Revenue-Maximizing Pricing and Capacity Expansion in a Many-Users Regime

Tamer Başar and R. Srikant

Abstract—

In this paper, we consider a network where each user is charged a fixed price per unit of bandwidth used, but where there is no congestion-dependent pricing. However, the transmission rate of each user is assumed to be a function of network congestion (like TCP), and the price per unit bandwidth. We are interested in answering the following question: how should the network choose the price to maximize its overall revenue? To obtain a tractable solution, we consider a single link accessed by many users where the capacity is increased in proportion to the number of users. We show the following result: as the number of users increases, the optimal price-per-unit-bandwidth charged by the service provider may increase or decrease depending upon the bandwidth of the link. However, for all values of the link capacity, the service provider's revenue-per-unit-bandwidth increases and the overall performance of each user (measured in terms of a function of its throughput, the network congestion and the cost incurred by the user for bandwidth usage) improves. Since the revenue per unit bandwidth increases, it provides an incentive for the service provider to increase the available bandwidth in proportion to the number of users.

Keywords—**Pricing, Capacity expansion, Congestion control, Many-users limit, Quality-of-Service**

I. Introduction

Recently, there has been much interest in the design of networks where very small queues are maintained at the routers (for example, see [10], [7], [17], [8], [13], [9], [15], [14]). This is accomplished through a combination of congestion control by the end users and limited congestion information feedback from the routers. A key assumption driving the design of such networks is the following well-known large deviations result: when the number of users in the network is large and the capacity of the network is large, then the probability that the arrival rate will exceed the available capacity is small. Thus, the probability of queue build-up is small [3]. In this paper, we examine the economics of providing large capacity from a service provider's point of view.

To make the problem tractable, we consider a single link accessed by many users, where each user is subject to congestion control and pricing. Many models have been suggested for pricing in networks. A pricing scheme that can be viewed as a natural extension of the behavior of TCP, the widely-used congestion control protocol, is to charge users for every congestion indication signal (for example, an ECN mark) that they receive [7], [13]. While this scheme leads to optimal resource allocation in a network, it may not be very practical. It is debatable as to whether end users would agree to a pricing scheme that dynamically varies with the load in the network. At the other extreme would be a pricing scheme where each user pays a fixed access charge to use the network. Such a scheme is widely practiced today by Internet service providers. However, this scheme suffers from the limitation that users utilizing different amounts of network resources are all treated equally; thus this scheme lacks any control over resource allocation.

In this paper, we consider an intermediate scheme where we assume that each user pays a price proportional to the amount of bandwidth that he/she uses. In addition, we assume that the network provides congestion indication signals that allow the users to adapt their transmission rates in response to network congestion. This congestion control action on the part of the users is assumed to be voluntary as in the case of TCP. However, if there is a need to police to unresponsive users, the network could impose a congestion price, but we assume that this is negligible compared to the bandwidth cost paid by the user.

Given the above pricing model, our goal is to understand if it is profitable for a service provider to increase the network capacity in proportion to the number of users. Our contributions in this paper can be summarized as follows:

- We model the interaction between the service provider and the users as a Stackelberg game [1]. The service provider sets the price per bandwidth and the users respond to the price by presenting a certain amount of flow to the network. The users do not cooperate among themselves, which thus leads to a Nash game among the users. We note that the solution to the Nash game is identical to the

Research supported by NSF Grants ANI-9813710, CCR 00-85917 ITR, and DARPA Grant F30602-00-2-0542.

Tamer Başar is with the Coordinated Science Laboratory and Department of Electrical and Computer Engineering, University of Illinois, tbasar@control.csl.uiuc.edu

R. Srikant is with the Coordinated Science Laboratory and Department of General Engineering, University of Illinois, rsrikant@uiuc.edu

congestion control solution of Kelly et al in [10] for the specific cost structure used in this paper. However, our primary focus here is on the Stackelberg game in order to understand the interplay between the price set by the network and congestion control.

- To understand the possible equilibrium solutions to the Stackelberg game, we consider a special case with a fixed link capacity of one unit and obtain explicit expressions for the users' flows and the optimal price. In addition to illustrating the difficulties in obtaining explicit solutions to the problem, this special case identifies certain key features of the optimal solution.
- We then consider a limiting regime where the number of users and the capacity of the link are increased simultaneously. In this case, we obtain explicit asymptotic solutions to the optimal price and the users' responses. Our main result shows that, as the number of users increases, the service provider's revenue per unit bandwidth as well as the performance of each user improve.

We note that there is an extensive literature on game-theoretic models of routing and flow control in communication networks (for example, see [19], [11], [12], [16], [4], [5]). These papers have presented conditions for the existence and uniqueness of an equilibrium. This has allowed, in particular, the design of network management policies that induce efficient equilibria [11]. This framework has also been extended to the context of repeated games in which cooperation can be enforced by using policies that penalize users who deviate from the equilibrium [16]. Our paper differs from the above papers due to the fact that our goal is to devise a revenue-maximizing pricing scheme for the service provider. Thus, a noncooperative (Nash) flow control game is played by the users (followers) in a Stackelberg game where the goal of the leader is to set a price to maximize revenue.

II. GAME-THEORETIC FORMULATION

Consider a link of capacity nc accessed by n users. Let p be the price per unit bandwidth charged by the network, and let x_i denote the transmission rate of User i. User i's objective is to maximize the following function with respect to x_i over $[0, nc - x_{-i})$:

$$F_i(x_i, x_{-i}; p) = w_i \log(1+x_i) - \frac{1}{nc - \sum_j x_j} - px_i \,, \quad (1)$$

where $w_i \log(1 + x_i)$ is the utility of the flow x_i to User i, $1/(nc - \sum_j x_j)$ represents the congestion cost on the link, and $x_{-i} := \sum_j x_j - x_i$. Unless otherwise stated, we will use $\sum_j x_j$ to denote $\sum_{j=1}^n x_j$, which we will also occasionally write as $\bar{x}$. If we assume that the queueing process at the link is $M/M/1$, then the above congestion cost is simply the delay on the link. For a given p, this defines a noncooperative game between the users of the network, with the underlying solution being the Nash equilibrium [1]. For each fixed $p > 0$, a Nash equilibrium for this n-player game is an n-tuple $\{x_i^*(p) \geq 0\}_{i=1}^n$ satisfying, for all $i \in \mathbf{N}$ (with $\mathbf{N} := \{1, \dots, n\}$),

$$\max_{0 \leq x_i \leq nc - x^*_{-i}} F_i(x_i, x^*_{-i}; p) = F_i(x_i^*, x^*_{-i}; p) \,. \quad (2)$$

Assuming that the game admits a unique Nash equilibrium (which we will prove to be the case), we associate with the service provider a revenue maximization problem to determine the *optimum* price to charge, namely

$$\max_{p \geq 0} L(p; \bar{x}^*(p)) \,, \qquad L(p; \bar{x}) := p\bar{x}$$

where $\bar{x}^*(p) := \sum_j x_j^*(p)$. What we have here is therefore a Stackelberg game [1], with one leader (with objective function L) and n noncooperative Nash followers (with objective functions F_i's).

Remark II.1: The utility function that we have chosen for User i is $w_i \log(1 + x_i)$, which is closely related to the utility function $w_i \log x_i$ that leads to proportionally fair resource allocation [8], [9]. If we use the utility function $w_i \log x_i$, however, then a user is forced to present a nonzero flow to the network since its utility becomes $-\infty$ if $x_i = 0$. Our utility function $w_i \log(1 + x_i)$ allows each user to exercise distributed admission control, i.e., decide whether to join the network or not. As we will see later, this ensures nontrivial and meaningful solutions to the Stackelberg game. On the other hand, if we use the utility function $w_i \log x_i$ for User i, then, for a given p, the first order necessary conditions (which are also sufficient) for a Nash equilibrium among the followers are given by

$$px_i = w_i - \frac{x_i}{(nc - \sum_j x_j)^2}, \quad i \in \mathbf{N},$$

leading to

$$p \sum_j x_j = \sum_j w_j - \frac{\sum_j x_j}{(nc - \sum_j x_j)^2}.$$

Thus, the optimal value of the revenue $p \sum_j x_j$ is $\sum_j w_j$, which is obtained by letting p go to ∞, which drives the x_j's to zero. Such an (unrealistic) situation arises due to the fact that each user is forced to send a nonzero flow over the link which the service provider can exploit by setting an arbitrarily high price.

As a final comment on the utility function, we note that it is straightforward to modify the results of this paper for the case where the utility function for User i is of the form $w_i \log(\delta + x_i)$, for any $\delta > 0$. ∎

III. Complete solution for a special case

We first consider here a special case of the problem formulated above, where $c = 1$. This special case allows for derivation of an explicit expression for the (unique) solution and provides insight into the solution of the more general case to be discussed in the next section.

Hence we have as the objective function of User $i \in \mathbf{N}$:

$$F_i(x_i, x_{-i}; p) = w_i \log(1 + x_i) - \frac{1}{n - \bar{x}} - px_i$$

where x_{-i} and $\bar{x}$ are as defined earlier, and p is a fixed price picked by the service provider. Now note that adding the quantity

$$\sum_{j \neq i} w_j \log(1 + x_j) - \sum_{j \neq i} px_j$$

to F_i and treating the resulting function as the new objective function of User i will not affect the Nash equilibrium, and hence in essence (as far as the Nash equilibrium is concerned) the original game is equivalent to one where all users have the identical objective function

$$F(x_1, \ldots, x_n; p) = \sum_{j=1}^{n} w_j \log(1+x_j) - \frac{1}{n - \bar{x}} - p\bar{x}\,. \quad (3)$$

Note that, for all $x_i,\ i \in N$ such that $\bar{x} < n$,

$$\begin{aligned} F_{x_i x_i} &= -\frac{w_i}{(1 + x_i)^2} - \frac{2}{(n - \bar{x})^3} < 0\,, \quad i \in \mathbf{N} \\ F_{x_i x_j} &= -\frac{2}{(n - \bar{x})^3} < 0\,, \quad j \neq i\,; \quad i, j \in \mathbf{N}\,. \end{aligned}$$

It is now easy to see that the *Hessian* matrix of F is negative definite, and thus F is strictly concave in the nonnegative orthant bounded by the hyperplane $\bar{x} = n$. Furthermore, $F \downarrow -\infty$ as $\bar{x} \uparrow n$, and hence the optimization problem

$$\max_{x_i \geq 0, i \in N; \bar{x} < n} F(x_1, \ldots, x_n; p)$$

admits a unique solution, which is also the only *person-by-person optimal* (that is, Nash equilibrium) solution. The optimal flows of the users can be obtained by solving the set of first-order conditions: for $i \in \mathbf{N}$,

$$\begin{aligned} F_{x_i}(x_1^*, \ldots, x_n^*; p) &= 0 \quad \text{if } x_i^* > 0 \\ &< 0 \quad \text{if } x_i^* = 0 \end{aligned}$$

where

$$F_{x_i} = \frac{w_i}{1 + x_i} - \frac{1}{(n - \bar{x})^2} - p\,.$$

Let us first consider the case when $x_i^* > 0\ \forall\, i \in \mathbf{N}$. It follows from $F_{x_i} = 0$ that (suppressing * on x_i throughout):

$$\frac{w_i}{1 + x_i} = \frac{w_j}{1 + x_j}\,, \qquad \forall\, i, j \in \mathbf{N}\,,$$

or equivalently,

$$y_i = \frac{w_i}{w_j} y_j\,, \qquad \forall\, i, j \in \mathbf{N}\,,$$

where $y_i := 1 + x_i$. Letting

$$\bar{y} := \sum_j y_j\,,\ \bar{w} := \sum_j w_j\,,$$

one can express y_i's in terms of $\bar{y}$:

$$y_i = \frac{w_i}{\bar{w}} \bar{y}\,, \qquad \forall\, i \in \mathbf{N}\,.$$

Using this in the first-order conditions $F_{x_i} = 0\,,\ i \in \mathbf{N}$, we obtain

$$g(\bar{y}) =: \frac{\bar{w}}{\bar{y}} - \frac{1}{(2n - \bar{y})^2} - p = 0\,, \quad (4)$$

for which we seek a solution in the interval $(n, 2n)$. Note that g is strictly decreasing in the interval $(n, 2n)$, and furthermore $g(2n^-) = -\infty$ and $g(n) = \bar{w} - n^{-2} - p$. Hence, equation (4) admits a solution (and a unique one) in the given interval if, and only if, $g(n) > 0$, that is

$$p < w_{av} - \frac{1}{n^2} =: \hat{p}\,, \quad (5)$$

where $w_{av} := (1/n)\bar{w}$. Let us denote the unique solution to (4) for p satisfying (5) by $\bar{y}(p)$. Clearly, the constraint on p is necessary for the existence of a positive Nash equilibrium, since if p exceeds the given bound, then there will not exist a feasible solution to (4), meaning that the only Nash equilibrium (equivalently, the only maximizing solution to (3)) will then dictate some of the users not to transmit at all. Note that the constraint (5) becomes progressively less restrictive as the number of users, n, increases, asymptotically reaching the bound w_{av}, assuming of course that the latter is well-defined as $n \to \infty$.

Hence, the discussion above shows that the constraint (5) is *necessary* for existence of a positive solution to the maximization problem with objective function (3); it is, however, not sufficient, as $\bar{y} > n$ does not imply positivity of the individual x_i's. For this, we also need

$$\frac{w_i}{\bar{w}} \bar{y}(p) > 1\,, \qquad \forall\, i \in \mathbf{N}\,. \quad (6)$$

Note that this condition brings in dependence on the individual w_i's, whereas (5) required only the knowledge of their sum, $\bar{w}$. The condition (6) says that w_i's should not be too far from their average value $w_{av} := (1/n)\bar{w}$, that is they should cluster around w_{av}. If this is not the case, the users with smaller w_i's will have to drop out of the game, that is not transmit at all, as we will see later in this section. But for the moment, since we are interested only in the positive Nash equilibrium, let us assume that for a given p satisfying only (5), w_i's are distributed in such a way that condition (6) also holds; a precise condition for this will be derived shortly.

The next step in the solution process is to obtain a revenue-maximizing price for the service provider, that is to solve the maximization problem

$$\max_{0<p<\hat{p}} L(p;\bar{x}(p))\,, \qquad L := p\,\bar{x}(p)\,,$$

where $\bar{x}(p)$ is the sum of individual users' transmission rates in the Nash game, in response to a fixed price p; this quantity is of course also equal to $\bar{y}(p)-n$. Since there is a one-to-one correspondence between $p \in (0,\hat{p})$ and $\bar{y} \in (n, 2n)$ through (4), the service provider's optimization problem can equivalently be carried out with respect to $\bar{y}$ using (4) as a constraint, that is:

$$\max_{n<\bar{y}<2n} L(p(\bar{y});\bar{y}-n)\,,$$

where

$$L(p(\bar{y});\bar{y}-n) = (\bar{y}-n)\left(\frac{\bar{w}}{\bar{y}} - \frac{1}{(2n-\bar{y})^2}\right).$$

This is again a strictly concave function in the given open interval, and since it becomes unbounded negative as $\bar{y} \uparrow 2n$ and is *zero* at $\bar{y} = n$, it admits an inner maximum (which is also unique) if, and only if, the first-order stationarity condition leads to a solution in the given interval. The condition (obtained by simply differentiating $L(p(\bar{y});\bar{y}-n)$ above with respect to $\bar{y}$) is:

$$-(\bar{y}-n)\left(\frac{\bar{w}}{\bar{y}^2} - \frac{2}{(2n-\bar{y})^3}\right) + \frac{\bar{w}}{\bar{y}} - \frac{1}{(2n-\bar{y})^2} = 0\,,$$

which admits the unique solution[1]

$$\bar{y}^* = \frac{2n(n\bar{w})^{\frac{1}{3}}}{1+(n\bar{w})^{\frac{1}{3}}}$$

provided that

$$n\bar{w} > 1$$

which ensures that $\bar{y}^* > n$, that is the total throughput is positive. Now, for the individual throughput levels to be positive, we need also the condition (6):

$$\frac{w_i}{\bar{w}} \cdot \frac{2n(n\bar{w})^{\frac{1}{3}}}{1+(n\bar{w})^{\frac{1}{3}}} > 1\,, \quad \forall\, i \in \mathbf{N}\,,$$

[1]The fact that the solution can be obtained explicitly (in closed form) here is precisely the reason why we have discussed the case of $c = 1$ separately. For $c \neq 1$ such an explicit solution cannot be obtained, except asymptotically for large n, as we will see in the next section.

which can be re-arranged and written in the more appealing form, in terms of w_{av}:

$$\left(\frac{2w_i}{w_{av}} - 1\right)(n^2 w_{av})^{\frac{1}{3}} > 1\,, \quad \forall\, i \in \mathbf{N}\,. \tag{7}$$

This condition is of course more restrictive than the earlier one, $n^2 w_{av} > 1$, which will therefore be dropped.

The revenue-maximizing price (for the service provider) can now be obtained from (4), to read (also written in terms of w_{av}):

$$p^* = \frac{w_{av}}{2}(1+(n^2w_{av})^{-\frac{1}{3}}) - \frac{1}{4n^2}(1+(n^2w_{av})^{\frac{1}{3}})^2 \tag{8}$$

which can easily be checked to be positive provided that $n^2 w_{av} > 1$, a condition already assumed to hold. It is also easy to check that $p^* < \hat{p}$. The following theorem now captures the complete positive solution to the problem of this section. In the statement of the theorem, we use the notation $f(n) \sim h(n)$ to mean $\lim_{n\to\infty}(f(n)/h(n)) = 1$.

Theorem III.1: The special case of the general problem of this paper, characterized by $c = 1$, admits a solution where all n users transmit at positive rates if, and only if, condition (7) holds. In this case, the solution is unique, with the optimal price charged by the service provider being (8), and with the Nash equilibrium transmission rates of the users given by

$$x_i^* = \frac{w_i}{w_{av}}(x_{av}^*+1) - 1 \tag{9}$$

where x_{av}^* is the optimal average throughput, given by

$$x_{av}^* = 1 - \frac{2}{1+(n^2w_{av})^{\frac{1}{3}}}\,. \tag{10}$$

Asymptotically, for large n, and with the smallest w_i bounded away from *zero*, condition (7) is satisfied, and

$$p^* \sim \frac{w_{av}}{2} + \frac{(w_{av})^{\frac{2}{3}}}{4}n^{-\frac{2}{3}}\,, \quad x_{av}^* \sim 1 - 2(w_{av})^{-\frac{1}{3}}n^{-\frac{2}{3}}\,. \tag{11}$$

Again asymptotically, the revenue per unit bandwidth, which is the product of p^* and x_{av}^*, becomes

$$\text{Revenue/bw} \sim \frac{w_{av}}{2} - \frac{3}{4}(w_{av})^{\frac{2}{3}}n^{-\frac{2}{3}}\,, \tag{12}$$

the congestion cost is given by

$$D^* = \frac{1}{n(1-\bar{x}^*)} \sim \frac{1}{2}(w_{av})^{\frac{1}{3}}n^{-\frac{1}{3}}\,, \tag{13}$$

and the equilibrium net utility levels of the users admit the expressions

$$F_i^* = w_i \log \frac{2w_i}{w_{av}} + \frac{1}{2}w_{av} - w_i - \frac{1}{2}(w_{av})^{\frac{1}{3}} n^{-\frac{1}{3}} . \quad (14)$$

■

Remark III.1: An interesting observation that can be made from the asymptotic results of the theorem above is that for n sufficiently large, the optimum price decreases with increasing n, while the average (per user) throughput increases with n. The net effect on the revenue per unit bandwidth is that it increases with n, and so do the utility levels of individual users–all very appealing features. Note also that the congestion cost goes asymptotically to *zero*, with the convergence being monotonic for large n. All these features will be discussed further in the next section, where we will see that the solution to the original problem (that is, with general c) also shares them. ■

The result of Theorem III.1 relies on condition (7), which we have already seen to be satisfied if n is sufficiently large. The other extreme case is the single-user problem ($n = 1$), in which case it is also satisfied, provided that $w_1 > 1$. In this extreme case, one can easily see that if $w_1 \leq 1$ then the optimum throughput is simply $x_1^* = 0$. Now, to obtain the complete solution for the general n case, let us first adopt the convention that the users are ordered in decreasing values of the w_i's, that is $w_i > w_j \Rightarrow i < j$, with $w_1 > 1$. Let n^* be the largest integer $\tilde{n} \leq n$ for which the following inequality is satisfied:

$$\left(\frac{2w_{\tilde{n}}}{\tilde{w_{av}}} - 1\right)(\tilde{n}^2 \tilde{w_{av}})^{\frac{1}{3}} > 1 \, , \quad (15)$$

where $\tilde{w_{av}}$ is defined as w_{av} with n replaced by $\tilde{n}$. Clearly, if the original n-user problem is replaced with a new one that retains only the first n^* users, this new one will admit a unique solution that is covered by Theorem III.1, with n replaced by n^*. Furthermore, this unique solution, appended with $x_j = 0, \;\; j > n^*$, constitutes the unique solution to the original n-user problem. This follows because of the uniqueness of the solution to the noncooperative game faced by the users for each p, the fact that $\hat{p}$ is strictly decreasing in n, and the property that with $\tilde{n}$ fixed, no other ordering of the users can lead to a higher value for the left-hand-side of (15). We now summarize this result (the complete solution to the problem of this section) in the following theorem.

Theorem III.2: Consider the problem of this section, with the indexing of the users such that $w_i > w_j \Rightarrow i < j$. Further, let $w_1 > 1$, and n^* be defined as above (the largest integer $\tilde{n}$ satisfying (15)). Then, $n^* \geq 1$, and the problem admits a unique solution with x_i^*, $i \leq n^*$, and p^* given as in Theorem III.1 with n^* replacing n, and $x_j^* = 0, \; j > n^*$. ■

Example III.1: Consider a network with $w_1 = 3, w_i = 1 \;\forall i \geq 2$. Checking the condition of Theorem III.2 for different values of n, we have:

$$n^* = 1 \quad \text{if} \;\; n \leq 4; \qquad n^* = n \quad \text{for} \;\; n \geq 5 \, .$$

Computing the optimum throughputs, price levels, revenue per unit bandwidth, and user's net utility levels, F_i^*, for selected n, we have:

$$n = 1 \Rightarrow x_{av}^* = 0.1811, \; p^* = 1.0489, \; p^* x_{av}^* = 0.1899, \\ F_1^* = -1.00428$$

$$n = 5 \Rightarrow x_{av}^* = 0.5317, \; p^* = 0.7316, \; p^* x_{av}^* = 0.3890, \\ x_1^* = 2.2823, \; x_j^* = 0.0941, j \geq 2, \\ F_1^* = -0.548245, \; F_j^* = -0.456863, j \geq 2$$

$$n = 10 \Rightarrow x_{av}^* = 0.6629, \; p^* = 0.6337, \\ p^* x_{av}^* = 0.4200, \; x_1^* = 3.15722, \; x_j^* = 0.3857, j \geq 2, \\ F_1^* = -0.440802, \; F_j^* = -0.399363, j \geq 2$$

$$n \to \infty \Rightarrow x_{av}^* = 1, \; p^* = 0.5, \; p^* x_{av}^* = 0.5, \\ x_1^* = 5, \; x_j^* = 1, j \geq 2, \\ F_1^* = -0.165546, \; F_j^* = -0.19897, j \geq 2 \, .$$

The numerical results above indicate that the asymptotic behavior alluded to earlier holds even for relatively small values of n, as in going from $n = 1$ to 5 and then to 10, we see the average throughput, x_{av}^*, increasing, price, p^*, decreasing , the revenue per unit bandwidth increasing, and the utility levels of each user improving, thus showing an overall monotonic trend toward the values reached as $n \to \infty$. Another interesting feature of the results above is that adding additional capacity for potentially new users is not always a revenue-improving move for the service provider. As we see from the above, as compared with the single-user case, a benefit accrues to the service provider only if 4 or more additional users (with lower w_i parameter values) are admitted to the network (with an additional unit of capacity added for each new user). ■

IV. The general solution

We now return to the original problem with a general c, and develop the solution by following essentially the same steps as in the previous section. By a slight abuse, we will use the same notation as in the previous section, where appropriate. Also, for convenience we will order the users according to their w_i values, so that $w^i > w^j$ will imply $i < j$.

We first consider the Nash game faced by the n users, for each fixed $p > 0$. The following lemma settles the issues of existence and uniqueness of Nash equilibrium for this game.

Lemma IV.1: For each fixed $p > 0$, the n-person noncooperative game where Player i's objective function (to be maximized) is given by (1), admits a unique Nash equilibrium $\{x_i^*(p) \geq 0\,;\; i \in \mathbf{N}\}$, with $\bar{x}^*(p) = \sum_j x_j^*(p) < nc$.

Proof: Following the line of reasoning in the previous section, every Nash equilibrium solution of this game is also a Nash equilibrium solution of a game with the following objective for User i:

$$F(x_1, \ldots, x_n; p) = \sum_j w_j \log(1 + x_j) - \frac{1}{nc - \bar{x}} - p\bar{x}\,.$$

Note that this objective function is identical for all users, and is strictly concave in the n-tuple $(x_1, \ldots, x_n)$ which is restricted to the nonnegative orthant bounded by the hyperplane $\bar{x} = nc$ on which F is unbounded from below. Then, from standard results in finite-dimensional optimization [2], it follows that F has a unique maximum in this bounded region (where it is finite, except on the given hyperplane), and every *person-by-person optimal* solution is also globally optimal. Hence, the Nash equilibrium exists and is unique. Clearly, the maximizing solution cannot lie on the hyperplane, thus the strict inequality on $\bar{x}^*(p)$. ∎

Depending on the value of p and values of other parameters, the unique solution alluded to in the lemma above could lie on the lower boundary of the constraint region (that is some of the x_i's could be *zero*). If this does not happen, we say that the Nash equilibrium is *inner* or *positive*. We now first obtain necessary and sufficient conditions for the equilibrium to be positive, and obtain a characterization for it. Clearly, from first-order conditions (which are also sufficient), the Nash equilibrium will be positive if, and only if, the following set of equations admits a positive solution (for x_i's):

$$\frac{w_i}{1 + x_i} - \frac{1}{(nc - \bar{x})^2} - p = 0, \tag{16}$$

or, equivalently,

$$w_i = (1 + x_i)\left(\frac{1}{(nc - \bar{x})^2} + p\right).$$

Taking the sum of the above equation over all i and defining $\bar{y} = \bar{x} + n$, we get

$$g(\bar{y}) := \frac{\bar{w}}{\bar{y}} - \frac{1}{(nc + n - \bar{y})^2} - p = 0\,. \tag{17}$$

Using (17) in (16), and solving for x_i, we obtain

$$x_i = \frac{w_i}{\bar{w}}\bar{y} - 1\,, \quad i \in \mathbf{N}\,, \tag{18}$$

which are required to be positive. Hence, the Nash equilibrium is positive for a given p if, and only if, there exists a $\bar{y} > n$ solving (17) and further making (18) positive. Since $g(\bar{y} \uparrow n(c+1)) = -\infty$, and $g(\bar{y} = n) = w_{av} - (nc)^{-2} - p$, and g is strictly decreasing in $[n, (c+1)n)$, similar to the development in the previous section, there will exist a unique solution to (17) in the open interval $(n, (c+1)n)$ if, and only if, $g(n) > 0$, that is

$$p < w_{av} - \frac{1}{(nc)^2} =: \hat{p} \tag{19}$$

which is the counterpart of (5) for the more general game of this section. Hence, for any p satisfying (19), the unique Nash equilibrium is positive if, and only if, the unique solution, $\bar{y}^*(p)$, to (17) satisfies (in view of (18) and the ordering of the users) the condition

$$\frac{w_n}{\bar{w}}\bar{y}^*(p) > 1\,. \tag{20}$$

Now proceeding on to the maximization problem (2) faced by the service provider, because of the one-to-one correspondence between $\bar{y}$ (or equivalently $\bar{x}$) and p through the constraint (17), a problem equivalent to (2) is maximization of the following function with respect to $\bar{y} > n$ (obtained by substitution of p from (17) in terms of $\bar{y}$):

$$\tilde{L}(\bar{y}) = \bar{w}(1 - \frac{n}{\bar{y}}) - \frac{\bar{y} - n}{(n(c+1) - \bar{y})^2}\,.$$

We seek a solution to this maximization problem in the open interval $(n, (c+1)n)$. $\tilde{L}$ is analytic over this interval, and

$$\tilde{L}_{\bar{y}} = \frac{n\bar{w}}{\bar{y}^2} - \frac{n(c-1) + \bar{y}}{(n(c+1) - \bar{y})^3}\,, \quad \tilde{L}_{\bar{y}\bar{y}} < 0.$$

Hence, $\tilde{L}$ is strictly concave, and further since it becomes unbounded negative at the upper end of the interval, it follows that it has a unique maximum in the half-open interval $[n, (c+1)n)$. Moreover, the situation $\bar{y} = n$ can be avoided by requiring that $\tilde{L}_{\bar{y}}(n) > 0$, which translates into the simple condition

$$n^2 c^2 w_{av} > 1\,. \tag{21}$$

Under this condition, there exists a unique solution, $\bar{y}^* \in (n, (c+1)n)$, to $\tilde{L}_{\bar{y}}(\bar{y}) = 0$ which we rewrite here for future reference:

$$\frac{n\bar{w}}{\bar{y}^2} - \frac{n(c-1)+\bar{y}}{(n(c+1)-\bar{y})^3} = 0\,. \tag{22}$$

The corresponding value of p, which in fact maximizes $L(p; \bar{x}^*(p))$, is then obtained directly from (17):

$$p^* = \frac{\bar{w}}{\bar{y}^*} - \frac{1}{(nc+n-\bar{y}^*)^2} \tag{23}$$

which, by construction, satisfies (19) and the positivity constraint. To complete the solution to the problem, however, we still have to require that (20) holds with p replaced by p^*, or equivalently

$$\frac{w_n}{\bar{w}}\bar{y}^* > 1\,, \tag{24}$$

with $\bar{y}^*$ solving (22). The counterpart of this condition for the case $c = 1$ is (7), which is more explicit because $\bar{y}^*$ admitted a closed-form expression in that case; in the present more general case, $\bar{y}^*$ satisfies a third-order polynomial equation, which does not admit an easily expressible solution.

Now, if the two conditions above, that is (21) and (24), do not hold, then one has to look for a boundary solution, that is a solution where some users are dropped out of the network. In this scheme, users will have to be dropped in reverse order, until the first time both these conditions hold; let this number be n^*. That is, n^* is the maximum number of ordered users for which conditions (21) and (24) hold. As in the previous section, no other ordering will lead to a higher number of users for whom the conditions hold. Then, the problem is essentially equivalent to a network with only n^* users, for which we again have an inner Nash equilibrium.

This now brings us to the following theorem, which is the counterpart of Theorem III.2 for this general case.

Theorem IV.1: Consider the general problem of this section, with the indexing of the users such that $w_i > w_j \Rightarrow i < j$. Further let $c^2 w_1 > 1$, and n^* be defined as above (the largest integer n for which conditions (21) and (24) both hold). Then, $n^* \geq 1$, and the problem admits a unique solution $\{x_1^* \ldots, x_n^*; p^*\}$ as follows: $x_i^* = (w_i/\bar{w}^*)\bar{y}^* - 1,\ i \leq n^*$, where $\bar{w}^*$ is the sum of the first n^* w_i's and $\bar{y}^*$ is the unique solution of (22) with n replaced by n^*; p^* is given by (23) with again n^* replacing n; and $x_j^* = 0,\ j > n^*$.

Proof: Proof follows readily from the discussion preceding the statement of the theorem, in view of also the discussion on the special ($c = 1$) case in the previous section. We should simply point out here that the reason why $n^* \geq 1$ under the condition $c^2 w_1 > 1$, is because this condition is precisely (21) when $n = 1$, and there is no need for (24) in the single user case (the condition is automatically satisfied). ∎

Remark IV.1: In deriving the solution to the Stackelberg game, we have assumed that the users respond to the service provider's price by reaching a Nash equilibrium. Given a price p, our model fits into the framework developed by Kelly et al [10], and hence the Nash equilibrium can be reached in a distributed fashion by the users, with each user implementing a congestion control algorithm. Consider the congestion control algorithm

$$\dot{x}_i = \kappa_i(w_i - p - px_i - (1+x_i)q(\sum_j x_j))$$

where

$$q(\lambda) := \frac{1}{(nc-\lambda)^2}.$$

If the Nash equilibrium solution is positive, then the equilibrium point of this system of congestion controllers is also the unique Nash equilibrium solution. Following [10], it is now easy to verify that $F(x_1, x_2, \ldots, x_n; p)$, defined in the proof of Lemma IV.1, is a Lyapunov function for this system of congestion controllers and hence, the system converges to its equilibrium point. Even in the presence of noise in the system due to various sources of randomness such as departures and arrivals of short-lived flows, inaccurate congestion feedback, etc., it has been established that, under appropriate limiting regimes, the above congestion control equations are valid [21]. Recently, local convergence of these congestion controllers has been established even in the presence of delays for appropriate choice of the controller gains κ_i [6], [18], [22]. Similar results for the dual version of the optimization problem can be found in [20]. ∎

V. MANY-USERS REGIME

We now study the behavior of the solution obtained in the previous section for large n. Studying this *many-users regime* will allow us to obtain an explicit expression for $\bar{y}^*$ (or $\bar{x}^*$), which was not possible for finite n, unless $c = 1$.

The study will also enable us to ask (and answer) questions like whether it is possible for the service provider to admit new users to the network by increasing the capacity, and whether the existing users would benefit (measured in terms of their utilities) from a "crowding" of the network.

An underlying assumption (or rather convention) throughout this section is that as $n \to \infty$, the sequence $\{w_{av}(n) := (1/n)\sum_{j=1}^{n} w_j\}$ has a well-defined limit, w_{av}. This would be the case, for example, when beyond a finite n the users added to the network all have the same w_i's, that is $w_i = w \ \ \forall\, i \geq n+1$. In this case, of course, $w_{av} = w$. An immediate implication of this assumption is that now the condition (21) is readily satisfied.

Now, to study the asymptotic behavior, it is also convenient to work with the arithmetic mean of the x_i's (or y_i's), rather than their sum, denoted for the former as

$$x_{av}(n) = \frac{1}{n}\sum_{j=1}^{n} x_j \,.$$

Then, we can re-write (22) as

$$\frac{w_{av}(n)}{(x_{av}(n)+1)^2} = \frac{c + x_{av}(n)}{n^2(c - x_{av}(n))^3} \,.$$

Under our assumption that $w_{av}(n) \to w_{av}$ as $n \to \infty$, a positive solution to $x_{av}(n)$ exists for large n if, and only if,

$$\lim_{n\to\infty} n^2(c - x_{av}(n))^3 = \alpha$$

for some $\alpha > 0$. Substituting this in (17), we obtain

$$p \sim \frac{w_{av}}{c+1} + \frac{2c-1}{\alpha^{2/3} n^{2/3}} \,, \tag{25}$$

where we have again used the notational convention that $f(n) \sim h(n)$ if $\lim_{n\to\infty}(f(n)/h(n)) = 1$. Using (25) in (22), and letting $n \to \infty$, yields

$$\alpha = \frac{2c(c+1)^2}{w_{av}} \,. \tag{26}$$

Then,

$$x_{av}(n) \sim c - n^{-2/3}\alpha^{1/3} = c - \left(\frac{2c(c+1)^2}{w_{av}}\right)^{\frac{1}{3}} n^{-\frac{2}{3}} \,, \tag{27}$$

and from (18),

$$\begin{aligned} x_i(n) &= \frac{w_i}{w_{av}(n)}(x_{av}(n)+1) - 1 \\ &\sim (c+1)\frac{w_i}{w_{av}} - 1 - \frac{w_i}{w_{av}}\alpha^{\frac{1}{3}} n^{-\frac{2}{3}} \end{aligned} \tag{28}$$

Now, for n sufficiently large, a positivity condition on the x_i's can be obtained from (28) by simply requiring the *zeroth*-order term to be positive:

$$\min_i w_i > \frac{w_{av}}{c+1} \,, \tag{29}$$

which is essentially condition (24) in the limit as $n \to \infty$. An interpretation for this for each i is that w_i for User i must be more than the limiting value of the price-per-unit-bandwidth for the user to be able to send nonzero flow over the link. Overall, (29) is precisely the condition for all users in the many-users regime to send positive flow over the network, as dictated by the optimal solution. Users whose w_i values fall short of the bound $w_{av}/(c+1)$ drop out of the network. If, on the other hand, there is a high population of users with the lowest w_i values, in which case asymptotically $w_{av}(n) \sim \min_i w_i$, the condition will be satisfied for all positive c, thus imposing no specific conditions on c; otherwise, the value of c will have to be increased to accommodate the users with small w_i values.

It is worth also noting that the optimal price charged by the network is positive for sufficiently large n, but whether it is an increasing or decreasing function of n depends on the specific value of c. For $c > 1/2$, it decreases with n, whereas for $c < 1/2$, it increases with n. In spite of this c-dependent behavior of the optimum price, the revenue per unit bandwidth exhibits a c-independent trend–increasing with n in the many-users regime:

$$\text{Revenue/bw} = \frac{p x_{av}}{c} \sim \frac{w_{av}}{c+1} - 3\alpha^{-\frac{2}{3}} n^{-\frac{2}{3}} \,.$$

Suppose that $w_{av}/(c+1)$ is larger than the cost of adding one unit of bandwidth, then, as the number of users increases, the service provider's profit increases. Thus, the service has an incentive to increase the link capacity which drives the congestion cost to zero as shown next.

The congestion cost decreases with n in the many-users regime:

$$\text{Congestion cost} = \frac{1}{n(c - x_{av}(n))} \sim \alpha^{-\frac{1}{3}} n^{-\frac{1}{3}} \,,$$

and so does the net utility of each user:

$$F_i^* = w_i \log\frac{(c+1)w_i}{w_{av}} - w_i + \frac{w_{av}}{c+1} - \alpha^{-\frac{1}{3}} n^{-\frac{1}{3}} \,.$$

Remark V.1: In the limit as $n \to \infty$, the congestion costs go to zero. Thus, it would be interesting to explore if the limiting solution coincides with the solution to the Stackelberg game where the followers (users) do not have congestion costs in their objectives. Accordingly, suppose that User i's objective is to maximize the following function with respect to x_i over $[0, nc - x_{-i})$:

$$F_i(x_i, x_{-i}; p) = w_i \log(1 + x_i) - p x_i \,. \tag{30}$$

subject to the constraint $\sum_j x_j \leq nc$. For a fixed p, as argued earlier, the Nash equilibrium solution coincides with the solution to a team problem with the following objective function:

$$F(x_1, x_2, \ldots, x_n; p) = \sum_j w_j \log(1 + x_j) - p \sum_j x_j,$$

which is subject to the same constraint as above. Using Lagrange multiplier approach [2], the unique solution to the above constrained optimization problem can be obtained as follows:

$$x_i = \frac{w_i}{p + \lambda} - 1$$

where λ satisfies

$$\lambda(\sum_j x_j - nc) = 0; \qquad \lambda \geq 0.$$

If $\lambda = 0$, the above conditions lead to

$$p \geq \frac{\bar{w}}{(n+1)c}$$

and

$$x_i(p) = \frac{w_i}{p} - 1.$$

On the other hand, if $\lambda > 0$, then the inequality constraint would be an equality, leading to

$$\lambda = \frac{\bar{w}}{(n+1)c} - p > 0,$$

and

$$x_i(p) = \frac{w_i(n+1)c}{\bar{w}}.$$

Maximization of $p \sum_j x_j(p)$ over $p \geq 0$ dictates that the revenue-maximizing price is given by

$$p^* = \frac{w_{av}}{c+1}$$

which coincides with the *zeroth* order term in (25). However, the fact that this is indeed the *zeroth* order term cannot be established without going through the full solution to the Stackelberg game and carrying out the asymptotic analysis. Further, the asymptotic expansion provides insight into the behavior of the price, the service provider's total revenue, revenue per unit bandwidth, and the net utilities of the users, for finite, but large, n. ∎

VI. Conclusions

The roles of the congestion controllers at the end users and active queue management schemes at the routers in designing low-loss, low-queueing-delay networks are well documented. The key assumption required for providing such a high quality-of-service is that the capacities of the links in the network are large, in proportion to the number of users in the network, which is typically large. In this paper, we have examined the question of whether there is an economic incentive for a service provider to provision large link capacities. For a simple single-link model, we have established the following result: under a pricing scheme that charges users for bandwidth (bits-per-second) usage and a collection of congestion-controlled sources subject to distributed admission control, the optimal revenue of the service provider increases with the number of users if the provisioned capacity is proportional to the number of users.

References

[1] T. Başar and G. J. Olsder. *Dynamic Noncooperative Game Theory*. SIAM Series in Classics in Applied Mathematics, Philadelphia, PA, 1999.

[2] D. Bertsekas. *Nonlinear Programming*. Athena Scientific, Belmont, MA, 1995.

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

[4] E. Altman, T. Başar and R. Srikant. Nash equilibria for combined flow control and routing in networks: Asymptotic behavior for a large number of users. *IEEE Transactions on Automatic Control*, 2002 (to appear).

[5] E. Altman, T. Başar, T. Jiminez and N. Shimkin. Competitive routing in networks with polynomial costs. *IEEE Transactions on Automatic Control*, 2002 (to appear).

[6] R. Johari and D. Tan, "End-to-end Congestion Control for the Internet: Delays and Stability," preprint. A shorter version was presented at the *Workshop on Stochastic Networks*, Madison, WI, June, 2000.

[7] R.J. Gibbens and F.P. Kelly. Resource pricing and the evolution of congestion control, Automatica, 35(6):1969–1985, 1999.

[8] F.P. Kelly. Models for a self-managed Internet . *Philosophical Transactions of the Royal Society* A358:2335-2348, 2000.

[9] F.P. Kelly. Mathematical modelling of the Internet. In *Mathematics Unlimited - 2001 and Beyond* (Editors B. Engquist and W. Schmid). Springer-Verlag, Berlin, pages 685-702, 2001.

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

[11] Y. Korilis, A. Lazar, and A. Orda. Architecting noncooperative networks. *IEEE Journal on Selected Areas in Communications*, 13:1241–1251, 1995.

[12] Y. A. Korilis, A. A. Lazar, and A. Orda. Achieving network optima using Stackelberg routing strategies. *IEEE/ACM Transactions on Networking*, 5:161–173, Feb. 1997.

[13] S. Kunniyur and R. Srikant. End-to-end congestion control: utility functions, random losses and ECN marks. In *Proceedings of INFOCOM 2000*, Tel Aviv, Israel, March 2000.

[14] S. Kunniyur and R. Srikant. Analysis and design of an adaptive virtual queue algorithm for active queue management. In *Proceedings of ACM Sigcomm*, San Diego, CA, August 2001.

[15] S. Kunniyur and R. Srikant. A time-scale decomposition ap-

proach to adaptive ECN marking. In *Proceedings of the IEEE INFOCOM*, Anchorage, AK, April 2001.

[16] R. J. La and V. Anantharam. Optimal routing control: a game theoretic approach. In *Proceedings of the 36th IEEE Conference on Decision & Control*, San Diego, CA, December 1997.

[17] S. H. Low and D. E. Lapsley. Optimization flow control, I: Basic algorithm and convergence. *IEEE/ACM Transactions on Networking*, 35(6):861–875, December 1999.

[18] L. Massoulie, "Stability of Distributed Congestion Control with Heterogeneous Feedback Delays," Microsoft Research Technical Report, 2000.

[19] A. Orda, R. Rom, and N. Shimkin. Competitive routing in multiuser communication networks. *IEEE/ACM Transactions on Networking*, 1(5):510–521, October 1993.

[20] F. Paganini, J. Doyle and S. Low. "Scalable laws for stable network congestion control." *Proceedings of the IEEE Conference on Decision and Control*, pp. 185–190, Orlando, FL, December 4-7, 2001.

[21] S. Shakkottai and R. Srikant. Mean FDE Models for Internet Congestion Control Under a Many-Flows Regime. 2001. Available at http://www.comm.csl.uiuc.edu/˜srikant/pub.html

[22] G. Vinnicombe, On the stability of end-to-end congestion control for the Internet, *University of Cambridge Technical Report*, 2001. Also in *Proceedings of the 39th Allerton Conference on Control, Communications and Computing* (October 3-5, 2001), 2002 (to appear).