there are n_i sources of type i, and they are assigned resources (B_i, C_i). Then, based on approximation (13), there is a function R_i, such that

$$R_i(C_i, B_i, n_i) \geqslant \Delta_i \tag{14}$$

indicates that the n_i sources get the required QoS; obviously R_i is (minus) the log of the overflow probability when resources (B_i, C_i) are available to n_i sources.

Formulation as convex programming problem. To optimize the admissible region (for a fixed connection mix η) we seek to solve

$$\begin{aligned}
&\text{Maximize } n \\
&\text{subject to} \quad R_i(C_i, B_i, n\eta_i) \geqslant \Delta_i \quad \forall\, i, \\
&\qquad\qquad \sum_i C_i \leqslant C, \qquad \sum_i B_i \leqslant B.
\end{aligned} \tag{15}$$

We first observe that, at optimum, all the loss constraints in (15) would hold with equality, as otherwise we may reduce B_i and/or C_i for the corresponding class and admit more connections.[1] We hence suppose that, for fixed n (in particular the optimal value), we may invert (14) to obtain the buffer as a function of the other parameters, i.e., $B_i = B_i(C_i, n\eta_i, \Delta_i)$. As derived in the previous section, this function is convex in C_i under very mild conditions.

Now consider the intermediate problem:

$$\begin{aligned}
&\text{Minimize} \ \sum_i B_i(C_i, n\eta_i, \Delta_i) \\
&\text{subject to} \quad \sum_i C_i \leqslant C.
\end{aligned} \tag{16}$$

Obviously, we have to find the largest n such that the mathematical program (16) has a solution that does not exceed B. Due to the convexity of (B_i, C_i) curves, (16) can be recognized as a standard convex minimization to which the *Strong Lagrangian Principles* can be applied, see, for example, [10]. It yields the following *Kuhn–Tucker* conditions:

$$-\frac{\mathrm{d}B_i}{\mathrm{d}C_i} = \frac{\partial R_i/\partial C_i}{\partial R_i/\partial B_i} = \lambda$$

for some global (class-independent) non-negative Lagrange multiplier λ.

[1] We ignore the fact that n needs to be integral, but this is of minor consequence when $n \gg 1$.

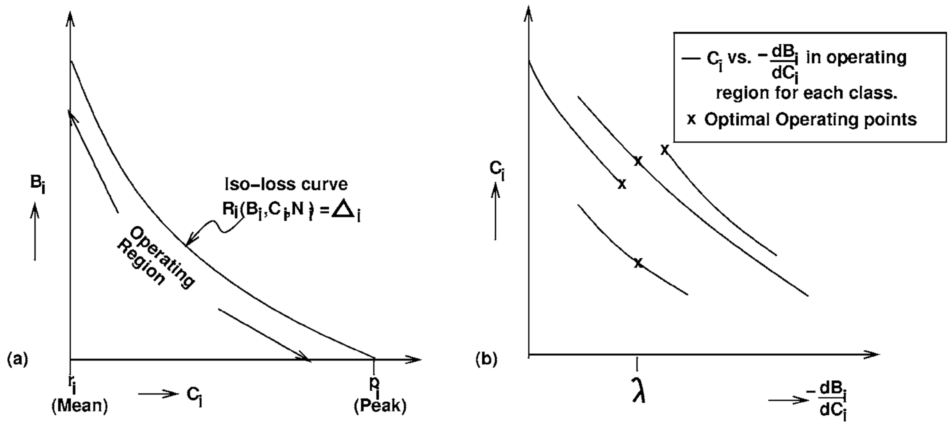

Figure 2. Illustration of optimal partitioning solution.

It, hence, follows that the maximum value of n retains feasibility of the following conditions for some λ:

$$\begin{aligned} &R_i(C_i, B_i, n\eta_i) = \Delta_i, \\ &-\frac{\mathrm{d}B_i}{\mathrm{d}C_i} = \frac{\partial R_i/\partial C_i}{\partial R_i/\partial B_i} = \lambda, \\ &\sum_i C_i \leqslant C, \qquad \sum_i B_i \leqslant B. \end{aligned} \tag{17}$$

Note that λ represents the slope of each of the bandwidth-buffer trade-off curves at the optimal operating point. This is pictorially illustrated in figure 2(a).

Interpretation. Interestingly, the optimal λ equals $t^\star$, cf. equation (8). As shown by Wischik [11], $t^\star$ can be interpreted as the most likely time to overflow (i.e., the duration of a typical busy period of the queue, leading to overflow). In other words: the system capacity is maximized, if all the $t^\star$'s of the subsystems match. This seems reasonable, since if one subsystem has overflow while other subsystems are still not completely utilized, the chosen partitioning of the B_i and C_i seems suboptimal.

Clearly C_i should always be chosen such that the corresponding system is stable, i.e., it should be larger than the mean input rate. The smallest possible C_i could also be constrained by other considerations, such as a limit on the *maximum delay*, D_i. In this case, it is simple to add the requirement $B_i/C_i \leqslant D_i$, yielding $B_i^{\max}$, $C_i^{\min}$ as the intersection of the line $B_i = D_i C_i$ with the loss trade-off curve $R_i(B_i, C_i, n_i) = \Delta_i$. Also, in many practical situations traffic streams are restricted to a peak rate, i.e., a maximum on the traffic rate generated by a single source. These bounds on C_i imply that the λ condition in (17) may not be satisfiable for all classes, see figure 2(b). Then the algorithm has to be adapted such that λ is replaced by the closest achieved value for class i, and choose the corresponding operating point (B_i, C_i) on the trade-off curve. These modifications are justified by standard results from theory of convex programming, and other convex constraints can be imposed on the operating region without altering our methods significantly provided the constraints do not destroy the convex program structure. The resulting operating point is hence uniquely specified for given λ. In conclusion, the optimal solution satisfies the following property: *There exists non-negative* λ *such that the choice*

$$\frac{\mathrm{d}B_i}{\mathrm{d}C_i} = \begin{cases} \lambda & \forall\, i\colon\ \lambda_i^{\min} \leqslant \lambda \leqslant \lambda_i^{\max}, \\ \lambda_i^{\min} & \forall\, i\colon\ \lambda < \lambda_i^{\min}, \\ \lambda_i^{\max} & \forall\, i\colon\ \lambda > \lambda_i^{\max}, \end{cases} \tag{18}$$

is the optimal solution to (15). $\lambda^{\min}$ ($\lambda_i^{\max}$) is the minimum (maximum) absolute slopes realized within the operating region of the class i trade-off curves.

Example. We now give an illustration of the optimal partitioning solution (18). First consider $n\eta$ fractional Brownian motion sources (without drift), with Hurst parameter H, using a queue with resources (B, C). Using $\log \mathbb{E} \exp(\theta A(t)) = \frac{1}{2}\theta^2\sigma^2 t^{2H}$, it easily follows from (13) that the buffer-bandwidth trade-off curve is given by

$$B^{1-H} C^{H} = \sqrt{2\Delta n\eta\sigma^2 H^H (1-H)^{1-H}}, \tag{19}$$

when the QoS requirement Δ is imposed. Notice that the buffer-bandwidth trade-off is indeed convex, and is described through a so-called Cobb–Douglas substitution function.

Now consider the setting of section 3.1. Let type i be fractional Brownian motion with parameter H_i. With evident notation, for any i, it follows from (19) that there exist $K_i = K_i(n)$ such that

$$B_i = K_i(n) C_i^{-\gamma_i}, \quad \text{where } \gamma_i := \frac{H_i}{1-H_i}, \quad \text{and} \quad K_i(n) = M_i n^{1/(2(1-H_i))},$$

for some M_i (independent of n). From equations (17), there is a class-independent number λ such that

$$\lambda = -\frac{\mathrm{d}B_i}{\mathrm{d}C_i} = K_i(n)\gamma_i C_i^{-\gamma_i - 1},$$

or

$$C_i(n, \lambda) = \left(\frac{K_i(n)\gamma_i}{\lambda}\right)^{1/(\gamma_i+1)} = \left(\frac{M_i\gamma_i}{\lambda}\right)^{1-H_i} \sqrt{n}$$

and

$$B_i(n, \lambda) = \left(\frac{K_i(n)\lambda^{\gamma_i}}{\gamma_i^{\gamma_i}}\right)^{1/(\gamma_i+1)} = \left(\frac{M_i\lambda^{\gamma_i}}{\gamma_i^{\gamma_i}}\right)^{1-H_i} \sqrt{n}.$$

From the buffer and bandwidth constraints $B = \sum_{i=1}^{K} B_i$ and $C = \sum_{i=1}^{K} C_i$ we obtain

$$C \sum_{i=1}^{K} \left(\frac{M_i\lambda^{\gamma_i}}{\gamma_i^{\gamma_i}}\right)^{1-H_i} = B \sum_{i=1}^{K} \left(\frac{M_i\gamma_i}{\lambda}\right)^{1-H_i}.$$

From the fact that the left-hand side is increasing in λ, whereas the right-hand side is decreasing in λ, it immediately follows that there is a unique solution.

4.2. Relation to GPS and buffer management

The above partitioning model can be used to offer differentiated QoS. An alternative mechanism to do so is Generalized Processor Sharing (GPS). In GPS, each class has its own buffer and bandwidth, but the bandwidth left unused by one class can be used by another class; for a more detailed explanation see, e.g., [4]. Hence, the partitioned system discussed in section 3.1 is a 'conservative description' of the corresponding GPS system. Therefore the (B_i, C_i) values of the partioned system can be used to determine the weights and buffer sizes in the corresponding GPS system.

Given the optimal (B_i, C_i) split of the partitioned system, the GPS weight ϕ_i for each class i connection can obviously be set conservatively as $\phi_i = C_i/(n_iC)$. This holds when the GPS is implemented such that there is *per-connection* queueing; if on the contrary the queueing is *per-class*, the weights would be $\Phi_i = C_i/C$. The corresponding buffer space must be guaranteed as well. This can be accomplished using the technique of virtual partitioning as described in [5]. The per-connection (respectively per-class) allocations described therein are easily recognized to be $b_i = B_i/n_i$ (respectively B_i) from the optimal partition.

5. Conclusions and remarks

We have proved that the buffer-bandwidth trade-off curve is convex in the asymptotic regime of many sources. One wonders whether this convexity holds under more general conditions. This question remains open.

The partitioned system is more conservative than the corresponding GPS system, since in the latter there is a higher degree of sharing of the link capacity across classes. The capacity of the partitioned system is determined in section 3, but the capacity of the corresponding GPS system (under the many-sources scaling) is, to our best knowledge, not known yet.

References

[1] D. Botvich and N. Duffield, Large deviations, the shape of the loss curve, and economies of scale in large multiplexers, Queueing Systems 20 (1995) 293–320.

[2] C. Courcoubetis and R. Weber, Buffer overflow asymptotics for a buffer handling many traffic sources, J. Appl. Probab. 33 (1996) 886–903.

[3] N. Duffield, Economies of scale in queues with sources having power-law large deviation scalings, J. Appl. Probab. 33 (1996) 840–857.

[4] A. Elwalid and D. Mitra, Design of generalized processor sharing schedulers which statistically multiplex heterogeneous QoS classes, in: *Proc. of IEEE INFOCOM 1999*, 1999, pp. 1220–1230.

[5] K. Kumaran and D. Mitra, Performance and fluid simulations of a novel shared buffer management system, in: *Proc. of IEEE INFOCOM 1998*, 1998, pp. 1449–1461.

[6] N. Likhanov and R. Mazumdar, Cell loss asymptotics in buffers fed with a large number of independent stationary sources, J. Appl. Probab. 36 (1999) 86–96.

[7] M. Mandjes and S. Borst, Overflow behavior in queues with many long-tailed inputs, Adv. in Appl. Probab. 32 (2000) 1150–1167.

[8] A. Simonian and J. Guibert, Large deviations approximation for fluid queues fed by a large number of on/off sources, IEEE J. Selected Areas Commun. 13 (1995) 1017–1027.

[9] A. Weiss, A new technique of analyzing large traffic systems, Adv. in Appl. Probab. 18 (1986) 506–532.

[10] P. Whittle, *Optimization under Constraints* (Wiley, New York, 1971).

[11] D. Wischik, Sample path large deviations for queues with many inputs, Ann. Appl. Probab. (2000) submitted.