Fair Background Data Transfers of Minimal Delay Impact

Costas Courcoubetis
Department of Informatics
Athens University of Economics and Business
Patission 76, Athens 10434, Greece
Email: courcou@aueb.gr

Antonis Dimakis
Department of Informatics
Athens University of Economics and Business
Patission 76, Athens 10434, Greece
Email: dimakis@aueb.gr

Abstract—In this paper we present a methodology for the design of congestion control protocols for background data transfers that have a minimal delay impact on short TCP transfers and compete for a target share of the leftover average capacity with other background TCP transfers. We analytically compute the optimal policy and show that it's delay performance can be well approximated by a weighted TCP policy, which maintains a target proportion of TCPs bandwidth at all times in order to achieve the same share of excess capacity. The relative approximation error is always less than 17.2% while it quickly decreases to zero as the number of coexisting background TCP flows increases.

Next, we consider a general utility-based fairness criterion for sharing the leftover average capacity, including a penalty term capturing the delay impact to short flows. The criterion is jointly optimized over all allocations of excess capacity to background flows (including TCP ones) on long timescales, and all bandwidth sharing policies on fast time scales. Even though the delay optimal sharing policy that solves the above optimization problem does not lead to distributed congestion control algorithms and more significantly, requires knowledge of the number of competing background TCP flows, both problems disappear under the weighted TCP policy. A distributed weight adjustment policy is considered where, at equilibrium, the overall performance is nearly optimal, with a quickly vanishing relative optimality error as the number of background TCP flows increases. We illustrate the methodology by giving two examples of congestion control algorithms for background transfers. Both achieve low delay for short flows relative to TCP, but at the same time they present strong incentives for adoption against incumbent low priority solutions in public environments.

I. Introduction

A key element of the success of the internet architecture is the ability to accommodate current and future needs of very diverse applications. Connection rates differ by few orders of magnitude, while file transfer sizes vary by more than ten orders of magnitude. Nevertheless this is achieved using only a handful of transport protocols, mainly TCP and its variants, which in essence allocate network bandwidth to flows continuously so as to achieve fair sharing at all times. Indeed TCP "fairness" or "friendliness" [1] has become a popular prescription for congestion control algorithms which intends to ensure equal sharing between flows.

Not all applications value instantaneous bandwidth equally though. It is valued more by web browsing flows than, say, background data transfers such as batch software updates. The former serve interactive tasks where low transfer delays are important, while the latter are indifferent to small temporal variations of their bandwidth share, provided the volume downloaded over a long time period remains the same. Recognizing this, specialized congestion control algorithms for background data transfers have recently emerged, e.g., TCP-LP [2], TCP-nice [3], uTorrent transport protocol [4], LEDBAT [5]. Such protocols are typically designed to behave as low priority traffic, i.e., opting for low delay impact on short flows, without considering fairness relative to other background transfers *carried by* TCP. Thus, when they coexist with TCP they are designed to yield giving full priority to the latter; even if starvation is precluded, the amount of bandwidth share relative to TCP is commonly unspecified and arbitrary. This unfairness between TCP and low priority transfers is manifested in weakened incentives for adoption of such new protocols instead of the incumbent TCP in public environments.

At the other end of the fairness spectrum, the authors in [6] give a "long term" twist on the notion of TCP fairness, tailored for background transfers, by requiring the average throughput obtained over a long period of time to match that of a TCP flow. In fact a congestion control algorithm for background transfers offering this type of fairness is introduced, where the delay of short flows is strictly decreased compared to the case where TCP is used instead [6], [7]. A side result of our analysis, suggests that the delay improvement brought by such flows is never more than 17.2%, and quickly deteriorates as the number of coexisting background TCP flows increases. Actually as shown in section II, this holds for any bandwidth sharing algorithm which is TCP fair in the long term sense. Thus, significant delay improvements are not possible unless the fairness requirement is weakened.

In this paper we identify a fundamental tradeoff between long term fairness between background flows when a subset of these uses TCP, and the impact on the delay of short TCP flows. We introduce a design methodology for end-to-end congestion control algorithms which explicitly accounts for this tradeoff. Similarly to the framework laid down by Kelly [8] we characterize fairness as the solution of an optimization problem involving utilities, but where a cost term is included accounting for the impact on delay of short flows. The optimization is performed not only over the long term average throughputs of background flows (including TCP ones), but also over all bandwidth sharing algorithms - operating on fast timescales- that attain these long term averages. Optimization over the latter is necessary in order to achieve the minimum possible delay costs. Using a linear programming formulation of a Markovian decision problem, we establish that the minimum delay is achieved by a threshold policy on the number of short flows which is not practical to implement and more significantly, it requires global information.

A basic result of this paper is that the more standard weighted TCP policy, which maintains a target proportion of TCP's bandwidth over a fast timescale, closely approximates the performance of the minimum delay policy *uniformly* well, and the relative error is quickly decreasing to zero as the number of coexisting

background TCP flows increases. More importantly, in contrast to the delay optimal algorithm, the weighted TCP policy allows for a distributed end-to-end implementation of a weight adjustment algorithm which approximately solves the overall optimization problem mentioned in the previous paragraph. Again, the relative optimality error decreases as the number of background TCP flows increases. Weighted TCP algorithms have been widely considered eversince they first appeared in [8], [9] and recently proposed for background data transfers in [10].

As an illustration of the methodology based on the optimization problem, in section IV-A we give two oppositely behaving algorithms which correspond to two different choices of the delay penalty term in the optimization problem. The first algorithm's aggressivity ranges between that of TCP and low priority, depending on whether the delay impact on short flows is low or high respectively. In contrast, the second algorithm which exhibits the opposite behavior, attempts at being a competitive alternative to TCP, except when excess capacity is abundant. Such a behavior may be favoured by users who are not willing to use a protocol that obtains a smaller share of the excess capacity than TCP, unless their throughput is already high.

In [6] the authors propose a utility-based framework concerning long term fairness, considering a model without short flow arrivals. In section II we introduce a system model with Poisson arrivals of short flows of exponentially distributed file sizes at a single bottleneck link, and establish that threshold policies minimize the delay of short flows. Bounds and formulas for the minimum delay are also given in the case where the offered load of short flows is high. In section III we assess the delay incurred by weighted TCP and compare it with the optimal. The general optimization problem ivolving the utility-based fairness criterion including delay costs is considered in section IV. The generic congestion control algorithm expounding our methodology is described in the same section, along with the two example algorithms mentioned above. In V further issues regarding the use of the methodology are discussed. All proofs are relegated to section VI.

II. DELAY OPTIMAL BANDWIDTH SHARING

Consider a link of capacity C shared by $k + l$ background flows of infinite duration, where k of these use TCP while the rest are referred to as FAIRBAT (fair background traffic) flows, abbreviated as FB. Here and in the next sections we seek to optimize the bandwidth sharing policies used by the l FB flows. Short TCP transfers of files with independent exponentially distributed file sizes (hence of finite duration), of mean μ^{-1}, arrive at the link according to a Poisson process with rate λ arrivals per unit time. Let x_n denote the download bandwidth of each TCP flow when the number of active short TCP flows is n. This number evolves according to a Markov chain with transition rates:

$$n \to \begin{cases} n+1 \, , & \text{with rate } \lambda \, , n \geq 0 \, , \\ n-1 \, , & \text{with rate } \mu n x_n \, , n \geq 1 \, . \end{cases} \quad (1)$$

The load brought in the system by short TCP flows is $C\rho$, where $\rho = \lambda/(\mu C)$. Clearly, if $\rho \geq 1$ the Markov chain is not positive recurrent regardless of the choice of x_n's; thus from now on $\rho < 1$ is assumed to hold. The amount of capacity $C(1-\rho)$ leftover by short TCP flows, is the *excess capacity* and is consumed in its entirety by the background flows.

Now, the choice of $(x_n, n = 0, 1, \ldots)$ determines how (actual, not excess) capacity is shared between (short or background) TCP and FB flows, since at state n the TCP flows use bandwidth $(k+n)x_n$ while the FB flows consume the remaining $C - (k + n)x_n$. The problem we solve in this section is the following: *what is the optimal sharing policy $(x_n, n = 0, 1, \ldots)$ such that the average delay experienced by short TCP flows is minized, under the constraint that the FB flows aggregately get a fraction f, of the excess capacity?* For example, $f = \frac{l}{k+l}$ would require each background flow to get an equal share of excess capacity regardless of its type, i.e., TCP or FB. Insofar as we only deal with the delay impact on short flows, we do not deal with how the f fraction is attributed between FB flows, until fairness is considered in section IV.

For any sharing policy which attains a fraction f for FB flows, let $N(k, f, \rho)$ be the average number of short flows under the stationary distribution of (1). For future reference, define the *delay impact* as

$$\delta(C(1-\rho)f) = \frac{N(k, f, \rho)}{C\rho} - \frac{1}{C(1-\rho)} \, , \quad (2)$$

where the last term is the delay of a unit size short flow if no background flows existed in the system[1], i.e., if $k = l = 0$. Thus, $\delta(C(1-\rho)f)$ measures the net impact of the background traffic on the average delay experienced by a short flow of unit size, if FB flows consume $C(1-\rho)f$ units of excess capacity. The minimization of $\delta(C(1-\rho)f)$ is equivalent to that of $N(k, f, \rho)$, so we arrive at the following problem:

$$N_*(k, f, \rho) = \min \sum_{n=0}^{\infty} n\pi_n \quad (3)$$

$$\text{such that: } \lambda\pi_{n-1} = \mu n x_n \pi_n \, , n = 1, 2, \ldots \quad (4)$$

$$\sum_{n=0}^{\infty} \pi_n = 1 \quad (5)$$

$$x_n \leq \frac{C}{k+n} \, , n = 0, 1, \ldots \quad (6)$$

$$\sum_{n=0}^{\infty} x_n \pi_n = \frac{C(1-\rho)(1-f)}{k} \quad (7)$$

$$\text{over } x_n \geq 0, \pi_n \geq 0 \, , n = 0, 1, \ldots \quad (8)$$

Equalities (4) are the local balance equations which the chain must obey, (6) is the link capacity constraint, and (7) is the constraint that the FB flows attain the target fraction f.

The following theorem states that the optimal policy has a structure of a threshold policy on the number of short flows.

Theorem 1 (Structure of the optimal policy). *The optimal sharing policy $(x_n, n = 0, 1, \ldots)$ satisfies:*
If $(1-\rho)^k \leq f$ then

$$x_n = \begin{cases} 0 \, , & \text{for each } n \leq n_* \, , \\ \frac{C}{k+n} & \text{for each } n \geq n_* + 2 \end{cases}$$

for some finite nonnegative integer n_*.
If $(1-\rho)^k > f$ we have

$$x_n = \begin{cases} \frac{C[(1-\rho)^k - f]}{k(1-\rho)^k} \, , & \text{for } n = 0 \, , \\ \frac{C}{k+n} \, , & \text{for each } n \geq 1 \, . \end{cases}$$

[1]In this case $x_n = C/n$ and the chain (1) is a M/M/1 queue.

In words, FB flows get the entire capacity at times where the number of flows is no more than n_*, while they get zero bandwidth in states greater than $n_* + 1$. When the number of short TCP flows is exactly $n_* + 1$, the FB bandwidth share will be just enough so that their average throughput satisfies (7), i.e.,

$$\pi_{n_*} C + \pi_{n_*+1} \left[C - (k + n_* + 1)x_{n_*+1} \right] = C(1-\rho)f \ .$$

If $(1-\rho)^k > f$, the FB flows get their target share while $n = 0$, so they do not need to compete with short TCP flows.

Interestingly, the optimal threshold n_* is determined by considering an associated blocking system, as described in the following proposition.

Proposition 1 (Determination of optimal threshold). *The optimal threshold n_* satisfies*

$$E\left(n_* + k, k, \frac{1-\rho}{\rho} \right) \le f < E\left(n_* + 1 + k, k, \frac{1-\rho}{\rho} \right) \ , \quad (9)$$

where

$$E(m, q, r) = \frac{\binom{m}{q} r^q}{\sum_{i=0}^{q} \binom{m}{i} r^i} \ , m \ge q \ ,$$

is the Engset formula of blocking probability for a loss system with q cicrcuits and m independent users, each offering traffic equal to r Erlangs.

Translating a threshold policy into an end-to-end congestion control algorithm for FB flows is a challenging task. A basic problem is that the FB flows would need to consume the entire link capacity when $n \le n_*$. During such periods, the packet loss rate at the link must considerably increase in order for TCP flows to decrease their congestion windows significantly. As a result, packet loss becomes a poor signal of subsequent upcrossings of n_*. Instead of threshold policies, we consider the simpler "weighted TCP" sharing policies described in the next section. Before turning on these, we evaluate the optimal delay performance in order to compare it with that under weighted TCP sharing.

The mimimum average number of short TCP flows $N_*(k, f, \rho)$ is obtained by invoking the following proposition, which holds for any (not necessarily optimal) threshold policy:

Theorem 2 (Performance of threshold policies). *Consider any threshold policy with*

$$x_n = \begin{cases} 0 \ , & \text{for each } n \le n_0 \ , \\ \frac{C}{k+n} & \text{for each } n > n_0 \end{cases}$$

for finite nonnegative integer n_0. The average number of short TCP flows under this policy at stationarity is

$$N_{n_0} = \frac{(k+1)\rho}{1-\rho} + n_0 E\left(n_0 + k, k, \frac{1-\rho}{\rho} \right) \ .$$

In particular, under the optimal policy, $N_{n_*} \le N_*(k, f, \rho) < N_{n_*+1}$.

As ρ approaches 1, the associated loss system in Proposition 1 is well approximated by a standard Erlang loss system. This simplifies the determination of both n_* and $N_*(k, f, \rho)$:

Corollary 1. *As $\rho \to 1$ the optimal threshold n_* satisfies $n_*(1-\rho) \to a_f$, where a_f is the unique solution of $B(k, a_f) = f$, and $B(k,a) = \frac{a^k}{k!} \left(\sum_{i=0}^{k} \frac{a^i}{i!} \right)^{-1}$ is the Erlang blocking formula for a system with k circuits under a load of a Erlangs.*
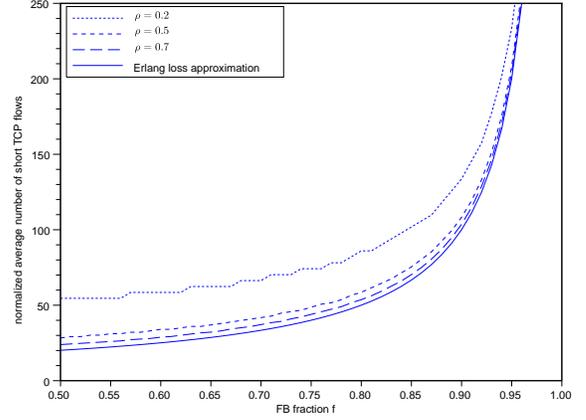

Fig. 1.   Average number of short TCP flows (normalized by $\rho/(1-\rho)$) under the optimal policy, when FB flows get a fraction f of the excess capacity and coexist with 10 background TCP flows.

Moreover, the average number of flows $N_*(k, f, \rho)$ satisfies $N_*(k, f, \rho)(1-\rho) \to k + 1 + a_f f$, as $\rho \to 1$.

In Fig. 1, $N_*(10, f, \rho)$ is plotted against the target fraction f under various load levels ρ, after being normalized by $\rho/(1-\rho)$ (the average number of short flows if background flows were absent). Notice the sharp increase to $+\infty$ as $f \to 1$. The solid curve is the approximation provided by the Erlang loss system in Corollary 1, which is fairly accurate for $\rho > 0.5$.

III. A WEIGHTED TCP SHARING POLICY

In this section we show that the optimal delay performance can be well approximated by a weighted TCP sharing policy. The relative error of the approximation vanishes as the number of backgound TCP flows k increases, *uniformly* over all fractions f. In particular, numerical evaluation shows that the loss is not more than 17.2% when $k = 1$ and quickly decreases in k.

Consider the same system as before but where FB flow i now takes, at all times, a fixed proportion w_i of TCP's instantaneous bandwidth. Thus, the l FB flows collectively behave as a set of $W = \sum_{i=1}^{l} w_i$ regular (i.e., weight 1) TCP flows, and $x_n = C/(k + W + n)$ at each state $n \ge 1$.

If the target fraction of FB flows using weighted TCP sharing is set to f, W must satisfy $W/(k+W) = f$. Thus, by Theorem 2 for $n_0 = 0$ and $k + W$ replacing k, the resulting average number of short TCP flows is

$$N_{\text{w}}(k, f, \rho) = \frac{(k + W + 1)\rho}{1-\rho} = \frac{\rho}{1-\rho} \left(k + 1 + \frac{kf}{1-f} \right) \ . \tag{10}$$

By the definition of optimality, $N_*(k, f, \rho) \le N_{\text{w}}(k, f, \rho)$ always holds. It turns out that $N_{\text{w}}(k, f, \rho)$ closely approximates $N_*(k, f, \rho)$ for large k, and ρ close to 1. In particular:

Theorem 3 (Optimal delay approximation error of weighted TCP sharing). *The following hold:*

1) *For every $0 \le f < 1$,*

$$\lim_{\rho \to 1} \frac{N_w(k, f, \rho) - N_*(k, f, \rho)}{N_w(k, f, \rho)} \le B(k-1, a_f) - f \ . \tag{11}$$
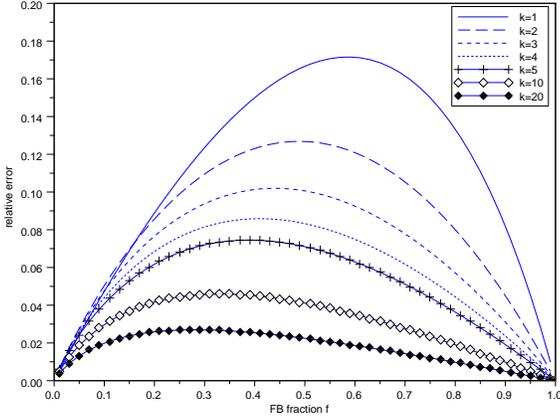
Fig. 2. The relative distance between the delay attained by the weighted TCP sharing policy and the optimal delay. The error decreases to 0 uniformly over all FB target fractions f, and the maximum value is approximately 17.2%, attained when $k = 1$.

2) *For each $k \geq 1$ let $b_k = \sup_{0 \leq f < 1} [B(k-1, a_f) - f]$, i.e., the worst upper bound* (11). *Then, $b_k \downarrow 0$ as $k \to \infty$.*

Although the upper bound b_k suffices to prove uniform convergence of the approximation error to 0, it is not tight. In fact, $b_1 = 1$ while the maximum approximation error for $k = 1$ is $3 - 2\sqrt{2} \approx 17.2\%$, attained at $f = 2 - \sqrt{2} \approx 60\%$. Fig. 2 depicts the relative error in the left hand side of (11), as a function of f. Note that the error is decreasing in k for most values of f. Thus, for practical purposes, weighted TCP sharing is a fairly good approximation to the optimal policy, even for small k.

The above bounds also apply to the performance of the "long term" TCP fair congestion control algorithms in [6], [7]. As shown in [6], these algorithms achieve higher throughput than TCP (over long periods of time) while at the same time they decrease the delay of short TCP flows. If the number of these "farsighted" flows is l, then they end up obtaining at least a $l/(k+l)$ fraction of excess capacity. Now the resulting delay cannot be lower than that achieved by the optimal sharing policy with $f = l/(k+l)$. In turn, the bounds above suggest that this optimal delay is within 17.2% of the delay achieved by weighted TCP sharing with $W = kf/(1-f) = l$, i.e., when these l flows use (unweighted) TCP. This implies that the delay decrease brought by farsighted flows cannot be more than 17.2%. Moreover, from Theorem 3, this improvement vanishes with higher k.

IV. FAIRNESS AND DELAY TRADEOFF

In the previous sections the bandwidth share f taken by FB flows was exogenous and fixed. The choice of f determines how fairly the excess capacity is distributed between background TCP and FB flows. However the choice should not be entirely based on bandwidth fairness criteria since as we saw in the previous sections, the delay experienced by short flows also depends on f (e.g., see Fig. 1). Thus, it may be desirable to explicitly account for delay in any fairness considerations. In the sequel we define fairness for the sharing of the excess capacity $C(1-\rho)$ using a utility based approach such as in Kelly [8] but also include the delay penalty to the short flows.

Let u_0 be the utility function associated with each one of the k (background) TCP flows and u_i associated with the i-th FB flow present in the system. For all $i = 0, \ldots, l$, u_i is differentiable, strictly concave, nonpositive[2] and $u_i'(0+) = +\infty$. The utilities model the preferences of background flows on *long term average* throughput, and should not be confused with bandwidth sharing of the actual capacity occuring on a faster timescale (see [8], [11]). The problem we consider is the following:

$$
\max \quad k u_0(y_0) + \sum_{i=1}^{l} u_i(y_i) - \int_0^{\sum_{i=1}^{l} y_i} \delta(z)^2 F(\delta(z)) dz
$$
$$
\text{such that } k y_0 + \sum_{i=1}^{l} y_i = C(1-\rho) , \tag{12}
$$
$$
\text{over } y_0, \ldots, y_l \geq 0 ,
$$

where y_i is the long term average throughput of flow $i = 1, \ldots, l$, y_0 is the corresponding throughput for each background TCP flow, and F is any nonnegative, nonincreasing function such that $\delta^2 F(\delta)$ is nondecreasing in $\delta > 0$. The last term in the objective function represents a measure of the total delay cost incurred to short flows by each unit of FB flow, and the reason it has this form is because it leads to simple to implement congestion control algorithms for FB flows which optimize (12) at equilibrium. Yet it is sufficiently flexible to cover cases of practical interest as seen in section IV-A.

Recall that the delay impact $\delta(z)$ in (2), is determined by the bandwidth sharing policy on fast timescales, which dictates how actual capacity is shared when n short flows are present, as in section II. Hence, in the case of delay optimal sharing where $\delta(z)$ is given by (2) for $N_*(k, z/(C(1-\rho)), \rho)$, (12) can be thought as optimizing the sharing of excess capacity, as specified by the utility functions, including a measure of the delay externalities to short flows *under the assumption that the delay optimal policy is employed by FB flows*. In fact, this policy yields the maximum possible optimal value of (12), since $\delta^2 F(\delta)$ is nondecreasing.

Remark. *The best possible optimal value of* (12), $V_*(k, \rho)$, *over all bandwidth sharing policies, is attained under the delay optimal bandwidth sharing policy.*

It is beneficial to consider the weighted TCP sharing policy instead. In this case the impact on delay is (see (10))

$$
\delta_{\mathrm{w}} \left(\sum_{i=1}^{l} y_i \right) = \frac{1}{C\rho} N_{\mathrm{w}} \left(k, \frac{\sum_{i=1}^{l} y_i}{C(1-\rho)}, \rho \right) - \frac{1}{C(1-\rho)} = \frac{1}{y_0} . \tag{13}
$$

When $\delta(z) = \delta_{\mathrm{w}}(z)$, problem (12) can be solved by a congestion control algorithm operating on two separate timescales. On the fast timescale, as stated, weighted TCP sharing is used with weight w_i for flow $i = 1, \ldots, l$. As a consequence, the long term throughputs $y_i, i = 0, \ldots, l$ with $y_i = w_i y_0$ and $k y_0 + \sum_{i=0}^{l} y_i = C(1-\rho)$ result, provided the weights w_i change on a slower timescale.

The use of weighted TCP sharing, instead of the optimal policy, yields a higher delay impact $\delta \left(\sum_{i=1}^{l} y_i \right)$ than necessary for (12) to attain $V_*(k, \rho)$. Nevertheless, it is a cost we opt to accept since a distributed gradient projection solution is then possibe.

---

[2]Since the argument of u_i is always upper bounded by $C(1-\rho)$ in (12), there is no loss of generality in assuming $u_i \leq 0$ since the amended utility functions $u_i(y) - u_i(C(1-\rho))$ generate the same optimal solutions in (12).

Specifically, the partial derivative of the objective function with respect to y_i is then,

$$-u_0'(y_0) + u_i'(y_i) - \frac{1}{y_0{}^2} F\left(\frac{1}{y_0}\right) , i = 1, \dots, l, \qquad (14)$$

where we used (13). Notice that (14) depends only on y_i, w_i, i.e., on local information, and requires knowledge of the functions u_0', F, u_i'. This is acceptable since in our approach, u_0', F, u_i' are parameters in the design space of the algorithm designer (see section V). Interestingly, the knowledge of the number k of background TCP flows is *not* needed.

On a slower timescale -slow enough for reliable estimates of y_i to be available- the weight parameteters w_i are locally adjusted such that the zeros of (14) are found. One way to do this is by the following algorithm:

$$\dot{w}_i = -u_0'\left(\frac{y_i}{w_i}\right) + u_i'(y_i) - \left(\frac{w_i}{y_i}\right)^2 F\left(\frac{w_i}{y_i}\right) , i = 1, \dots, l. \qquad (15)$$

The next proposition establishes that the value of the objective function in (12) evaluated at the equilibrium $w_1, \dots, w_l$ is close to the value $V_*(k, \rho)$ of (12) under the optimal sharing policy. In fact, their relative distance vanishes as the number k of background TCP flows increases, and ρ is high.

Theorem 4 (Approximation error of weighted TCP sharing). *Let $k \geq 1$. Then, there are unique equilibrium weights $w_1, \dots, w_l$ for algorithm* (15) *under the weighted TCP sharing policy. The resulting equilibrium throughputs $y_1^w, \dots, y_l^w$ are the unique maximizers of* (12) *with $\delta(z) = \delta_w(z)$.*

Let $V_w(k, \rho)$ be the corresponding optimal value. If $C(1-\rho)$ remains constant[3] in order for the excess capacity sharing problem to make sense as $\rho \to 1$, then for every $k \geq 1$,

$$\limsup_{\rho \to 1} \left| \frac{V_*(k, \rho) - V_w(k, \rho)}{V_*(k, \rho)} \right| \leq \frac{a_k}{1 - a_k},$$

where $a_k = \min(2b_k, 1)$ (see Theorem 3).

Therefore by Theorem 3, the weighted TCP sharing policy is asymptotically optimal as $k \to \infty$.

A. Two example algorithms

In this section we apply the above methodology and devise two algorithms which make sense in (different) practical situations.

Consider problem (12) with $u_i(y) = \log y - \log(C(1 - \rho))$ for every $y > 0, i = 0, \dots, l$. If delay was not important, i.e., $F(\delta) = 0$, the optimal allocation would be the one where each background flow gets an equal share of excess capacity. When $F(\delta) > 0$, (12) yields an allocation which deviates from equal sharing only as long as this brings a sufficient delay decrease.

Now, assume a delay penalty with $F(\delta) = \gamma$ for every $\delta > 0$, where $\gamma > 0$ is a constant. Setting the partial derivatives (14) equal to zero yields $y_0/y_i = 1 + \gamma y_0^{-1}$ for all $i = 1, \dots, l$. That is, the resulting type of FB flow is always less aggressive than TCP. Notice also that the delay impact on short flows is y_0^{-1}. Recalling that the delay impact on short flows is y_0^{-1}, notice that the larger this impact is, the less aggressive the FB flows become relative to TCP, thus effectively behaving as low priority traffic. On the other hand, the smaller y_0^{-1} gets, the more similar to TCP the FB flows behave. Indeed, there is little reason in treating

[3]More generally, we could assume $C(1 - \rho)$ converges to a constant as $\rho \to 1$, at the expense of having to use more symbols.
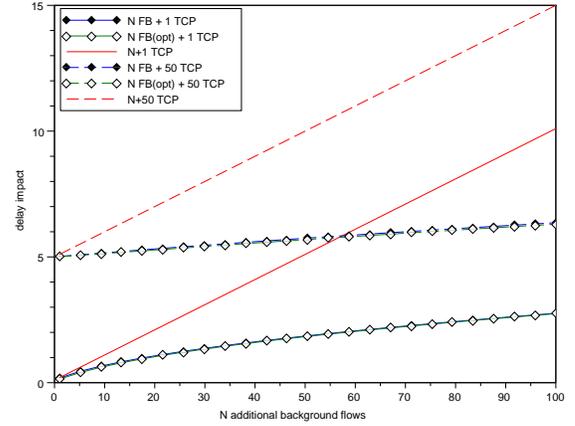

Fig. 3. Impact of the number and type of background flows on the delay of short TCP flows when N background flows are added. (FB flows use the first algorithm of section IV-A, FB(opt) use the optimal solution of (12) with $F(\delta) = \gamma$.)

background data as low priority unless their impact on delay is significant.

In Fig. 3 the delay impact δ is depicted, as the result of adding N background flows, in a numerical example with $C(1 - \rho) = 10, \gamma = 1$. TCP flows cause a linear delay increase, while in the case of FB the delay increases more slowly. FB(opt) flows correspond to the optimal solution of (12) under the optimal sharing policy of section II. Notice that the delay impact of FB essentially coincides with the optimal.

In Fig. 4 the equilibrium bandwidth share of each FB flow relative to that of a background TCP flow is depicted. When the number of background flows is low, the delay penalty is small and so the equilibrium shares of FB and TCP flows are nearly equal, i.e., fair. As the number of background flows increases and the delay penalty becomes more dominant, fairness is traded for lower delay cost. Notice that the optimal algorithm yields slightly more fair shares with the same delays, as seen in Fig. 3, when a single background TCP exists in the system.

To obtain the second algorithm, consider $F(\delta) = \gamma \delta^{-2}$. Setting (14) equal to zero yields $y_i = y_0/(1 + \gamma y_0)$ for all $i = 1, \dots, l$ where again, the FB flows are always less aggressive than TCP but, in contrast to the first algorithm, they behave as low priority traffic only when y_0 is high. When y_0 is small, then $y_i \approx y_0$. Such a behavior may be favoured by users who are not willing to use a protocol that obtains a smaller share of the excess capacity than TCP, unless their throughput is already high.

In Fig. 5 the delay impact of the second algorithm is depicted, for the same parameters $C(1-\rho), \gamma$ as before. For a small number of background flows, the impact is minimal since FB flows receive a sufficiently high throughput. As more background flows join the system and the excess capacity is divided among a larger number of flows, the FB flows become more aggressive as seen in Fig. 6. As a consequence, the delay starts to increase in a linear fashion, similarly to the case where all flows use TCP. Notice that in this algorithm, the delay penalty term in (12) is the same for both FB and FB(opt). Thus, the equilibrium shares of FB and FB(opt), relative to TCP, coincide. By optimality, FB(opt) yields lower delays than FB, but as seen in Fig. 5 the error between the two is insignificant. Specifically, the error is much smaller than the
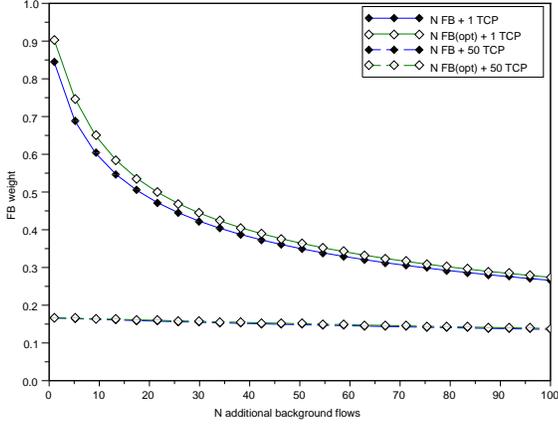
Fig. 4. Equilibrium bandwidth share of an FB (or FB(opt)) flow relative to that of a background TCP flow, i.e., the y_i/y_0 ratio. For small N, the delay penalty is low and so the equilibrium shares of FB and TCP flows are nearly equal, i.e., fair. (FB flows use the first algorithm of section IV-A, FB(opt) use the optimal solution of (12) with $F(\delta) = \gamma$.)

Fig. 6. Equilibrium bandwidth share of an FB (or FB(opt)) flow relative to that of a background TCP flow, i.e., the y_i/y_0 ratio. (FB flows use the second algorithm of section IV-A, FB(opt) use the optimal solution of (12) with $F(\delta) = \gamma\delta^{-2}$.)
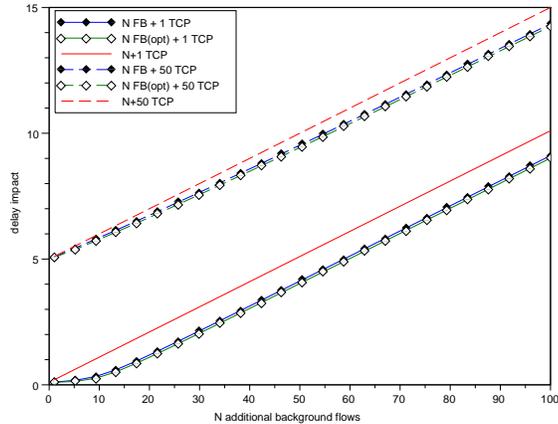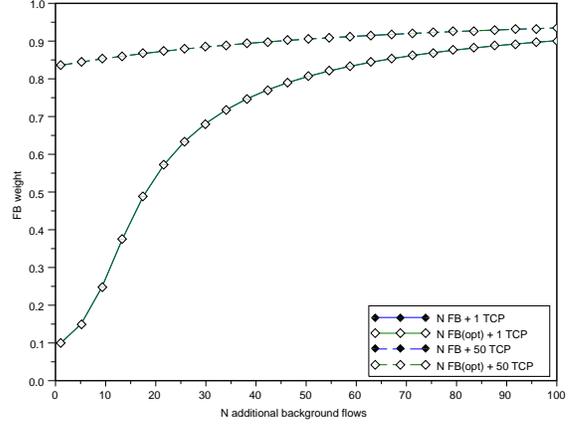

Fig. 5. Impact of the number and type of background flows on the delay of short TCP flows. (FB flows use the second algorithm of section IV-A, FB(opt) use the optimal solution of (12) with $F(\delta) = \gamma\delta^{-2}$.)

17.2% upper bound found numerically in section III.

V. DISCUSSION

The design methodology proposed in this paper is based on an explicit specification of the tradeoff between long term fairness and the delay impact to short flows, such as the one given by the optimization problem (12). Presumably, the algorithm designer decides on the fair allocation of excess capacity between FB and background TCP flows, in the absence of delay costs, by first choosing u_0, u_i. This is what we did for the two algorithms in section IV-A, where fairness was defined as equal sharing of the excess capacity. This led us to choose $u_i = u_0$, while the choice of a logarithmic form was made to yield simple expressions for $(u'_i)^{-1}$. The optimality conditions (obtained by equating (14) with zero) give a relation between y_0, y_i and $\delta^2 F(\delta)$, where δ is the delay impact. Now, the designer chooses F in order for y_i, the TCP share y_0 and the delay impact satisfy a desired relation.

Notice that fairness is defined with respect to TCP share y_0 (through u'_0), and so the latter becomes a kind of common currency, generalizing the use of TCP friendliness. This common currency makes possible the coexistence of FB flows with different utilities u_i and the optimality of the distributed algorithm (15) under weighted TCP sharing. Furthermore, under weighted TCP, the delay impact is given solely in terms of y_0, a value which is known to all FB flows through $y_0 = y_i/w_i$.

Whereas in the utility-based framework of Kelly [8], [11] the shadow price of the resource constraint is signalled to the users residing at the endpoints, through congestion indications coming from the network (e.g., packet loss), in our case the function F must be known (in order for the marginal cost $\delta^2 F(\delta)$ to be computable) by each FB flow, since the delay cost is a choice of the algorithm designer. It is interesting to note that in certain cases, the information about the delay cost can be implicitly signalled by the network. As an example, assume each packet loss incurs a cost $\beta > 0$ to each FB flow, and let p_n be the packet loss probability when n short flows are present. By the TCP square root formula [12], $x_n = \frac{1}{T}\sqrt{\frac{2\pi}{p_n}}$ where T is the round-trip time, and $x_n = C/(k + W + n)$ is the TCP share at state n. Thus, the average[4] loss rate for FB flow i is

$$w_i E(p_n x_n) = \frac{w_i 2\pi}{T^2}\left(\frac{k+W}{C} + \frac{(k+W+1)\rho}{C(1-\rho)}\right) \approx \frac{2\pi w_i}{T^2 y_0},$$

for $C(1-\rho)$ posessing a positive limit as $\rho \uparrow 1$. Since the throughput of flow i is $w_i y_0$, the average price per unit of throughput is $\beta 2\pi/(T^2 y_0{}^2)$. Now if FB flows adjust their weight according to $\dot{w}_i = u'_i(y_i) - \beta 2\pi/(T^2 y_0^2)$ such that they attempt to maximize their surplus, are in essense solving (12) for e.g., $u_0(y) = -\beta\pi/(T^2 y)$ and $F(\delta) = \beta\pi/T^2$ for all δ.

In the future we plan to explore more the relation between the average congestion indications obtained in the fast timescales, and the implicit delay cost functional.

---

[4]The expectation is with respect to the stationary distribution π_n arising from the choice of the x_n's, in (1).

VI. Proofs

A. Proof of Theorem 1

Consider the case $(1-\rho)^k > f$ first. If $x_n, n = 0, 1, \ldots$ are defined as in Theorem 1, FB flows do not affect the dynamics of the Markov chain. Hence, the stationary distribution is given by Lemma 4 for $n_0 = 0$. In fact, the average bandwidth constraint (7) is readily shown to hold for this stationary distribution. Since the average number of TCP flows is minimized by setting x_n at their maximum value $\frac{C}{k+n}$ for each $n = 1, \ldots$, we conclude that the allocations $x_n, n = 0, 1, \ldots$ are optimal.

For the remainder of this section we consider the case $(1-\rho)^k \le f$.

Lemma 1. *If* $(1-\rho)^k \le f$ *then* $x_0 = 0$.

Proof: Assume first that $x_n = \frac{C}{k+n}$ for all $n \ge 1$. Then the stationary distribution is given by Lemma 4 for $n_0 = 0$, and

$$\pi_0 x_0 + \sum_{n=1}^{\infty} \pi_n x_n = (1-\rho)^{k+1} x_0 + \frac{C(1-\rho) - C(1-\rho)^{k+1}}{k} \,.$$

Plugging this into (7) yields $f = (1-\rho)^k (1 - kx_0/C)$ which it does not hold unless $x_0 = 0$.

Now assume that there exists $n \ge 1$ for which the constraint (6) is not met. If $x_0 > 0$ then we could decrease x_0 and increase x_n such that (7) remains true. Notice though, that the increase of x_n does decrease the average number of short TCP flows, while the increase of x_0 does not have any effect whatsoever. Thus, the optimal allocation x_0 must be zero. ∎

We now transform the optimization problem into an equivalent convex problem. For each $n \ge 0$ and x_n, π_n which satisfy (4)-(8), define

$$\bar{\pi}_n = x_n \pi_n \frac{k}{C(1-\rho)(1-f)} \,. \tag{16}$$

The following holds:

Lemma 2. *The numbers* $\bar{\pi}_n, n = 0, 1, \ldots$ *as defined in* (16) *satisfy the constraints*

$$\sum_{n=0}^{\infty} \bar{\pi}_n = 1 \,, \tag{17}$$

$$\rho \bar{\pi}_n \le \frac{n+1}{k+n} \bar{\pi}_{n+1} \,, n = 0, 1, \ldots \,, \tag{18}$$

$$\sum_{n=0}^{\infty} n \bar{\pi}_n = \frac{\rho k}{(1-\rho)(1-f)} \,, \tag{19}$$

$$\bar{\pi}_n \ge 0 \,, n = 0, 1, \ldots \tag{20}$$

Conversely, for any $\bar{\pi}_n, n = 0, 1, \ldots$ *satisfying* (17)-(20) *there exist unique* $\pi_n, x_n, n = 0, 1, \ldots$ *for which* (4)-(8) *and* (16) *hold.*

Proof: Summing (16) over $n = 0, 1, \ldots$ and using (7) yields (17). (18) follows by mutliplying both sides of (6) by π_n and utilizing (4),(16). Since $\rho < 1$, the average rate of departures must equal λ, that is,

$$\sum_{n=0}^{\infty} \mu n x_n \pi_n = \lambda \Leftrightarrow \sum_{n=0}^{\infty} n \bar{\pi}_n = \frac{\rho k}{(1-\rho)(1-f)} \,,$$

which proves (19).

To show the converse part, define

$$\pi_n = \frac{(n+1)(1-\rho)(1-f)}{\rho k} \bar{\pi}_{n+1}$$

and $x_n = \begin{cases} \frac{C\rho \bar{\pi}_n}{(n+1)\bar{\pi}_{n+1}} \,, & \text{if } \bar{\pi}_{n+1} > 0 \,, \\ 0 \,, & \text{if } \bar{\pi}_{n+1} = 0 \,, \end{cases}$ for each $n = 0, 1, \ldots$

Now, (4) follows by noting that $\frac{n\bar{\pi}_n}{(n+1)\bar{\pi}_{n+1}} = \frac{\pi_{n-1}}{\pi_n}$ and substitution in the definition of x_n above. Also,

$$\sum_{n=0}^{\infty} \pi_n = \sum_{n=0}^{\infty} \frac{(n+1)(1-\rho)(1-f)}{\rho k} \bar{\pi}_{n+1}$$
$$= \sum_{n=0}^{\infty} \frac{n(1-\rho)(1-f)}{\rho k} \bar{\pi}_n = 1 \,,$$

by making use of (17). (6) follows by (18) and the definition of x_n above, while (7) follows by replacing π_n, x_n in (7).

This deals with existence; to establish uniqueness note that any collection of $\pi_n, x_n, n = 0, 1, \ldots$ which satisfy (16) and (4) defines π_n uniquely by the latter equation. Thus, x_n is defined uniquely by (16) for every $n \ge 1$, while x_0 is determined by (7). ∎

Now, if we multiply both sides of (4) by $n - 1$ and sum over $n = 1, 2, \ldots$ we get

$$\lambda \sum_{n=1}^{\infty} (n-1)\pi_{n-1} = \frac{\mu C(1-\rho)(1-f)}{k} \left(\sum_{n=1}^{\infty} n^2 \bar{\pi}_n - \sum_{n=1}^{\infty} n \bar{\pi}_n \right)$$
$$= \frac{\mu C(1-\rho)(1-f)}{k} \sum_{n=1}^{\infty} n^2 \bar{\pi}_n - \lambda \,,$$

when (19) holds. Thus, the minimization of the objective in (3) is equivalent to that of the second moment of the distribution $\bar{\pi}_n, n = 0, 1, \ldots$, when (19) holds. Combining this with Lemma 2 leads to the following equivalent formulation:

$$\text{Minimize } \sum_{n=0}^{\infty} n^2 \bar{\pi}_n \text{ over } (\bar{\pi}_n, n = 0, 1, \ldots)$$

$$\text{such that (17)-(20) hold.} \tag{21}$$

The Lagrangian for this problem is

$$\sum_{n=0}^{\infty} n^2 \bar{\pi}_n + \xi \left(\sum_{n=0}^{\infty} \bar{\pi}_n - 1 \right) + \psi \left(\sum_{n=0}^{\infty} n \bar{\pi}_n - \frac{\rho k}{(1-\rho)(1-f)} \right)$$
$$+ \sum_{n=0}^{\infty} \zeta_n \left(\rho \bar{\pi}_n - \frac{n+1}{k+n} \bar{\pi}_{n+1} \right)$$
$$= \sum_{n=0}^{\infty} \left(n^2 + \xi + \psi n + \rho \zeta_n - \frac{n}{k+n-1} \zeta_{n-1} \right) \bar{\pi}_n$$
$$- \xi - \psi \frac{\rho k}{(1-\rho)(1-f)} \,, \tag{22}$$

where $\psi, \xi \in \mathbb{R}, \zeta_{-1} = 0, \zeta_0, \zeta_1, \ldots \in \mathbb{R}_+$ are the Lagrange multipliers.

By Theorem 1 of section 8.3 in [13], there exist values for these multipliers, for which the infimum of (22) over $\bar{\pi}_n \ge 0, n = 0, 1, \ldots$ is attained, and coincides with the minimizer of (21). By abusing notation, denote the corresponding values again by ψ, ξ, ζ_n, and let $\bar{\pi}_n, n = 0, 1, \ldots$ be the minimizer of (21). Also, note by complementary slackness that

$$\zeta_n \left(\rho \bar{\pi}_n - \frac{n+1}{k+n} \bar{\pi}_{n+1} \right) = 0 \,, \quad \text{for each } n = 0, 1, \ldots \tag{23}$$

Define $n_* = \max\{m \geq 0 : \bar{\pi}_n = 0 \,, \text{ for all } n \leq m\}$, which is well-defined since $\bar{\pi}_n = 0$ by Lemma 1. For such n_* the following holds:

Lemma 3. $\zeta_n > 0$ for all n with $n \geq n_* + 2$.

Proof: Since the infimum of (22) is attained, we have

$$\rho\zeta_n \geq \frac{n}{k+n-1}\zeta_{n-1} - n^2 - \psi n - \xi \,, \quad \text{for each } n = 0, 1, \ldots, \tag{24}$$

where the inequality becomes equality when $\bar{\pi}_n > 0$.

By the definition of n_* we have $\bar{\pi}_{n_*+1} > 0$, so (18) implies $\bar{\pi}_n > 0$ for all $n \geq n_* + 1$. Thus (24) holds with equality for all $n \geq n_* + 1$. Moreover, $\rho\bar{\pi}_{n_*} < \frac{n+1}{k+n}\bar{\pi}_{n_*+1}$, so $\zeta_{n_*} = 0$. Thus,

$$-(n_* + 1)^2 - \psi(n_* + 1) - \xi = \rho\zeta_{n_*+1} \geq 0$$
$$\text{and} \quad -n_*^2 - \psi n_* - \xi \leq -\frac{n_*}{k+n_*-1}\zeta_{n_*-1} \leq 0 \,.$$

Therefore the smallest root of the quadratic $x \mapsto -x^2 - \psi x - \xi$ lies in $[n_*, n_* + 1]$, while the other is greater than or equal to $n_* + 1$. Let $n \geq n_* + 2$ and assume $-n^2 - \psi n - \xi \leq 0$. Then $-(n+1)^2 - \psi(n+1) - \xi < 0$, and (24) gives

$$0 > -(n+1)^2 - \psi(n+1) - \xi \geq \frac{n+1}{k+n}\zeta_n \,,$$

i.e., $\zeta_n > 0$.

If on the other hand, $-n^2 - \psi n - \xi > 0$, then (24) implies $\zeta_n > 0$ again. ∎

By the lemma above, the complementary slackness condition (23) implies $\rho\bar{\pi}_n = \frac{n+1}{k+n}\bar{\pi}_{n+1}$ for all $n \geq n_0 + 2$, which in turn gives $x_n = C/(k+n)$.

B. Proof of Proposition 1

Lemma 4. *The stationary distribution of the number of short TCP flows is*

$$\pi_n = \begin{cases} \dfrac{\binom{n+k}{k}(1-\rho)^{k+1}\rho^n}{\sum_{i=0}^{k}\binom{n_0+k}{i}(1-\rho)^i\rho^{n_0+k-i}} \,, & n \geq n_0 \,, \\ 0 \,, & n < n_0 \,. \end{cases}$$

for the threshold policy with threshold n_0. In particular, $\pi_n = P(X = n | X \geq n_0)$, where X is the sum of $k + 1$ independent geometric random variables with parameter $1 - \rho$.

Proof: Let X be the sum of $k + 1$ independent geometric distributions with parameter $1 - \rho$. Then,

$$P(X = n) = \binom{n+k}{k}(1-\rho)^{k+1}\rho^n \,,$$

for every $n = 0, 1, \ldots$. This distribution is readily shown to satisfy the detailed balance equations

$$P(X = n)\rho = P(X = n+1)\frac{n+1}{n+1+k} \,, n = 0, 1, \ldots,$$

which correspond to a system with the threshold set to zero. Since the Markov chain is reversible, the stationary distribution for the case $n_0 > 0$ is $\pi_n = P(X = n | X \geq n_0)$, i.e.,

$$\pi_n = \frac{\binom{n+k}{k}(1-\rho)^{k+1}\rho^n}{\sum_{i=0}^{k}\binom{n_0+k}{i}(1-\rho)^i\rho^{n_0+k-i}} \,, n = n_0, n_0+1, \ldots \tag{25}$$

and 0 for $n < n_0$. ∎

Consider the threshold policies with threshold value n_0 but where x_{n_0+1} is allowed to vary within $\left(0, \frac{C}{n_0+1+k}\right]$. We obtain an explicit expression for the throughput

$$C\pi_{n_0} + \pi_{n_0+1}[C - (n_0 + k + 1)x_{n_0+1}] \tag{26}$$

of FB flows as a function of x_{n_0+1}. Applying Lemma 4 to the system with threshold $n_0 + 1$ yields

$$\frac{\pi_{n_0+1}}{\sum_{i=n_0+1}^{\infty}\pi_i} = (1-\rho)E\left(n_0 + k + 1, k, \frac{1-\rho}{\rho}\right) \,, \tag{27}$$

i.e., $\pi_{n_0+1} = (1 - \pi_{n_0})d$ where d is the expression in the right hand side of (27) This together with local balance between states n_0 and $n_0 + 1$ give,

$$\pi_{n_0} = \frac{\lambda d}{(n_0+1)x_{n_0+1}\mu d + \lambda} \,,$$
$$\text{and} \quad \pi_{n_0+1} = \frac{(n_0+1)x_{n_0+1}\mu d}{(n_0+1)x_{n_0+1}\mu d + \lambda} \,, \tag{28}$$

and substitution in (26) yields

$$d + \frac{d\mu(n_0+1)x_{n_0+1}}{d\mu(n_0+1)x_{n_0+1} + \lambda}\left(1 - \rho - d - \frac{k\rho}{n_0+1}\right) \,. \tag{29}$$

For $x_{n_0+1} = 0$, this expression is greater than $(1 - \rho)E\left(n_0 + k, k, \frac{1-\rho}{\rho}\right)$, i.e., the value obtained at $x_{n_0} = C/(n_0 + 1 + k)$, by Lemma 4. Thus (29) is a decreasing function of x_{n_0+1}, so there exists unique threshold n_* which satisfies (9). ∎

C. Proof of Theorem 2

The average number of short TCP flows at stationarity is, by Lemma 4,

$$E(X | X \geq n_0) = n_0 - 1 + \sum_{n=n_0}^{\infty} P(X \geq n | X \geq n_0) \tag{30}$$

$$= n_0 - 1 + \frac{\sum_{i=0}^{k} G_i}{P(X \geq n_0)} \,, \tag{31}$$

where $G_i = \sum_{n=n_0}^{\infty}\binom{n+k}{i}(1-\rho)^i\rho^{n+k-i}$.

Now,

$$G_i = \sum_{n=n_0}^{\infty}\binom{n+k}{i}(1-\rho)^i(-\rho^{n+1+k-i} + \rho^{n+k-i})$$
$$+ \sum_{n=n_0}^{\infty}\binom{n+k}{i}(1-\rho)^i\rho^{n+1+k-i}$$
$$= \binom{n_0+k}{i}(1-\rho)^i\rho^{n_0+k-i}$$
$$+ \sum_{n=n_0}^{\infty}\left[\binom{n+k+1}{i} - \binom{n+k}{i}\right](1-\rho)^i\rho^{n+1+k-i}$$
$$+ \sum_{n=n_0}^{\infty}\binom{n+k}{i}(1-\rho)^i\rho^{n+1+k-i}$$
$$= \binom{n_0+k}{i}(1-\rho)^i\rho^{n_0+k-i} + (1-\rho)G_{i-1} + \rho G_i \,,$$

where summation by parts is used to get the second equality. Thus,

$$G_i = \frac{1}{1-\rho}\binom{n_0+k}{i}(1-\rho)^i\rho^{n_0+k-i} + G_{i-1}$$
$$= \frac{1}{1-\rho}\sum_{j=0}^{i}\binom{n_0+k}{j}(1-\rho)^j\rho^{n_0+k-j} \,,$$

and so,

$$\sum_{i=0}^{k} G_i = \sum_{i=0}^{k} \frac{k+1-i}{1-\rho}\binom{n_0+k}{i}(1-\rho)^i \rho^{n_0+k-i} .$$

Plugging this back to (30) gives

$$E(X|X \geq n_0) = n_0 - 1 + \frac{k+1}{1-\rho}$$

$$- \frac{(n_0+k)\sum_{i=0}^{k-1}\binom{n_0+k-1}{i}(1-\rho)^i \rho^{n_0+k-1-i}}{\sum_{i=0}^{k}\binom{n_0+k}{i}(1-\rho)^i \rho^{n_0+k-i}}$$

$$= \frac{(k+1)\rho}{1-\rho} + n_0 E\left(n_0+k, k, \frac{1-\rho}{\rho}\right) ,$$

after some algebra.

D. Proof of Corollary 1

Consider a sequence of threshold policies indexed by ρ for which the threshold level n_ρ satisfies $n_\rho(1-\rho) \to a$ as $\rho \to 1$. Since the number of users $n_\rho + k$, in the loss system described in Proposition 1, grows with ρ while the total load is kept fixed at a, the call arrival sequence is approximated by a Poisson process with rate a. Hence,

$$E\left(n_\rho+k, k, \frac{1-\rho}{\rho}\right) \to B(k, a) , \text{ as } \rho \to 1 ,$$

Observe that since $B(k, 0) = 0, B(k, +\infty) = 1$, and $B(k, a)$ is increasing in a, there is a unique a_f for which $B(k, a_f) = f$ holds. Thus, $n_*(1-\rho) \to a_f$ as $\rho \to 1$, and both the lower and upper bounds in Proposition 1 converge to $B(k, a_f)$.

E. Proof of Theorem 3

$$\lim_{\rho \to 1} \frac{N_w(k, f, \rho) - N_*(k, f, \rho)}{N_w(k, f, \rho)} \leq \frac{\frac{kf}{1-f} - a_f f}{k + \frac{kf}{1-f}}$$

$$= \frac{1-f}{k} a_f \left(B(k-1, a_f) - f\right) \leq B(k-1, a_f) - f .$$

by using the identity $kB(k, a_f)/(1 - B(k, a_f)) = a_f B(k-1, a_f)$ and the definition $B(k, a_f) = f$. The last step follows by noting that the average number of ongoing calls $a_f(1 - B(k, a_f))$, in the associated loss system, is less than the number of circuits k.

To show the second part, first notice that $\sup_{0 \leq f \leq 1}[B(k-1, a_f) - f] = \sup_{a \geq 0}[B(k-1, a) - B(k, a)]$, and for every $a \geq 0$, $F(k, a) = B(k-1, a) - B(k, a) \to 0$ as $k \uparrow \infty$. Moreover, since $B(k, a)$ concave in k [14], $F(k, a)$ is nonincreasing in k, and so $F(k, a) \downarrow 0$ as $k \uparrow \infty$. Thus the convergence is uniform over intervals of the form $[0, a_0]$. It is uniform over the entire positive axis if $\sup_{k \geq k_0, a \geq a_0} F(k, a)$ can be made arbitrarily small by some choice of k_0, a_0. But this follows since for every sequence $a_k \uparrow +\infty$ we have $F(k, a_k) \leq F(k_0, a_k)$ for any $k \geq k_0$, and $F(k_0, a_k) \to 0$ as $k \uparrow \infty$.

F. Proof of Theorem 4

Since the objective function is strictly concave and the parameters range over a compact set, there exists unique maximizer $y^w = (y_0^w, \ldots, y_l^w)$ and the optimality conditions obtained by equating (14) with zero, yield y^w as their unique solution. Since $u_i'(0+) = +\infty$ and $k \geq 1$, we have $y_i^w > 0$ for each $i = 0, \ldots, l$.

Hence, there are unique weights $w_i, i = 1, \ldots, l$ with $y_i^w = w_i y_0^w$ which are also the unique weights satisfying $\dot{w}_i = 0$ in (15).

Now, for every $y = (y_0, \ldots, y_l) \in \mathbb{R}_+^{l+1}$ let $\mathcal{U}(y) - \mathcal{D}(y)$ be the decomposition of the objective in (12) into the utility term $\mathcal{U}(y)$ and the delay penalty term $\mathcal{D}(y)$. For δ_* given by (2) with $N_*(k, f, \rho)$, let $\mathcal{D}_*(y), \mathcal{D}_w(y)$ be the resulting delay penalties arising from δ_* and δ_w, respectively. First, notice that for any $0 \leq z < C(1-\rho)$ we have

$$\delta_w(z)^2 F(\delta_w(z)) - \delta_*(z)^2 F(\delta_*(z)) \leq (\delta_w(z)^2 - \delta_*(z)^2)F(\delta_w(z))$$

$$\leq (\delta_w(z) - \delta_*(z))2\delta_w(z)F(\delta_w(z))$$

$$\leq 2b_k \delta_w(z)^2 F(\delta_w(z)) ,$$

where we used the monotonicity of F, and (11) in the final step. Thus, $\mathcal{D}_w(y) - \mathcal{D}_*(y) \leq 2b_k \mathcal{D}_w(y)$, where $a_k = \min(2b_k, 1)$ for any $k \geq 1$. Now let $y^* \in \mathbb{R}_+^{l+1}$ be the maximizer of (12) under the optimal bandwidth sharing policy and notice

$$V_*(k, \rho) = \mathcal{U}(y^*) - \mathcal{D}_*(y^*)$$

$$\leq a_k \mathcal{U}(y^*) + (1 - a_k)\left[\mathcal{U}(y^*) - \mathcal{D}_w(y^*)\right]$$

$$\leq a_k \mathcal{U}(y^*) + (1 - a_k)V_w(k, \rho) ,$$

where we have used (11), and so $(1 - a_k)\left[V_*(k, \rho) - V_w(k, \rho)\right] \leq a_k \mathcal{D}_*(y^*)$, i.e.,

$$\left|\frac{V_*(k, \rho) - V_w(k, \rho)}{V_*(k, \rho)}\right| \leq \frac{a_k}{1-a_k}\frac{\mathcal{D}_*(y^*)}{|V_*(k, \rho)|} \leq \frac{a_k}{1-a_k} . \quad (32)$$

The last inequality follows because $\mathcal{U}(y^*) \leq 0$.

REFERENCES

[1] S. Floyd, M. Handley, J. Padhye, and J. Widmer, "Equation-based congestion control for unicast applications," in *ACM SIGCOMM Computer Communication Review*, vol. 4, no. 30, October 2000, pp. 43–56.

[2] A. Kuzmanovic and E. W. Knightly, "Tcp-lp: low-priority service via end-point congestion control," *IEEE/ACM Trans. Netw.*, vol. 14, pp. 739–752, August 2006. [Online]. Available: http://dx.doi.org/10.1109/TNET.2006.879702

[3] R. Venkataramani, R. Kokku, and M. Dahlin, "Tcp nice: A mechanism for background transfers," in *5th Symposium on Operating Systems Design and Implementation (OSDI 2002)*, Boston, USA, December 2002.

[4] A. Norberg, *uTorrent transport protocol*, 2009, draft.

[5] S. Shalunov, *Low Extra Delay Background Transport (LEDBAT)*, 2009, draft-shalunov-ledbat-congestion-00.

[6] P. Key, L. Massoulie, and M. Vojnovic, "Farsighted users harness network time-diversity," in *INFOCOM 2005. 24th Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings IEEE*, vol. 4, 2005, pp. 2383–2394.

[7] L. Shai, M. Vojnovic, and D. Gunawardena, "Competitive and considerate congestion control for bulk data," in *15th IEEE International Workshop on Quality of Service*, June 2007, pp. 1–9.

[8] F. P. Kelly, "Charging and rate control for elastic traffic," *European Transactions on Telecommunications*, vol. 8, pp. 33–37, 1997.

[9] R. Gibbens and F. Kelly, "Resource pricing and evolution of congestion control," *Automatica*, vol. 35, pp. 1969–1985, 1999.

[10] B. Briscoe, "A fairer, faster internet protocol," *IEEE Spectrum*, vol. Dec 2008, pp. 38–43, Dec. 2008. [Online]. Available: http://spectrum.ieee.org/telecom/standards/a-fairer-faster-internet-protocol

[11] F. Kelly, A. Maulloo, and Tan, "Rate control in communication networks: shadow prices, proportional fairness and stability," *Journal of the Operational Research Society*, vol. 49, pp. 237–252, 1998.

[12] M. Matthis, M. Semke, J. Mahdavi, and T. Ott, "The macroscopic behavior of tcp congestion avoidance algorithm," *Computer Communication Review*, vol. 27, no. 3, 1997.

[13] D. G. Luenberger, *Optimization by Vector Space Methods*, 1st ed. New York, NY, USA: John Wiley & Sons, Inc., 1997.

[14] W. Karush, "A queuing model for an inventory problem," *Operations Research*, vol. 5, no. 5, pp. pp. 693–703, 1957. [Online]. Available: http://www.jstor.org/stable/167469