Distributed algorithms for fair bandwidth allocation to elastic services in broadband networks

Haïkel Yaïche
Dept. of Elec. and Comp. Engg.
Ecole Polytechnique de Montreal
Montreal, P.Q. H3C 3A7, Canada
yaiche@comm.polymtl.ca

Ravi R. Mazumdar
School of ECE
Purdue University
W. Lafayette, IN 47907, USA
mazum@ecn.purdue.edu

Catherine Rosenberg
School of ECE
Purdue University
W. Lafayette, IN 47907 USA
cath@ecn.purdue.edu

Abstract—**The Nash arbitration scheme from cooperative game theory provides a natural framework to address the allocation of available bandwidth in network links which is network (Pareto) optimal and satisfies precise notions of fairness. In this paper we propose two distributed bandwidth allocation schemes that allocate available badwidths to elastic sources according to the Nash arbitration scheme. We prove convergence to the desired allocations for both algorithms. Finally we show how such a scheme can be implemented in a real network.**

I. Introduction

A major aspect that has to be considered in the design of elastic services is flow control since elastic sources have to continually change their rates on the basis of notifications sent by the network [9]. There are two main types of flow control [7]: window flow control in which source adjust dynamically their window that represents the maximum number of packets (or cells) allowed into the network, and rate flow control in which a sources adjusts a rate representing the maximum rate at which it can send data. In this paper we are interested in the development of rate flow control schemes.

The role of flow control for elastic services in broadband networks consists in allocating the available capacities within the network to competing (or active) connections in an efficient and fair way. The issue of fairness has been addressed in the context of the Available Bit Rate service in ATM networks by considering mainly the max-min fair allocation and its variants [12]. In a previous work [19], we have addressed the issue of allocating bandwidths to connections using a game theoretic approach, and have proposed a comprehensive allocation policy based on the Nash arbitration scheme notion that satisfies some fairness axioms. In this paper, we consider this fairness criterion as an allocation criterion that an elastic service *rate* control algorithm should satisfy.

The Nash arbitrated allocation takes into account the bandwidth requirements of an elastic connection defined as a Minimum Rate (MR) and a Peak Rate (PR). It is characterized by a *global optimization* problem requiring the knowledge of connection bandwidth requirements, link available capacities, and connection-link incidence matrix. A rate control scheme in which allocated rates are computed by a central system is unreliable because the communication links connecting the central to the network are subject to failures, and has high communication overheads since it requires a global knowledge of the varying state of the network.

It is necessary to devise a distributed and fair rate control scheme in which network nodes (switches or routers) and elastic sources run local procedures and cooperate by exchanging relevant information. To this end, we address in the present paper the issue of determining the Nash arbitrated allocation in a distributed way.

Many recent works have used an optimization approach to derive rate-based control schemes for elastic services. Optimization based algorithms for rate control have been introduced in [13] and [14]. They consider the maximization of the sum of user utilities under network and user bandwidth constraints and suggest distributed rate control algorithms. They prove the stability (convergence under quasi-static assumptions) based on mathematical idealizations as well as higher order differentiability assumptions on the utility functions of the users. It should be noted that optimizing sums of utility functions does not in general lead to a Nash arbitration point with plays an important role in that it satisfies the important notion of fairness. The case for operating systems at Nash arbitration points has recently been also underscored in the context of IP networks [8].

In this paper we propose a performance based criterion rather than an abstract utility based approach since these utilities are difficult to quantify mathematically. In the context of packet-switched networks Mazumdar *et al* [5] have used a game theoretic approach and derived a network global objective in order to characterize a fair performance-oriented flow control mechanism. Recently, the use of such a framework in the context of bandwidth allocation subject to both capacity and budget constraints for broadband services was considered in [19].

The paper is organized as follows. Section 2 studies the global optimization problem, (S), that characterizes the Nash arbitrated allocation. In section 3 two optimization algorithms solving (S) are presented. The first algorithm is based on the optimality conditions. The second one is based on the corresponding dual problem using a gradient-based projection technique. These two algorithms can be implemented in a distributed way. In Section 4 we propose implementations of rate-based control schemes for elastic services based on the two algorithms.

Due to space limitations we only present the proofs of the main results. Detailed proofs can be found in [20].

II. Global allocation problem

We consider a static model in which a fixed number, N, of active elastic connections (*users*) share the capacities of L network links. A guaranteed Minimum Rate (MR), a Peak Rate (PR), and an assigned path characterize each connection. The network available bandwidths to be allocated are assumed to be non-time-varying. We introduce the following notations which will be used throughout the document: x has its value in $\mathcal{R}^N$ and stands for an instance of a user rate allocated vector, C is the vector of link available capacities, and A is a LxN incidence matrix. a_{lp} is equal to 1 if the link l belongs to the path of user p and 0 otherwise.

For simplicity and without loss of generality, we assume that on each link the spare capacity is strictly superior to the sum of the MR_i's ($i \in \{1, ..., N\}$) of the connections crossing this link. If this assumption is not valid then our results are still valid for the subset of connections to which we can allocate more than the corresponding minimum rates.

In [19], a game theoretic approach has been adopted to address the issue of efficient and fair allocation of available bandwidths between competing users (elastic connections). It has been argued that the Nash arbitration scheme idea gives rise to a suitable allocation policy. It has been shown that the Nash arbitrated solution, efficient and fair rate allocated vector, can be computed in a centralized way. With respect to our model and assumptions, it is the unique solution of the following global optimization problem, (S):

$$\begin{cases} Max_{\{x\}} \prod_{i=1}^{N} (x_i - MR_i) \\ x_i \geq MR_i \quad i \in \{1, ..., N\} \\ x_i \leq PR_i \quad i \in \{1, ..., N\} \\ (Ax)_l \leq (C)_l \quad l \in \{1, ..., L\} \end{cases}$$

We consider the convex problem (P) (primal problem) equivalent to (S) since they have the same optimal solution:

$$\begin{cases} Min_{\{x\}} GL(x) = -\sum_{i=1}^{N} \ln(x_i - MR_i) \\ x_i > MR_i \quad i \in \{1, ..., N\} \\ x_i \leq PR_i \quad i \in \{1, ..., N\} \\ (Ax)_l \leq (C)_l \quad l \in \{1, ..., L\} \end{cases}$$

Let X be a subset of $\mathcal{R}^N$ defined by connection bandwidth constraints and let $\mathcal{L}$ be the Lagrangian function associated with (P) and defined over $X \times \mathcal{R}^L$. X and $\mathcal{L}$ are defined as follows:

$$\begin{aligned} X = \{x = (x_1, ..., x_N) \in \mathcal{R}^N / x_i > MR_i \\ \text{and } x_i \leq PR_i\} \\ \mathcal{L}(x, \mu) = -\sum_{i=1}^{N} \ln(x_i - MR_i) \\ + \sum_{i=1}^{N} (\sum_{l=1}^{L} \mu_l \, a_{li}) \, x_i - \sum_{l=1}^{L} C_l \, \mu_l \end{aligned}$$

The dual function, d, corresponding to (P) is then defined on $\mathcal{R}^L$ as follows:

$$d(\mu) = Min_{x \in X} \mathcal{L}(x, \mu) \quad , \quad \mu \in \mathcal{R}^L \qquad \text{(II.1)}$$

Since problem II.1 is separable and has a unique solution, d can be computed explicitly. Indeed, for each $\mu \in \mathcal{R}^L$,

$$\begin{aligned} d(\mu) = \sum_{i=1}^{N} [-\ln(g_i(\sum_{l=1, a_{li}=1}^{L} \mu_l) - MR_i) \\ + (\sum_{l=1, a_{li}=1}^{L} \mu_l) \, g_i(\sum_{l=1, a_{li}=1}^{L} \mu_l)] - \sum_{l=1}^{L} C_l \mu_l \qquad \text{(II.2)} \end{aligned}$$

where for each $i \in \{1, ..., N\}$, g_i is defined on $\mathcal{R}$ as follows:

$$g_i(p) = \begin{cases} PR_i & \text{if } p \leq \frac{1}{PR_i - MR_i} \\ MR_i + \frac{1}{p} & \text{if } p \geq \frac{1}{PR_i - MR_i} \end{cases}$$

The dual problem (D) is the following:

$$Max_{\mu \in \mathcal{R}^L_+} \, d(\mu)$$

Since X is convex, GL is convex over X, and there exists $x \in X$ such that $(Ax)_l < C_l$ for each $l \in \{1, ..., L\}$, there exists a Lagrange multiplier and therefore there is no duality gap (see [17] chapter 5, for a general definition of a Lagrange multiplier). Also, (D) has at least one optimal solution.

Let $\bar{U}$ be the set of solutions of the dual problem. This set is also the set of Lagrange multipliers. $\bar{U}$ is nonempty and can be characterized in many ways ([17], chapter 5). The saddle-point characterization allows us to show that $\bar{U}$ is compact. We know from duality theory that d is concave on $\mathcal{R}^L$. One can show readily ([17], Danskin's theorem) that d is also continuously differentiable and the partial derivatives are determined as follows:

$$\frac{\partial d}{\partial \mu_l}(\mu) = \sum_{i=1, a_{li}=1}^{N} g_i(\sum_{l=1, a_{li}=1}^{L} \mu_l) \; - C_l$$
$$\mu \in \mathcal{R}^L; , \; l \in \{1, ..., L\}$$

We propose rate-based allocation schemes that aim at allocating badwidths to elastic connections according to the Nash arbitration scheme. For reliability as well as reducing communication overheads the scheme is aimed to be distributed among network nodes and sources. These are the basic issues which will be addressed in the sequel.

III. Optimization-based rate allocation algorithms

A. *Optimality-conditions-based algorithm*

We first begin by introducing the optimality-condition based algorithm which is related to the primal formulation.

The characterization of the Nash arbitrated allocation and the set of Lagrange multipliers or dual solutions ($\bar{U}$) is stated in this following proposition [17, Chapter 5]:

Proposition III.1: Let $\bar{x}$ be the unique solution of the centralized problem (S). Let $\bar{\mu}$ be a Lagrange multiplier for the primal problem (P). Then, the following optimality conditions *characterize* the pair $(\bar{x}, \bar{\mu})$:

- for each $i \in \{1, ..., N\}$, let $\Sigma_i = \sum_{l=1, a_{li}=1}^{L} \bar{\mu}_l$ then:

$$\bar{x}_i = \begin{cases} PR_i & \text{if } \Sigma_i \leq \frac{1}{PR_i - MR_i} \\ MR_i + \frac{1}{\Sigma_i} & \text{if } \Sigma_i \geq \frac{1}{PR_i - MR_i} \end{cases}$$

- for each $l \in \{1, ..., L\}$,
  1. $(A\bar{x} - C)_l \leq 0$
  2. $(A\bar{x} - C)_l \, \bar{\mu}_l = 0$
  3. $\bar{\mu}_l \geq 0$

Remark III.1: The Lagrange multipliers $\{\mu_l\}$ have the interpretation of being the implied (congestion) prices associated with the links. For a complete discussion see [19].

We present an algorithm in order to solve the optimization problem (S). The algorithm is based on the optimality conditions described in proposition III.1. As will be seen later in this section, the algorithm can be interpreted as a gradient-based algorithm.

The idea of using the optimality conditions to develop algorithms that solve an optimization problem is not new. For example, *F. Kelly* [14], [1], and C. Courcoubetis [15] have used this approach. We propose to develop the idea of *F. Kelly* ([14]) further since we propose extensions adapted to our allocation problem (S). Moreover, we propose a discrete-step, implementable, algorithm for which we prove the convergence.

Before presenting the algorithm we introduce some mathematical definitions and some relevant theory. Let f be a function defined on $\mathcal{R}^L$ and have its value in $\mathcal{R}^L$. f is equal to $(f_1, ..., f_L)$ where for each $l \in \{1, ..., L\}$, f_l is a real-valued function. As will be seen later, f depends on a positive parameter, ε, which is assumed to be constant in this section. Before giving an explicit expression for f we define the following functions.

$\mathbf{x}$ is an N-element-vector-valued function defined on $\mathcal{R}^L$. For each $i \in \{1, ..., N\}$,

$$\mathbf{x}_i(\mu) = g_i\Big(\sum_{l=1,\ a_{li}=1}^{L} \mu_l\Big) \qquad \text{(III.3)}$$

where g_i is a real-valued function defined in the previous section.

For each $l \in \{1, ..., L\}$ and $\varepsilon > 0$, define the real-valued function $\omega_l : \Re \rightarrow \Re$

$$\omega_l(p) = \begin{cases} C_l \left(\frac{p}{p+\varepsilon}\right)^2 & \text{if } p \geq 0 \\ 0 & \text{if } p \leq 0 \end{cases}$$

Define the functions $\{f_l\}$ as follows. For each $l \in \{1, ..., L\}$:

$$f_l(\mu) = \sum_{i=1,\ a_{li}=1}^{N} \mathbf{x}_i(\mu) - \omega_l(\mu_l) \ , \ for\ \mu \in \mathcal{R}^L \qquad \text{(III.4)}$$

Then we can show the following result which is stated without proof.

Lemma III.1: f is Lipschitz and bounded on $\mathcal{R}^L$. Let K_1 be the Lipschitz constant.

Then:

$$K_1 = \sqrt{L}\Big(\sum_{i=1}^{N}(PR_i - MR_i)^2\, N(i) + \frac{8}{27}\frac{C_{max}}{\varepsilon}\Big)$$

where C_{max} is the maximal available link capacity in the network, and for each $i \in \{1, ..., N\}$, $N(i)$ be the number of links crossed by user i

We now propose and study an implementable algorithm, $A1$, which allows us to approximate the efficient and fair allocation as characterized by the optimization problem (S).

By introducing a C^1 function V_1 it will be clear that the algorithm uses in fact steepest-descent method in order to minimize the function V_1 over $\mathcal{R}^L$. V_1 is a Lyapunov function for the system. Proposition III.2 gives a condition on the stepsize which guarantees the stability of algorithm $A1$.

Algorithm $A1$:

The algorithm is defined by a positive constant stepsize α and by the following difference equation system of dimension L:

$$\mu^{(k+1)} = \mu^{(k)} + \alpha\, f(\mu^{(k)}), \quad k \geq 0 \qquad \text{(III.5)}$$
$$\mu^{(0)} \in \mathcal{R}^L$$

Let V_1 be a real-valued function defined on $\mathcal{R}^L$ as follows:

$$V_1(\mu) = \sum_{k=1}^{L} \int_0^{\mu_k} \omega_k(p)dp - \sum_{i=1}^{N} \int_0^{\sum_{k=1:a_{ki}=1}^{L} \mu_k} g_i(p)dp \qquad \text{(III.6)}$$

It is easy to see that that V_1 is C^1 on $\mathcal{R}^L$. The first partial derivatives of V_1 can be easily obtained. Indeed, for any $\mu \in \mathcal{R}^L$ and $l \in \{1, ..., L\}$ we have:

$$\frac{\partial V_1}{\partial \mu_l}(\mu) = -f_l(\mu) \qquad \text{(III.7)}$$

It is clear from equation III.7 that algorithm $A1$ implements a steepest-descent technique for the minimization of V_1. Moreover it can be shown that $V1$ is convex and proper and hence has a unique global minimum, denoted by $\tilde{\mu}$, which belongs to $\mathcal{R}_+^{*\,L}$.

The following proposition deals with the convergence of $A1$.

Proposition III.2: Let $\{\mu^{(k)}\}$ a sequence generated by III.5 such that $\mu^{(0)} \in \mathcal{R}^L$ and $\alpha \in]0\,, \frac{2}{K_1}[$. Then $\{\mu^{(k)}\}$ converges to $\tilde{\mu}$, the unique global minimum of V_1.

Proof

Let $\mu^0 \in \mathcal{R}^L$ and $\alpha \in]0\,, \frac{2}{K_1}[$. We know that $\nabla V_1 = -f$ and f is K_1-Lipschitz. ∇V_1 is then K_1-Lipschitz and by the descent Lemma [17, Appendix A.24] the following holds:

$$V_1(\mu^{(k+1)}) \leq V_1(\mu^{(k)}) - (\alpha - \frac{K_1}{2}\alpha^2)\|\nabla V_1(\mu^{(k)})\|^2 \quad \text{(III.8)}$$

Since $\alpha \in]0, \frac{2}{K_1}[$, then $\{V_1(\mu^{(k)})\}$ is a decreasing sequence and it is convergent since it has a lower bound (V_1 has a minimum on $\mathcal{R}^L$). We can rewrite equation III.8 as follows:

$$0 \leq \alpha(1 - \frac{K_1}{2}\alpha)\|\nabla V_1(\mu^{(k)})\|^2 \leq V_1(\mu^{(k)}) - V_1(\mu^{(k+1)}) \quad \text{(III.9)}$$

Hence $\|\nabla V_1(\mu^{(k)})\| \to 0$ as k goes to ∞.

Since V_1 is convex on $\mathcal{R}^L$ and has a unique global minimum, see [17, Appendix B, Prop. B.9], we conclude that $\{\mu \in \mathcal{R}^L | V_1(\mu) \leq V_1(\mu^0)\}$ is compact. Therefore, $\{\mu^{(k)}\}$ is bounded and then it has at least one limit point (Weierstrass theorem). Let μ be one limit point. μ is then a stationary point of $V1$ (i.e $\nabla V_1(\tilde{\mu}) = \vec{0}$) since ∇V_1 is continous and $\lim_{k\to\infty} \nabla V_1(\mu^{(k)}) = \vec{0}$. On the other hand we know that V_1 has only one stationary point, which is $\tilde{\mu}$. So, $\{\mu^{(k)}\}$ has only one limit point, $\tilde{\mu}$, which is its limit. ■

Remark III.2: Actually, a stepsize in the interval $]0, \frac{2}{K_1}[$ guarantees a decrease of the cost function, V_1, at each iteration. Taking, $\alpha = \frac{1}{K_1}$ guarantees a good decrease at each iteration.

Remark III.3: Since $\mathbf{x}$ is a continuous function the sequence $\{\mathbf{x}(\mu^{(k)})\}$ converges to $\mathbf{x}(\tilde{\mu})$.

Validation

Algorithm $A1$ depends on $\varepsilon > 0$ which is used in the definition of $\omega(.)$. We show that as $\varepsilon \to 0$ the corresponding solution converges to the Nash arbitrated solution.

For $\varepsilon > 0$, let $\bar{\nu}(\varepsilon)$ defined over $\Re_+^*$ denote the unique global minimum of $V_1(\cdot, \varepsilon)$ where the explicit dependence of V_1 on ε is pointed out. Likewise, let $\bar{\mathbf{x}}(\varepsilon)$ denote $\mathbf{x}(\bar{\nu}(\varepsilon))$. Let $\bar{x}$ be the Nash arbitrated allocation vector (solution of (S)). Proposition III.4 shows that as ε approaches 0 $\bar{\nu}(\varepsilon)$ approaches $\bar{U}$, the set of dual solutions. Before stating the proposition we first state the following result:

Proposition III.3:

$$\lim_{\varepsilon\to 0^+} d(\bar{\nu}(\varepsilon)) = d(\bar{\mu}), \quad \forall \bar{\mu} \in \bar{U}$$

Proof

On $\mathcal{R}_+^L$ V_1 can be obtained explicitly as follows,

$$V_1(\mu,\varepsilon) = -d(\mu) - \sum_{i=1}^{N} \ln(PR_i - MR_i) - 2\varepsilon \sum_{l=1}^{L} C_l \ln(\mu_l + \varepsilon) + \varepsilon \sum_{l=1}^{L} \frac{C_l \mu_l}{\mu_l + \varepsilon} + (\sum_{l=1}^{L} \mu_l) 2\varepsilon \ln \varepsilon \qquad \mu \in \mathcal{R}_+^L,\ \varepsilon > 0$$

Let $\bar{\mu}$ be an element of $\bar{U}$, ε a positive real number, and $\bar{\nu}(\varepsilon)$ the global minimum of $V_1(\cdot, \varepsilon)$. Then:

$$|d(\bar{\nu}(\varepsilon)) - d(\bar{\mu})| \leq |V_1(\bar{\nu}(\varepsilon),\varepsilon) + d(\bar{\mu}) + \sum_{i=1}^{N} \ln(PR_i - MR_i)| + 2\varepsilon \ln \varepsilon (\sum_{l=1}^{L} \bar{\mu}_l(\varepsilon)) + 2\varepsilon C_{max} \sum_{l=1}^{L} |\ln(\bar{\nu}_l(\varepsilon) + \varepsilon)| + \varepsilon \sum_{l=1}^{L} C_l \quad \text{(III.10)}$$

Since $\forall \mu \in \mathcal{R}_+{}^L, V_1(\mu,\varepsilon) \geq V_1(\bar{\nu}(\varepsilon),\varepsilon)$ and $d(\mu) \leq d(\bar{\mu})$ we obtain:

$$V_1(\bar{\nu}(\varepsilon),\varepsilon) + d(\bar{\mu}) + \sum_{i=1}^{N} \ln(PR_i - MR_i) \geq -2\varepsilon \sum_{l=1}^{L} C_l \ln(\bar{\nu}_l(\varepsilon) + \varepsilon) \quad \text{(III.11)}$$

and

$$V_1(\bar{\nu}(\varepsilon),\varepsilon) + d(\bar{\mu}) + \sum_{i=1}^{N} \ln(PR_i - MR_i) \leq 2\varepsilon \ln \varepsilon (\sum_{l=1}^{L} \bar{\mu}_l) - 2(\sum_{l=1}^{L} C_l)\varepsilon \ln \varepsilon + \varepsilon(\sum_{l=1}^{L} C_l) \quad \text{(III.12)}$$

Let $\varepsilon_0 > 1$ and let $\varepsilon \in]0, \varepsilon_0]$ By equation III.12 we obtain:

$$V_1(\bar{\nu}(\varepsilon),\varepsilon)) \leq -d(\bar{\mu}) - \sum_{i=1}^{N} \ln(PR_i - MR_i) + 2\varepsilon_0 \ln \varepsilon_0 (\sum_{l=1}^{L} \bar{\mu}_l) - \frac{2}{e}(\sum_{l=1}^{L} C_l) + \varepsilon_0(\sum_{l=1}^{L} C_l) \quad \text{(III.13)}$$

Let $l \in \{1, ..., L\}$. We would like to show that there exists $M_l > 0$ such that for all $\varepsilon \in]0, \varepsilon_0]$ $\bar{\nu}_l(\varepsilon) \leq M_l$. Let us assume the contrary, i.e.:

$$\forall M > 0 \ \exists \varepsilon \in]0, \varepsilon_0] \ / \ \bar{\nu}_l(\varepsilon) > M$$

Then one can construct a positive real sequence $\{\bar{\nu}_l(\varepsilon_n)\}$ (strictly increasing) such that $\forall n \geq 0$, $\varepsilon_n \in]0, \varepsilon_0]$ and $\lim_{n\to\infty} \bar{\nu}_l(\varepsilon_n) = \infty$. Also the following inequality holds:

$$V_1(\bar{\nu}(\varepsilon_n),\varepsilon_n)) \geq m\|\bar{\nu}(\varepsilon_n)\|_1 - 2\varepsilon_n(\sum_{l=1}^{L} C_l) \ln(\|\bar{\nu}(\varepsilon_n)\|_1 + \varepsilon_n)$$

$$-N\ln(M\|\bar{\nu}(\varepsilon_n)\|_1+1)-\sum_{i=1}^{N}\frac{PR_i}{PR_i-MR_i}$$

where

$$m=\min_{k\in\{1,\ldots,L\}}(C_k-\sum_{i=1,\ a_{ki}=1}^{L} MR_i)$$

and

$$M=\max_{i\in\{1,\ldots,N\}}(PR_i-MR_i)$$

Since $\forall n\geq 0,\ \ \varepsilon_n\in]0,\varepsilon_0]$ and $\varepsilon_0>1$ then,

$$\begin{aligned}V_1(\bar{\nu}(\varepsilon_n),\varepsilon_n))&\geq m\|\bar{\nu}(\varepsilon_n)\|_1-\\&2\varepsilon_0(\sum_{l=1}^{L}C_l)\ln(\|\bar{\nu}(\varepsilon_n)\|_1+\varepsilon_0)-N\ln(M\|\bar{\nu}(\varepsilon_n)\|_1+1)\\&-\sum_{i=1}^{N}\frac{PR_i}{PR_i-MR_i}\end{aligned}\qquad\text{(III.14)}$$

From III.14 we conclude that $\lim_{n\to\infty}V_1(\bar{\nu}(\varepsilon_n),\varepsilon_n))=\infty$ since $\lim_{n\to\infty}\bar{\nu}_l(\varepsilon_n)=\infty$. This contradicts equation III.13. Hence we conclude that for each $l\in\in\{1,...,L\}$ there exists $M_l>0$ such that for all $\varepsilon\in]0,\varepsilon_0]\ \bar{\nu}_l(\varepsilon)\leq M_l$.

Having proved that, one can now easily show from equations III.11 and III.12 that $V_1(\bar{\nu}(\varepsilon),\varepsilon)+d(\bar{\mu})+\sum_{i=1}^{N}\ln(PR_i-MR_i)$ goes to 0 as ε goes to 0^+. And then that $\lim_{\varepsilon\to 0^+}|d(\bar{\nu}(\varepsilon))-d(\bar{\mu})|=0$, for each $\bar{\mu}$ in $\bar{U}$ (see equation III.10). ■

Proposition III.4:

$$\bar{\nu}(\varepsilon)\to\bar{U}\ \ as\ \varepsilon\to 0^+$$

Proof

Let $\varepsilon_0>0$ and $\varepsilon\in]0,\varepsilon_0]$. In the proof of the previous proposition we have shown that $\{\bar{\nu}(\varepsilon)\}_{\varepsilon\in]0,\varepsilon_0]}$ is bounded. Define the function λ as follows

$$\lambda(t)=\bar{\nu}(\frac{1}{t})\quad t\geq\frac{1}{\varepsilon_0}$$

Then, $\{\lambda(t)\}_{t\geq\frac{1}{\varepsilon_0}}$ is bounded. Let L^+ be the set of limit points of $\lambda(t)$. That is,

$$\begin{aligned}L^+\ &=\ \{\lambda\in\mathcal{R_+}^L\mid\exists\{t_n\}\ s.t\ \lim_{n\to\infty}t_n=\infty\\&\quad and\ \lim_{n\to\infty}\lambda(t_n)=\lambda\}\end{aligned}$$

From [16, Appendix A.2] , it follows that $\lambda(t)\to L^+\ \ as\ t\to\infty$. Let λ be an element of L^+. Then there exists a sequence $\{t_n\}$ such that $\forall n\geq 0\quad t_n\geq\frac{1}{\varepsilon_0}$ and $\lim_{n\to\infty}\lambda(t_n)=\lambda$. Since $\lambda(t_n)=\bar{\nu}(\frac{1}{t_n})$ and from Proposition III.3 we have the following:

$$\lim_{n\to\infty}d(\lambda(t_n))=d(\bar{\mu}),\quad\forall\bar{\mu}\in\bar{U}$$

Since d is continuous, $\lim_{n\to\infty}d(\lambda(t_n))=d(\lambda)$. Hence λ is an element of $\bar{U}$. As a result $L^+\subset\bar{U}$ and then $\lambda(t)\to\bar{U}\ \ as\ t\to\infty$. Hence, $\bar{\nu}(\varepsilon)\to\bar{U}\ \ as\ \varepsilon\to 0+$. ■

Finally we show that as $\varepsilon\to 0$, $\bar{x}(\varepsilon)$ converges to the Nash arbitrated allocation vector. This just follows from the continuity of $\bar{\mathbf{x}}(\varepsilon)$ and the above result.

Proposition III.5:

$$\lim_{\varepsilon\to 0+}\bar{\mathbf{x}}(\varepsilon)=\bar{x}$$

This completes our analysis of the first algorithm based on the optimality conditions.

B. Dual-based algorithm

We propose an algorithm that solves the dual problem (D) and which is based on a simple gradient-based projection method. This algorithm, $A2$, uses a constant stepsize. As will be shown later $\bar{U}$, the set of the solutions of the dual problem, is the set to which the algorithm converges. Moreover, the corresponding primal solutions converge to the unique Nash arbitrated allocated vector (see proposition III.7). The dual algortihm $A2$ is presented in the following.

Algorithm $A2$:

It is defined with a positive constant stepsize γ and by the following difference equation system of dimension L. For each $l\in\{1,...,L\}$, and $k\geq 0$

$$\mu_l^{(k+1)}=Max(0,\mu_l^{(k)}+\gamma(\sum_{i=1;a_{li}=1}^{N}\mathbf{x}_i(\mu^{(k)})\ -\ C_l))\qquad\text{(III.15)}$$

Where the functions $\{\mathbf{x}_i\}$ are defined as in the previous subsection and $\mu^{(0)}\in\mathcal{R_+}^L$.

The following lemma shows that the gradient of the dual function, d, is Lipschitz. Proposition III.6 gives a condition on the stepsize, γ, so that the algorithm converges to $\bar{U}$.

Lemma III.2: ∇d is Lipschitz on $\mathcal{R}^L$. Let K_2 be the Lipschitz constant. Let for each $i\in\{1,...,N\}$, $N(i)$ be the number of links crossed by user i. Then:

$$K_2=\sqrt{L}(\sum_{i=1}^{N}(PR_i-MR_i)^2N(i))$$

Proposition III.6: Let $\{\mu^{(k)}\}$ a sequence generated by III.15 such that $\mu^{(0)}\in\mathcal{R_+}^L$ and $\gamma\in]0\,,\,\frac{2}{K_2}[$. Then:

$$\mu^{(k)}\to\bar{U}\ \ as\ \ k\to\infty$$

Proof

Let $\gamma\ \in]0\,,\,\frac{2}{K_2}[$. d is C^1 over the closed and convex set $\mathcal{R_+}^L$. The gradient of d is K_2-Lipschitz. Hence via [17, Prop. 2.3.2, Chap. 2], every limit point of $\{\mu^{(k)}\}$ is an element of $\bar{U}$. We now show that the sequence $\{\mu^{(k)}\}$ is bounded. Since, d is concave on $\mathcal{R}^L$, $\mathcal{R}_+{}^L$ is a convex and closed set, and $\bar{U}$ nonempty and bounded the following set is compact: $\{\mu \in \mathcal{R}_+{}^L | -d(\mu) \leq -d(\mu^{(0)})\}$ (see [17, Appendix B, Prop. B.9]). Using the descent Lemma the projection characterization one can show that for each $k \geq 0$

$$-d(\mu^{(k+1)}) \leq -d(\mu^{(k)})$$

Hence, $\{\mu^{(k)}\}$ is bounded. Since the set of its limit points is included in $\bar{U}$, the result of the proposition follows. ■

Remark III.4: A stepsize in the interval $]0, \frac{2}{K_2}[$ guarantees an increase of the dual function, d, at each iteration.

Let $\bar{x}$ be the Nash arbitrated allocation vector (solution of (S)). Let for every $k \geq 0$, $x^{(k)}$ be equal to $\mathbf{x}(\mu^{(k)})$ where the sequence $\{\mu^{(k)}\}$ is generated by algorithm $A2$. Then the following holds.

Proposition III.7:

$$\lim_{k \to \infty} x^{(k)} = \bar{x}$$

Proof

It is very similar to the proof of proposition III.5. ■

Algorithms $A1$ and $A2$ suggest synchronous distributed implementations in which network nodes are synchronized (by iteration) and exchange information with users (connections). Indeed, iterations of type III.5 or III.15 can be run at each network node using local information (μ_l for link l) and information sent by relevant users. A user updates its local variable (x_i for user i) using information received from the links this user crosses (see III.3). After each update a user sends the new value to these links.

IV. Rate-based control schemes

We have proposed and studied two algorithms, $A1$ and $A2$, in order to solve the global optimization problem characterizing the Nash arbitrated allocation. These algorithms lend themselves to a distributed implementation in which network nodes and connections (or *sources*) play an active role. We propose two rate-based control schemes using explicit-rate-type notification ([12]) and motivated by the two optimization algorithms mentioned above.

The goal of the two rate-based control mechanisms is to allocate the available bandwidths inside the network between active elastic connections according to the fair criterion, the Nash arbitration scheme.

We first describe the rate control scheme based on algorithm $A2$, the other one being quite similar. We assume that elastic sources send regularly forward Resource Management (RM) packets in order to get feedback from the network about the congestion state or resource availability. In the context of the ATM ABR capacity transfer, a source sends an RM cell before a certain fixed number of data cells.

The information necessary for the operation of the control scheme is conveyed by the RM packets which are of two kinds; forward RM packets which are created by sources and conveyed along their corresponding paths and backward ones which are created by destinations that turn around the forward RM packets. The fields of an RM packet (figure 1) relevant to the description of the control scheme are "DIR" (direction: forward or backward), "MR" (connection minimum rate), "PR" (connection peak rate), "CP" (congestion price), and "ER" (explicit rate). "CP" is used by network nodes to communicate the value of the price variables (termed μ_l for link l previously) they control. "ER" stands for the maximal rate at which a given connection can transmit data.

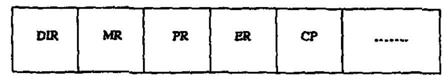

Fig. 1. An RM packet structure

There is a set of parameters associated with the control scheme: a constant stepsize, γ, used to update the price variables, some feedback intervals, FI, and some measurement intervals, MI. Each network link has its own feedback interval and measurement interval. A link price is updated at the beginning of each feedback interval and the total link input rate is measured during the measurement interval, figure 2.

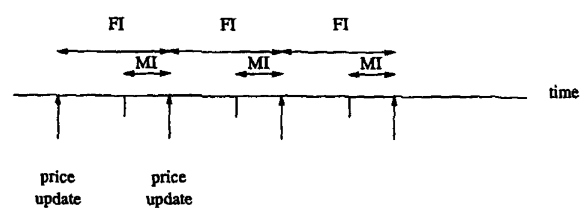

Fig. 2. A link updating and measuring process

The local procedures run by network nodes are based on the iterative algorithm $A2$. Indeed, if we interpret $\mathbf{x}_i(\mu^k)$ as the current data rate of connection i which is a function of the current network link price vector then in equation (III.15) the sum $\sum_{i=1,\ a_{li}=1}^{N} \mathbf{x}_i(\mu^{(k)})$ can be interpreted as the current total input rate at link l. It is important to note that the new price for a link l is computed when the information about current total input rate (the above sum) is available at the link. This helps to determine the right values for the feedback and measurement intervals associated with network links.

Now in the follwoing we describe the local procedures associated with the control scheme.

A source procedure:

• A source sends regularly a forward RM packet, and puts the minimum rate and the peak rate in the corresponding fields. Then, it sends the packet to the destination.

• At the reception of a backward RM packet a source adjusts its transmission data rate according to the explicit rate notification (ER) contained in the RM packet. This is done as follows: we consider that a source has a variable called "Allowed Rate" which is updated as follows: AR ← ER. AR is the maximal rate at which a source is allowed to transmit.

A destination procedure:

• At the reception of a forward RM packet, a destination creates a backward RM packet, puts zero in the "Congestion Price" field, and sends it back to its corresponding elastic source.

A network node procedure: *For a particular output link*

• At the beginning of each feedback interval (figure 2) the node updates the link *price* using the input rate measured during the previous measurement interval, Input, a constant stepsize, *Gamma*, and the link available capacity, C. The following illustrates the *price* updating:

$$price := \text{Max}\ (0\ ,\ price + Gamma\ (Input - C))$$

• At the reception of a backward RM packet, ER and CP are modifed using the current link price, the minimum rate (MR), and the peak rate (PR). The modifications are done as follows (ER is modified using the new value of CP):

$$CP := price + CP$$

$$ER := \begin{cases} PR & if\ \ CP \leq \frac{1}{PR-MR} \\ MR + \frac{1}{CP} & if\ \ CP \geq \frac{1}{PR-MR} \end{cases}$$

Once the modifications are completed the backward RM packet is relayed back to the source.

• A node measures regularily (figure 2) the total input rate at the link.

One can notice that the explicit rate (ER) contained in a backward RM packet does not increase when going through network nodes in the backward direction. In addition, the implementation of the rate control scheme does not differentiate between network access nodes and the other nodes as far as the update of ER is concerned.

For the good operation of the control scheme it is important to dimension for each link the feedback and measurement intervals. Indeed, FI should be large enough in order to allow the sources traversing a particular link to react to the new price (after update) conveyed by the backward RM packets and for a link to experience the result of sources reaction. The total input rate at a link should be measured during that period i.e when the response of sources to the new price has reached the link.

The second rate-based control scheme we propose differs only in the way the price variable (*price*) is updated. Indeed, for a link l it is updated using two constants, *Alpha* and ε, and according to the following (equation III.4):

$$price := price + Alpha\ (Input - \omega_l(price))$$

V. Concluding Remarks

Many issues concerning the good operation of the two rate control schemes have not been addressed in this paper. One of them is adaptivity which is basically the convergence of the control algorithm after a short transient period to the point at which sources get their Nash arbitrated share. One way is to use variable step sizes. A second issue is the issue of using network measurements, which introduce randomness for which stochastic stability will need to be studied. Robustness issues have also not been discussed which will be addressed in future work.

References

[1] J. Arrow and L. Hurwicz: *Decentralization and computation in resource allocation*. In *Essays in Economics and Econometrics*, University of North Carolina Press, 1960.

[2] ITU-T Recommendation I.371: *Traffic control and congestion control in B-ISDN*, Geneva, June 1996.

[3] D. Bertsekas and R. Gallager: *Data networks*. Prentice-Hall, 1987.

[4] J. Nash: *The bargaining problem*, Econometrica, Vol. 18, 1950, pp. 155-162.

[5] R. Mazumdar, L. Mason, and C. Douligeris: *Fairness in network optimal flow control: optimality of product forms*. IEEE Transactions on communications, vol. 39, no. 5, 1991, pp. 775-782.

[6] K. Bharathkumar and J.M. Jaffe: *A new approach to performance oriented flow control*. IEEE Trans. Comm., vol. COM-29, April 1981, pp. 427-435.

[7] L. Kleinrock *"On the modeling and analysis of computer networks"*. Proceedings of the IEEE, vol. 81, NO. 8, August 1993.

[8] P. Key: *Fixed-point models and congestion pricing for TCP and related schemes*, Workshop on Mathematical Modeling of TCP, Ecole Normale Superieure, Paris, Dec. 7-8, 1998. http:// wwww.dmi.ens.fr/ mistral/tcpworkshop.html

[9] S. Shenker: *"Fundamental design issues for the future Internet"*. IEEE Journal on Selected Areas in Communications, vol. 13, NO. 7, September 1995.

[10] R. Jain: *"Congestion control and traffic management in ATM networks: Recent advances and a survey*. Computer networks and ISDN systems, 1995.

[11] F. Bonomi and K. W. Fendick: *The rate-based flow control framework for available bit rate ATM services*, IEEE Network, March/April, 1995.

[12] European Transactions on Telecommunications, *Focus on Elastic Services over ATM networks*, R. Mazumdar and B. Doshi eds., ETT Vol. 8, N0. 1, 1997.

[13] S. H. Low and D. E. Lapsley: *Optimization flow control, I: Basic algorithm and convergence*. Submitted for publication, 1998.

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

[15] C. Courcoubetis, V. A. Siris, and G. D. Stamoulis: *Integration of pricing and flow control for available bit rate services in ATM networks*. Proc. IEEE Globecom'96, London.

[16] H. Khalil: *Nonlinear systems*. Prentice Hall. 1996.

[17] D. P. Bertsekas: *Nonlinear programming*. Athena Scientific, Belmont, Massachusetts, 1995.

[18] M. Minoux: *Mathematical programming: theory and algorithms*. Willey, Chichester, 1986.

[19] H. Yaiche, R. Mazumdar and C. Rosenberg: *A game theoretic framework for rate allocation and charging of Available Bit Rate (ABR) connections in ATM networks*. in Broadband Communications'98, P. Kuehn and R. Ulrich eds., Chapman and Hall, 1998, pp. 222-233.

[20] H. Yaiche, R. Mazumdar and C. Rosenberg: *Distributed algorithms for fair bandwidth allocation to elastic services*, preprint, May 1999.