Proceedings of the 17th International Symposium on Mathematical
Theory of Networks and Systems, Kyoto, Japan, July 24-28, 2006



# Scalable and Distributed Control Laws for Network Flow Optimization

Huigang Chen and John S. Baras
Institute for System Research
Department of Electrical & Computer Engineering
University of Maryland
College Park, Maryland 20742
Email: {huigang,baras}@isr.umd.edu

***Abstract*— We investigate the design of scalable and distributed control laws for flow control in a large network – equivalently an optimization problem. We identify the implications of the desired "plug-and-play" property of such protocols and propose design principles upon which our algorithm should be built: the class of our algorithms must achieve little extra communication cost and global asymptotic stability. General structural properties of such algorithms are presented and we finally provide an algorithm which satisfies our design principles for a network with heterogeneous delays.**

## I. Introduction

The design objective of congestion control algorithms of TCP type over the Internet [1] is to achieve efficient and fair usage of bandwidth for each user with limited information on the user's network environment. The necessity of the requirement for limited (or local) information is a natural consequence of the size of the system we deal with. By limited information we mean only information which can be measured or obtained by each user directly through her interaction only with the part of the network relevant to her flow. For example TCP operates explicitly on the knowledge of the losses of user's packets, which can be seen as a congestion message sent by the intermediate routers, and implicitly on the round-trip delay of each user's flow through a self-clock mechanism, which is a direct measurement by the end user. On the other hand efficiency and fairness are design goals which depend on various combinations of different flows, which are definitely non-local to each user per se. However, by adopting an optimization framework to interpret the efficient and fair bandwidth allocation [2]–[4], one can immediately reformulate the original large coupled problem into smaller decoupled problems via duality. In essence, each user tries to maximize her own utility function (induced by the fairness requirement) which is a function of her flow rate (primal variables of the optimization problem). At the same time congestion messages generated at each router by Active Queue Management (AQM) can be seen as dual variables (or Lagrange multipliers for the bandwidth constraints). Then the distributed algorithm is executed between all users and routers in the network through the exchange of primal and dual variables. The original design goal is translated into the ability of the distributed algorithm to reach the global optimal point eventually.

Given the separable nature of the network optimization problem in our context, an immediate candidate for distributed algorithms comes from dual gradient methods [3], [5], in which the dual variables are updated based on a gradient approach and the primal variables are obtained directly by solving the first-order optimality condition. This is generally termed "dual law" since only the calculation of the dual variables has dynamics. A variant of the dual law algorithm, in which primal variables are also updated according to a certain kind of dynamics, is called "primal/dual law" and this actually corresponds to the Lagrangian method in the theory of optimization [5]. The only equilibrium point of these two algorithms is the solution of the global optimization problem. In reality TCP with AQM, which has a pure integrator term, can be modelled as a primal/dual law algorithm. Furthermore, there is a class of "primal law" algorithms, which can model AQM with arbitrary random dropping functions, but in a strict sense those algorithms do not solve the network optimization problem, since their equilibrium points are not guaranteed to be the optimal solution, although they can be arbitrarily close to the optima [2].

A major cause of problems in the aforementioned distributed algorithms is the existence of delays in the network. Information obtained from the network in order to update primal or dual variables is usually subject to delays due to the time spent on computation, propagation, and queues. This information staleness is one of the major destabilizing factors for the algorithm dynamics and it is well known that TCP/AQM algorithms do not scale with large delay and bandwidth: they either result in low utilization of the network resources or display perpetual fluctuations of flow rates. Many research efforts have been devoted to this issue. First results on a scalable control law were proposed by Low and Paganini et al for a particular utility function [6], and subsequently they extended their result for general utility functions [7]. But both protocols are only verified (validated) for a linearized situation by Vinnicombe's results on TCP/AQM network control with heterogeneous delays [8]; however the global behavior results of their scalable control law are restricted to a single source/link network [9]. General approaches of global stability analysis include Lyapunov-Krasovskii methods, Lyapunov-Razumikhin methods [10], and contraction mapping methods. Stability conditions for the primal law and dual law algorithms are obtained by Fan et al [11] by employing a Razumikhin equivalent method, but their condition requires global information about the network. A contraction mapping method is used by Ranjan et al [12] to analyze a class of congestion control algorithms which enjoys stability with arbitrary large delays.

Our work intends to design a scalable and distributed control algorithm for the network flow optimization problem such that the algorithm has global stability and only requires local information for both users and routers. Specifically each user only needs to know the number of bottlenecks his flow traverses, the round-trip delay, and the aggregate congestion of his flow, and each router only needs to know the number of flows and the aggregate flow it has. In this way such a controller has a nice plug-and-play property which is desirable for actual implementation. The paper is organized as follows. In Section II we will present the network model and problem formulation. We will put forward our design principles there. In Section III the general properties that valid controllers must have are discussed based on some of our design principles. The scalable controller is then designed in Section IV and its global stability is proved. Final discussion and conclusions are given in Section V.

## II. Problem Formulation

The network considered in this paper is similar to the one in [7] and consists of $N$ users and $L$ bottleneck links (all those links whose bandwidths are fully utilized at equilibrium). We use the notation $[n]$ for the set $\{1, \cdots, n\}$ and the operator $|\cdot|$ for set cardinality. Therefore we have for the user set $[N]$ and for the bottleneck link set $[L]$. In reality network links other than bottleneck links may have effects on the dynamics of network flows. But for simplicity we only consider those bottleneck links, which we abbreviate as "links" hereafter. Each user $i$ has a fixed flow path $r_i \subset [L]$ to send a file with infinite length. In other words we only consider persistent flows. Also each router at link $j$ has a set $f_j \subset [N]$ of accessing flows. The routing matrix $R \in \{0,1\}^{L \times N}$ is defined as

$$R_{ji} \triangleq \begin{cases} 1, & j \in r_i, \\ 0, & j \notin r_i. \end{cases}$$

We denote by $x_i$ the flow rate of user $i$ and by $p_j$ the congestion information on link $j$. Due to the packet forward delay incurred during the transmission of flow packets, the aggregate flow rate seen by the router at link $j$ at time $t$ is

$$y_j(t) = \sum_{i \in f_j} x_i(t - \tau_{ij}^f) = R_{j\cdot} \begin{bmatrix} x_1(t - \tau_{1j}^f) \\ \vdots \\ x_N(t - \tau_{Nj}^f \end{bmatrix} \qquad (1)$$

where $\tau_{ij}^f$ is the forward delay from user $i$ on link $j$. Accordingly, the aggregate congestion information received by each user $i$ at time $t$ is

$$q_i(t) = \sum_{j \in r_j} p_i(t - \tau_{ij}^b) = [p_1(t - \tau_{i1}^b), \cdots, p_L(t - \tau_{iL}^b)] R_{\cdot i} \qquad (2)$$

where $\tau_{ij}^b$ is the backward delay from link $j$ to user $i$. For reasons of fast computation and small communication cost, routers cannot differentiate individual flows and users cannot differentiate congestion levels of individual links. All they have access to are aggregate information and we will show that these are actually sufficient for our purposes.

An important assumption made now is that both forward delays and backward delays are time invariant, which is a valid approximation when routers have small buffers compared to the product of bandwidth and propagation delays. Then the observation that $\tau_{ij}^b \geq \tau_{ij}^f$ usually holds if the reverse route is symmetric with respect to the forward route. We also use the following definition

$$\tau_i \triangleq \tau_{ij}^f + \tau_{ij}^b, \forall j \in r_i, \qquad (3)$$

which is the round-trip delay of flow $i$. Again it consumes extra communication bits to accumulate information about forward delays and backward delays separately, and in contrast it is straightforward for each user to measure the round-trip delay. Therefore it is much more desirable to design algorithms whose parameters depend not on the forward/backward delays separately but only on the round-trip delays.

As mentioned in Section I the problem of efficient and fair allocation of network bandwidths can be cast into the problem of network optimization over flows. This optimization problem is a classical convex programming problem with linear constraints:

$$\begin{array}{l} \max_{x_i \geq 0} \sum_{i \in [N]} U_i(x_i) \\ \text{s.t. } Rx \leq c. \end{array} \qquad (4)$$

where each function $U_i : \mathbb{R}_+ \to \mathbb{R}$, which is understood as the utility function associated with user $i$, is a strictly concave, continuously differentiable nondecreasing function and $c$ is a $L$-dimensional vector whose $j$th component represents the bandwidth of link $j$. As usual we assume that $U_i'(x) \to \infty$ as $x \to 0$. The relation between the role of utility functions and fairness criteria has been clarified by [2], [4]: it turns out that many practical concepts of fairness are equivalent to the right selection of utility functions. As a consequence of our assumptions the network optimization problem (4) has a unique solution at which all the constraints are satisfied with equality, i.e. we attain efficient usage of network resources. We use the notation $\cdot^*$ to denote the equilibrium value from (or induced by) the network optimization problem, for example $p_j^*$ is the equilibrium congestion information on link $j$.

The standard approach to solve this global optimization problem (4) in a distributed manner is to solve the dual problem instead:

$$\min_{p_j\geq 0}\sum_{i\in[N]}\max_{x_i\geq 0}(U_i(x_i)-x_iR_{\cdot i}p)+p^Tc. \tag{5}$$

This process decouples the coupling of the primal variables through the constraints of the original optimization problem and turns it into many small maximization problems, each of which can be handled by users with local information. The main algorithms derived from the gradient method and the Lagrangian method can be written in general form as shown below,

$$\text{Dual Law:}\begin{cases} x_i(t)=U_i'^{-1}(q_i(t)),\\ \dot{p}_j(t)=\Gamma_j(y_j(t)-c_j).\end{cases} \tag{6}$$

$$\text{Primal/Dual Law:}\begin{cases} \dot{x}_i(t)=K_i(U_i'(x(t))-q_i(t))\\ \dot{p}_j(t)=\Gamma_j(y_j(t)-c_j).\end{cases} \tag{7}$$

It is known that without delays both algorithms (6) and (7) converge to the optimal solution with any positive coefficients $K_i$s and $\Gamma_j$s [3]. When there are delays involved, global stability analysis of both algorithms in a heterogeneous network reveals that the stability condition depends on those coefficients in a complicated way. Although decentralized protocols exist in order to satisfy these stability conditions, they require extra communication costs and most importantly users have to reveal their own utility functions. This is illustrated by the following simple example:

*Example 1:* A single source/link network uses the following primal/dual algorithm for its flow control

$$\begin{aligned}\dot{x}(t)&=K(U'(x)-p(t-\tau)),\\ \dot{p}(t)&=\Gamma(x-c).\end{aligned}$$

By simple analysis it is required that $-U''(c)>2\tau\Gamma/\pi$ for the existence of a coefficient $K$ so that the system is locally stable. In order to set the right $\Gamma$ the router has to know the user's utility function. But it is clearly undesirable to transmit a function across the communication links, not to mention security reasons. When the user knows only her $U(\cdot)$ and the link knows its $c$, no one can calculate $U''(c)$.

Therefore we propose the following necessary design principles for our algorithms in order to meet the needs of real-world networks

1) Equilibrium of the algorithm should solve the optimization problem (4);
2) The input and the parameters of user and link controllers should be obtained from local information only. For an individual user the local information is that which is accrued along the path of his flow, and for an individual link the local information is that which is aggregated from its accessing flows. Additionally each user's utility function should be only known to himself and each link's bandwidth should also be kept to itself.
3) The dynamics of the algorithm are globally asymptotically stable given heterogeneous delays.

The scalable control laws by Paganini et al [7] satisfy Principles 1)-2) and partially 3) since only linear stability is verified for their algorithm. Their algorithm takes the following form in which $\zeta_i$ is an auxilary state variable at the user $i$'s side:

$$\begin{aligned}\tau_i\dot{\zeta}_i&=\beta_i(U_i'(x_i)-q_i),\\ x_i&=\bar{x}_ie^{\zeta_i-\frac{\alpha_iq_i}{|r_i|\tau_i}},\\ \dot{p}_j&=c_j^{-1}(y_j-c_j).\end{aligned} \tag{8}$$

Strictly speaking their protocol is not completely decentralized as defined in Principle 2), because their controller parameters depend on a global variable $\bar{\tau}$, which is the delay upper-bound of the whole network. Specifically in order to achieve linear stability, the following condition has to be satisfied,

$$\frac{\beta_i|r_i|}{\alpha_i}\bar{\tau}<\eta$$

for some constant $\eta$. Although this restriction might not seem to be significant, the future growth of the network may potentially require a global reset of user control coefficients and furthermore the existence of this condition on a global variable may intuitively result in slow performance due to its conservativeness. Therefore we aim at designing algorithms strictly satisfying the proposed Principles 1)-3).

## III. General Properties of Controllers

Before we start to design a specific distributed algorithm which satisfies all the principles introduced in the previous section, we first want to understand the structural implications of controllers based on Principles 1) and 2. The reason for investigating these principles first is that to some extent they reflect the "static" characteristics our controller must possess, while Principle 3) is more relevant to its "dynamical" characteristics. It is quite difficult to make a statement about the general properties of such controllers since the controller space is a very large functional space. Therefore we resort to focus on the linearized version of both user and link controllers and the results from the linearized controllers will give us necessary conditions as well as design guidance for the full-blown nonlinear controllers in the next section. For our purpose we only consider controllers which allow a unique euilibrium state in this section.

Again consider a single user/link network with the round-trip delay $\tau$ and let $F(s)(G(s))$ be the transfer function of the user (link) controller with congestion message as the input (output) signal and flow rate as the output (input) signal. Suppose both $F(s)$ and $G(s)$ are proper rational functions. Here we made another assumption that the user (link) dynamics do not explicitly depend on her delayed value of flow rate (congestion message). This is a valid assumption since in our problem formulation delays do not bring any benefits to our goals. Then the open loop gain of the system is $e^{-\tau s}G(s)F(s)$. First we give the condition for the user controller:

*Proposition 1:* Assume that the user dynamics (by themselves) do not involve any delays. Then the transfer function $F(s)$ of the linearized user controller is a valid

TABLE I

USER CONTROLLERS AND THEIR TRANSFER FUNCTIONS

| User Controller | Transfer Function |
|---|---|
| $x = U'^{-1}(q)$ | $1/\xi$ |
| $\dot{x} = K(U'(x) - q)$ | $K/(s + K\xi)$ |
| user controller by Paganini et al (8) | $K(s+v)/(s+Kv\xi)$ |

user controller for the optimization problem if and only if it is stable and $F(0) = \xi^{-1}$ where $\xi = -U''(x^*)$.

*Proof:* By definition we have $\delta x(s) = F(s)\delta q(s)$ where $\delta x = x - x^*$ and $\delta q = q - q^*$. Since in equilibrium $U'(x^*) = q^*$, we have $U'(x^* + \delta x) = U'(x^*) + U''(x^*)\delta x = q^* - \delta q$. Hence we conclude that $F(0) = \xi^{-1}$ is a necessary condition for $F(s)$ to be valid. We shall show that the condition is also sufficient. Without loss of generality assume that the user controller is strictly proper and has the following form

$$F(s) = \frac{a_1 s^{n-1} + a_2 s^{n-2} + \cdots + a_{n-1}s + a_n}{s^n + b_1 s^{n-1} + \cdots + b_{n-1}s + \xi a_n}.$$

Here $a_1, \cdots, a_n$ and $b_1, \cdots, b_{n-1}$ can be functions of $\xi$. It is well known that this transfer function can be realized in a controller canonical form [13]

$$\begin{aligned} \begin{bmatrix} \dot{z}_1 \\ \vdots \\ \dot{z}_n \end{bmatrix} &= \begin{bmatrix} -b_1 & \cdots & -b_{n-1} & -\xi a_n \\ 1 & 0 & \cdots & 0 \\ 0 & \ddots & 0 & \vdots \\ 0 & \cdots & 1 & 0 \end{bmatrix} \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix} \\ & \quad -[a_n, 0, \cdots, 0]^T \delta q \\ \delta x &= \tfrac{a_1}{a_n} z_1 + \cdots + \tfrac{a_{n-1}}{a_n} z_{n-1} + z_n. \end{aligned}$$

This is a local version of the following nonlinear dynamics:

$$\begin{aligned} \dot{z}_1 &= -b_1(-U''(z_n))z_1 - \cdots - b_{n-1}(-U''(z_n))z_{n-1} \\ & \quad + a_n(-U''(z_n))(U'(z_n) - q), \\ \dot{z}_k &= z_{k-1}, \quad 2 \le k \le n, \\ x &= \tfrac{a_1(-U''(z_n))}{a_n(-U''(z_n))} z_1 + \cdots + \tfrac{a_{n-1}(-U''(z_n))}{a_n(-U''(z_n))} z_{n-1} + z_n \end{aligned}$$

The result then can be easily verified from the fact that the equilibrium of the system when the input is $q^*$ is indeed $y_k = 0, 1 \le k \le n-1$ and $x = y_n = x^*$. ∎

As an illustration of this proposition we can observe the correspondence between previously proposed valid user controllers and their linearized forms in Table I.

Now we turn to the properties of valid link controllers.

*Proposition 2:* Assume that, like the user dynamics, the link dynamics by themselves do not involve any delays. Then the transfer function $G(s)$ of the linearized link controller is a valid link controller for the optimization problem if and only if it is stable and $G(s) = H(s)/s$ in an irreducible form where $H(s)$ is some rational transfer function.

*Proof:* First we verify the sufficiency part. Suppose the equilibrium point with $F(s)$ in Proposition 1 and $G(s)$ given under the current form is the optimal solution of the network optimization problem. From the definition we have $\delta p(s) = G(s)\delta y(s)$ where $\delta p = p - p^*$ and $\delta y = y - y^*$. The link controller can be realized in the following form,

$$\begin{aligned} \dot{u} &= y - c \\ \dot{v} &= Av + Bu \\ p &= Cv + Du \end{aligned}$$

where $(A, B, C, D)$ is a realization of the transfer function $H(s)$. Since $G(s)$ contains a pure integrator, the only input that achieves the internal stability is $\delta y = 0$, or in the realized system $y^* = c$. Along with the source controller we have the equations for the equilibrium state

$$\begin{aligned} U'(x^*) &= q^* = p^* \\ y^* &= x^* = c \end{aligned}$$

By the KKT conditions this equilibrium point is the optimal solution of the network optimization problem.

Next we show that this integrator form is also necessary. First since only the link knows its own bandwidth $c$ and the equilibrium point has to be $y^* = c$ for optimality, only the link controller can enforce the input $\delta y$ to be zero at equilibrium. Suppose the link controller is realized as shown below,

$$\begin{aligned} \dot{z} &= Az + B\delta y \\ \delta p &= Cz + D\delta y. \end{aligned}$$

The previous argument is equivalent to the condition $\text{rank}A < \text{rank}[A, B]$. It is sufficient to check the situation when $(A, C)$ is observable (since $(A, C)$ has to be detectable for stabilization, therefore the unobservable modes are asymptotically stable themselves regardless of input, so we only focus on the observable part). By a similarity transformation we can write the system in canonical observer form [13] as follows

$$\begin{aligned} A &= \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 1 & \cdots & 0 & a_1 \\ 0 & \ddots & 0 & a_2 \\ 0 & \cdots & 1 & a_{n-1} \end{bmatrix}, \\ B &= [b_1, \cdots, b_n]^T, \\ C &= [0, \cdots, 0, 1]. \end{aligned}$$

Here $A_{1n} = 0$ and $b_1 \neq 0$ due to the rank condition. Then it is straightforward to see that $G(s)$ must have a pure integrator term. ∎

*Remark 1:* The structural properties of valid user and link controllers indicated in the previous two propositions suggest that delay independent stability [12] may not be achievable given our design principles. To see this let us observe now that the open loop gain of a single user/link network can be written as $e^{-\tau s}H(s)F(s)/s$ where $F(0) = \xi^{-1}$ and $H(0) = h$ for some nonzero $h$. If the system is delay independent stable, then the Nyquist curve of its open loop gain should intersect the $x$-axis at points greater than -1 regardless of the value of $\tau$. But it is easy to see that for sufficiently large $\tau$, the Nyquist curve intersects the $x$-axis at the frequency $\omega \approx \frac{\pi}{2\tau}$ and the intersection point is approximatedly $-\frac{2h}{\pi\xi}\tau$ which can be made arbitrary smaller than -1. Therefore in order to achieve stability one must design the controllers based on the size of delays.

## IV. Design of Scalable Controller

We first focus on the design of scalable controllers for a single user/link network based on previous discussions and then extend the design to arbitrary networks with heterogeneous delays.

### A. The Case of Single User/Link Network

As in previous sections we denote by $c$ the link bandwidth and by $\tau = \tau^f + \tau^b$ the round-trip delay. Similar to the user controller in Paganini et al's algorithm (8), we can choose the transfer function of our user controller to be

$$F(s) = \frac{s + k/\tau}{\tau s + \xi k/\tau} \quad (9)$$

and our link controller to be

$$G(s) = \frac{1}{s}. \quad (10)$$

Here $\xi$ is defined as $-U''(c)$ as in Proposition 1 and $k$ is some constant. First by direct calculation we have

*Lemma 1:* The single user/link network with the user controller given by (9) and the link controller (10) is linearly asymptotically stable for arbitrary $\tau$ and $c$ if $0 < k \leq k_0 \approx 0.5474$. Here $k_0 \triangleq \omega_0/\tan\omega_0$ where $\omega_0 \in (0, \pi/2)$ is the solution of the equation $\omega \sin\omega = 1$.

From the proof of Proposition 1 we can realize our algorithm from its linearized form as follows:

$$\begin{aligned} \dot{z} &= \tfrac{k}{\tau^2}(U'(x) - p(t-\tau^b)), \\ x &= z - \tfrac{p(t-\tau^b)}{\tau}, \\ \dot{p} &= x(t-\tau^f) - c. \end{aligned} \quad (11)$$

*Remark 2:* It is worth discussing the initial dynamics of the above system. Since there is no guarantee at the beginning from $x = z - p(t-\tau^b)/\tau$ such that $x$ is kept positive, we have to resort to other means. A feasible solution to the initial dynamics is as follows,

$$\begin{cases} x(t) = z - p(t-\tau^b)/\tau, & \text{if } z > p(t-\tau^b)/\tau, \\ \dot{x}(t) = -\alpha x(t), & \text{otherwise}, \end{cases}$$

for any positive constant $\alpha$. Since from our dynamics (11) $p(t)$ is a continuous function of time, it is easy to see that once $x(t) > 0$, it stays positive thereafter. So the dynamics of $x$ will be of the form $\dot{x} = -\alpha x$ for at most a finite time duration at the beginning of the algorithm. This period can be regarded as a "probing" phase of the flow dynamics. Therefore it is sufficient for us to consider only the dynamics (11) thereafter.

The global stability of the system (11) can be studied from the observation that the system is actually of Lur'e type [14] by rewriting it into an equivalent form as follows

$$\begin{aligned} \dot{x} &= -\tfrac{k}{\tau^2}(p(t-\tau^b) - p^*) - \tfrac{1}{\tau}(x(t-\tau) - x^*) + \tfrac{k}{\tau^2}u, \\ \dot{p} &= x(t-\tau^f) - c, \\ u &= U'(x) - p^*. \end{aligned}$$

Taking $u$ as the input signal and $x$ as the output signal, the transfer function from $u$ to $x$ is

$$L(s) = \frac{k}{\tau^2}\left(s + \frac{k+\tau s}{\tau^2 s}e^{-\tau s}\right)^{-1},$$

while the mapping from $x$ to $u$ is a $(0,\infty)$-sector nonlinear mapping. In order to obtain nonlinear stability of (11) by Popov's criterion [15], [16] it remains to show that there exists $\eta \in \mathbb{R}$ such that $(1+\eta s)L(s)$ is positive real.

*Lemma 2:* $(1+\tau s/2)L(s)$ is positive real when $0 < k \leq 1/2$.

*Proof:* By Lemma 1 we only need to check whether $\mathrm{Re}(1+\tau i\omega/2)L(i\omega) \geq 0$ and this in turn is equivalent to whether $\mathrm{Re}(1+\tau i\omega/2)^{-1}L(i\omega)^{-1} \geq 0$. Hence the proof reduces to showing that

$$\frac{1}{2}\theta(\theta^2 - k\cos\theta - \theta\sin\theta) - k\sin\theta + \theta\cos\theta \geq 0 \quad (12)$$

where $\theta \triangleq \omega\tau$.

When $k = 1/2$ the above inequality (12) becomes

$$\frac{1}{2}\theta(\theta^2 - \frac{1}{2}\cos\theta - \theta\sin\theta) - \frac{1}{2}\sin\theta + \theta\cos\theta \geq 0 \quad (13)$$

which is correct by checking it with numerical means.

If $0 \leq \theta \leq \theta_0 \approx 2.2889$, in which $\theta_0$ is the smallest positive solution of the equation $\theta\cos\theta + 2\sin\theta = 0$, we have

$$\frac{1}{2}\theta\cos\theta + \sin\theta \geq 0.$$

But the left hand side of the above inequality is exactly the difference of the left hand sides of the inequalities (12) and (13) times $\frac{1}{2} - k$. Thus the inequality (12) holds for $0 \leq \theta \leq \theta_0$. Now consider the situation when $\theta > \theta_0$. In this case the left hand side of (12) is lower bounded by

$$\frac{1}{2}\theta(\theta^2 - \frac{1}{2} - \theta) - \theta - \frac{1}{2}.$$

One can directly check that this cubic polynomial achieves its minimum over $\theta \geq \theta_0$ at $\theta = \theta_0$, and that the minimum is positive. Therefore we conclude that the inequality (12) holds for all $\theta$ and $(1+\tau s/2)L(s)$ is positive real. ■

Then from Lemma 2 and Popov's criterion we immediately have:

*Proposition 3:* With the initial dynamics discussed in Remark 2, the system (11) is globally asymptotically stable for arbitrary values of $\tau$ and $c$ if $k \in (0, 1/2]$.

So we obtain a scalable controller which satisfies all the design principles in Section II for a single user/link network.

### B. The Case of General Network

A direct extension of the user and link controllers (9-10) from the previous subsection to the situation of a general network with heterogeneous delays is

$$F_i(s) = \frac{s + k/\tau_i}{\tau_i|r_i|s + \xi_i k/\tau_i} \quad (14)$$

for user $i$ and

$$G_j(s) = \frac{1}{|f_j|s} \quad (15)$$

for link $j$.

Define a $L \times N$ matrix-valued function $\hat{R}(s)$ on the frequency domain by

$$\hat{R}_{ji}(s) = R_{ji} e^{-\tau^f_{ij} s}$$

then the relation between the flow rate vector $x$ and the aggregate rate vector $y$ (1) can be written as

$$y(s) = \hat{R}(s)x(s).$$

From the definition of the round-trip delays (3), the relation between the congestion message vector $p$ and the aggregate congestion vector $q$ (2) can be equivalently expressed by

$$q(s) = \text{diag}(\{e^{-\tau_i s}\})\hat{R}^H(s)p(s).$$

where $\hat{R}^H$ is the Hermitian of $\hat{R}$.

Therefore combining these equations the open loop gain of the network system with tentative controllers (14-15) is given as follows

$$\begin{aligned} L(s) &= \text{diag}\left(\left\{\frac{s+k/\tau_i}{\tau_i|r_i|s+\xi_i k/\tau_i} e^{-\tau_i s}\right\}\right) \\ &\quad \times \hat{R}^H(s)\text{diag}\left(\left\{\frac{1}{|f_j|s}\right\}\right)\hat{R}(s). \end{aligned}$$

It would be desirable that this natural extension from the single user/link network (14-15) simply gives us stabilizing controllers for general networks. To examine this we need to study the eigenloci of the matrix $L(s)$, per the Generalized Nyquist Theorem [17]. Recall an elegant result by Vinnicombe [8]:

*Lemma 3 (Vinnicombe):* Assume $\Lambda = \text{diag}(\{\lambda_i\})$ and $M = M^T \geq 0$ are $N \times N$ matrices. Then the eigenvalues of $\Lambda M$ $\sigma(\Lambda M) \in \rho(M)\bar{\text{co}}(\{0, \lambda_1, \cdots, \lambda_N\})$. Here $\rho(\cdot)$ denotes the spectral radius and $\bar{\text{co}}(\cdot)$ denotes the convex hull.

Note that

$$\sigma(L(s)) = \sigma\left(\text{diag}(\{l_i(s)\})M(s)\right),$$

where we define

$$\begin{aligned} l_i(s) &\triangleq \frac{s + k/\tau_i}{s(\tau_i s + \xi_i k/(\tau_i|r_i|))} e^{-\tau_i s}, \\ M(s) &\triangleq \text{diag}(\{|r_i|^{-1/2}\})\hat{R}^H(s)\text{diag}(\{|f_j|^{-1}\}) \\ &\quad \times \hat{R}(s)\text{diag}(\{|r_i|^{-1/2}\}). \end{aligned} \tag{16}$$

We first calculate the upper bound of the spectral radius $\rho(M(i\omega))$ of $M(i\omega)$:

$$\begin{aligned} &\rho(M(i\omega)) \\ =\; &\rho(\text{diag}(\{|r_i|^{-1}\})\hat{R}^H(i\omega)\text{diag}(\{|f_j|^{-1}\})\hat{R}(i\omega)) \\ \leq\; &\sup_{i,\omega} \sum_{j,n} |r_i|^{-1}|\hat{R}_{ji}(i\omega)||f_j|^{-1}|\hat{R}_{jn}(i\omega)| \\ =\; &1 \end{aligned}$$

from the definitions of $|r_i|$ and $|f_j|$. Therefore a sufficient condition for linear stability is

$$-1 \notin \bar{\text{co}}(\{0, l_1(i\omega), \cdots, l_N(i\omega)\})$$

by Lemma 3 and the Generalized Nyquist Theorem. Since the Nyquist curves $l_i(i\omega)$ with arbitrary $\tau_i$ and $\xi_i$ are bounded by a single curve $l(\theta)$ on the Nyquist plane:

$$l(\theta) \triangleq -\frac{i\theta + k}{\theta^2} e^{-i\theta},$$

we only need to check whether

$$-1 \notin \bar{\text{co}}(0 \cup \{l(\theta), \forall\theta \geq 0\}).$$

However since $\angle l(\theta) \to -180^\circ$ with $\theta \to 0$ and part of the curve $l(\theta)$ lies on the second quadrant of the Nyquist plane, the convex hull of curve $l(\theta)$ contains -1. Therefore we cannot obtain linear stability[1] with controllers (14-15) and this is the main reason why the controllers (8) proposed by Paganini et al have to rely on a global variable $\bar{\tau}$.

Therefore we propose our scalable user controller, which is a modification of (14), as follows,

$$T_i(s) = \frac{s + \frac{\min\{\tau_i, 1\}}{2\tau_i}}{2\tau_i \max\{\tau_i, 1\}|r_i|s + \frac{\xi_i \min\{\tau_i, 1\}}{2\tau_i}}, \tag{17}$$

and together with our original link controller (15), we prove below that they are valid stabilizing linear controllers for our optimization problem (4).

*Lemma 4:* The flow dynamics of a network with heterogeneous delays where each user controller is given by (17) and each link controller is given by (15) are linearly asymptotically stable.

*Proof:* See Appendix. ■

By Proposition 1 one can realize the linear controllers (17) and (15) by

$$\begin{aligned} \dot{z}_i &= \frac{\min\{\tau_i, 1\}}{4\tau_i^2|r_i|\max\{\tau_i, 1\}}(U_i'(x_i) - q_i), \\ x_i &= z - \frac{q_i}{2\tau_i|r_i|\max\{\tau_i, 1\}}, \\ \dot{p}_j &= \frac{y_j - c_j}{|f_j|}. \end{aligned} \tag{18}$$

We also assume that for the above system we adopt initial dynamics similar to that in Remark 2, so that after an initial phase the system stays in the correct region of $x_i > 0$ with the above dynamics forever.

Just like the case of single user/link network, we can rewrite the system (18) into Lur'e form and the condition for global asymptotic stability is equivalent to the positive realness of the following matrix-valued function

$$\begin{aligned} &W(s) \\ \triangleq\; &(I + \text{diag}(\{\eta_i s\})) \\ &\times \begin{pmatrix} sI + \text{diag}\left(\left\{\frac{e^{-\tau_i s}}{\tau_i^2|r_i|s}\frac{\tau_i s + \min\{\tau_i, 1\}/2}{2\max\{\tau_i, 1\}}\right\}\right) \\ \times\hat{R}^H(s)\text{diag}(\{|f_j|^{-1}\})\hat{R}(s) \end{pmatrix}^{-1} \\ &\times \text{diag}\left(\left\{\frac{\min\{\tau_i, 1\}}{4\tau_i^2|r_i|\max\{\tau_i, 1\}}\right\}\right) \end{aligned}$$

for some real-valued $\{\eta_i\}$.

*Lemma 5:* The matrix-valued function $W(s)$ defined above is positive real for $\eta_i = \tau_i/2$.

[1]One may wonder whether the inability to prove the stability in our analysis is due to some conservativeness of Lemma 3. In fact we are able to construct a 7-user 5-link network with controllers (14-15) such that the flow dynamics of the network is unstable.

*Proof:* Since we already have linear stability from Lemma 4 it suffices to show that the following matrix-valued function is positive real

$$\begin{aligned}\tilde{W}(s) \triangleq\ & \text{diag}\left(\left\{\frac{\tau_i s}{1+\tau_i s/2}\right\}\right) I \\ & +\text{diag}\left(\left\{\frac{e^{-\tau_i s}}{\tau_i s(1+\tau_i s/2)}\frac{\tau_i s+\min\{\tau_i,1\}/2}{2\max\{\tau_i,1\}}\right\}\right) M(s).\end{aligned}$$

where $M(s)$ is defined in (16) and recall that $\rho(M(i\omega)) \leq 1$.

Since $M(i\omega)$ is positive semidefinite Hermitian, there exists a unitary matrix $U(\omega)$ such that

$$M(i\omega) = U^H(\omega)\text{diag}(\{\lambda_i(\omega)\})U$$

where $\{\lambda_i(\omega)\}$ are eigenvalues of $M(i\omega)$ which are real and satisfy $\lambda_i(\omega) \leq 1$. Therefore we can rewrite $\tilde{W}$ as

$$\begin{aligned}\tilde{W}(i\omega) =\ & \text{diag}(\{w_i(i\omega)\})U^H\text{diag}(\{\lambda_i\})U \\ & +\text{diag}\left(\left\{\frac{i\omega\tau_i}{1+i\omega\tau_i/2}\right\}\right) U^H\text{diag}(\{1-\lambda_i\})U,\end{aligned}$$

where

$$w_i(s) \triangleq \frac{\tau_i s}{1+\tau_i s/2} + \frac{e^{-\tau_i s}}{\tau_i s(1+\tau_i s/2)}\frac{\tau_i s+\min\{\tau_i,1\}/2}{2\max\{\tau_i,1\}}.$$

In fact the proof of Lemma 2 implies that $w_i(s)$ is positive real. Also it is easy to see that the function $\tau s/(1+\tau s/2)$ is positive real. By positive semidefiniteness of the matrices $U^H\text{diag}(\{\lambda_i\})U$ and $U^H\text{diag}(\{1-\lambda_i\})U$ and Lemma 3 it is straightforward that the matrix function $\tilde{W}(s)$ is positive real. Hence the proof is completed. ■

Now by Popov's criterion [15], [16] we finally reach the conclusion:

*Theorem 1:* Along with the initial dynamics introduced in Remark 2, the network optimization algorithm given by (18) is globally asymptotically stable and thus satisfies all the design principles proposed in Section II.

## V. Conclusions

We have succeeded in designing a scalable and distributed algorithm for the network optimization problem as promised at the beginning of the paper. We believe that our definition of the problem reflects the real meaning of the plug-and-play property for the network flow control problem and to the authors' knowledge our algorithm is the first to achieve this goal: to obtain efficient and fair bandwidth allocation for a network with the presence of delays and with minimum extra communication cost. We also believe that our approach presents a design methodology from which people can create various algorithms to meet different performance requirements while still maintaining the basic plug-and-play property, and our algorithm is just the simplest one in this class. Still many basic questions must be solved. For example we have not yet dealt with the case of time-varying delay. Also our analysis is based on fluid models of flow control mechanisms so it is interesting to see how to design real communication protocols based on our algorithm. We will address these questions in our future research.

## VI. Acknowledgments

This work was supported by the Maryland Hybrid Networks Center (HyNet), through the DARPA NMS Program under Contract N6600100C8063, and the NASA RPC Program under Cooperative Agreements NCC8-235 and NAG-2844.

## Appendix
## Proof of Lemma 4

Following the discussion in Section IV-B it is sufficient to check whether

$$-1 \notin \bar{\text{co}}(0 \cup \{g(i\omega,\tau), \forall\tau \geq 0\})$$

holds for every $\omega \geq 0$. Here we define

$$g(s,\tau) \triangleq \frac{e^{-\tau s}}{\tau^2 s^2}\frac{\tau s + \frac{1}{2}\min\{\tau,1\}}{2\max\{\tau,1\}}.$$

The proof then breaks into the examination of the 3 parts of the curve $g(i\omega,\tau)$ on the Nyquist plane for any fixed $\omega$. First we study the part of the curve after it crosses the real axis for the first time. Then we study the situation of the curve before it crosses the real axis, where the cases when $\tau \leq 1$ and $\tau > 1$ are studied separately.

First by direct calculation we obtain that for fixed $\omega$ the first intersection of $g(i\omega,\tau)$ with the real axis takes place at $\tau_0 = \omega_1/\omega$ if $\omega < \omega_1$, where $\omega_1 \approx 1.1656$ is the solution of the equation $2\omega = \tan\omega$, and $\tau_0 = (\arctan 2\omega)/\omega$ if $\omega \geq \omega_1$. The location of the intersection is

$$\begin{aligned} & -1/(2\omega\tau_0\sin\omega\tau_0\max\{\tau_0,1\}) \\ \geq\ & -1/(2\omega_1\sin\omega_1) \approx -0.4668.\end{aligned}$$

The maximum value of imaginary part attained by the curve $g(i\omega,\tau)\max\{\tau,1\}$ with fixed $\omega$ is obtained by maximizing

$$\text{Im}g(i\omega,\tau)\max\{\tau,1\} = \frac{\sin\omega\tau}{4\omega^2\tau^2} - \frac{\cos\omega\tau}{2\omega\tau}.$$

By numerical calculation the maximum value is $v_{\max} \approx 0.1824$ when $\omega\tau \approx 2.5288$. We can then show that the part of the curve $g(i\omega,\tau)$ at which $\tau \geq \tau_0$, or equivalently the part after passing the real axis, lies below the affine line $\mathcal{L}$ defined by $\text{Im}z = \omega(\text{Re}z+1)$, since the slope of the line passing through -1 under which our curve lies is less than

$$< \frac{\frac{v_{\max}}{\left(1-\frac{1}{2\omega_1\sin\omega_1}\right)\max\{\tau,1\}}}{\frac{v_{\max}}{\left(1-\frac{1}{2\omega_1\sin\omega_1}\right)\frac{\omega_1}{\omega}}} \approx 0.2964\omega.$$

Therefore the argument is valid.

Now let us inspect the part of the curve before passing the real axis. There are 2 situations. When $\tau \leq 1$ the curve can be written as

$$g(i\omega,\tau) = -\frac{e^{-i\tau\omega}}{\tau^2\omega^2}\frac{i\tau\omega+\tau/2}{2} = -\frac{e^{-i\tau\omega}}{\tau\omega}\left(i\frac{1}{2}+\frac{1}{4\omega}\right).$$

We will show that in this situation the curve lies below $\mathcal{L}$. By some manipulations this is equivalent to the following inequality

$$2\tau\omega^3 > (\omega^2+1)\sin\tau\omega - \frac{1}{2}\cos\tau\omega.$$

One can verify that

$$h_1(\omega) \triangleq 2\tau\omega^3 - (\omega^2+1)\sin\tau\omega + \frac{1}{2}\cos\tau\omega$$

is the integral of the following

$$\begin{aligned} h_1'(\omega) = \quad & 6\tau\omega^2 - \tau\omega^2\cos\tau\omega - \tfrac{\tau-1}{2}\cos\tau\omega \\ & -(\tau/2+2)\omega\sin\tau\omega \end{aligned}$$

with respect to $\omega$. The above is greater than zero for $\tau \in [0,1]$ since

$$h_1'(\omega) > 5\tau\omega^2 - (\tau/2+2)\omega\sin\tau\omega > \frac{5}{2}\omega(\tau\omega - \sin\tau\omega) > 0.$$

Since $h_1(0) = 1/2 > 0$, $h_1(\omega) = h_1(0) + \int_0^\omega h_1'(u)du > 0$. We then conclude that when $\tau \in [0,1]$ the curve lies below the line $\mathcal{L}$.

In the other situation when $\tau > 1$ the curve can be expressed as

$$g(i\omega,\tau) = -\frac{e^{-i\tau\omega}}{\tau^2\omega^2}\frac{i\tau\omega + 1/2}{2\tau}.$$

We will just consider the curve

$$\tilde{g}(i\omega,\tau) = -\frac{e^{-i\tau\omega}}{\tau^2\omega^2}\frac{i\tau\omega + 1/2}{2}.$$

since the curve $\tilde{g}$ lies above the curve $g$. Again by using the simplifying notation $\theta = \tau\omega > \omega$ and after some algebraic manipulations, it suffices to show the validity of the following inequality

$$4\omega\theta^2 > \omega(2\theta\sin\theta + \cos\theta) + (\sin\theta - 2\theta\cos\theta).$$

It is actually sufficient to check the above inequality when $\theta = \omega$. So we will only need to show

$$4\theta^3 > (2\theta^2+1)\sin\theta - \theta\cos\theta.$$

Similarly to the situation when $\tau \leq 1$, we define a function

$$h_2(\theta) \triangleq 4\theta^3 - (2\theta^2+1)\sin\theta + \theta\cos\theta.$$

Its derivative is

$$h_2'(\theta) = 12\theta^2 - 2\theta^2\cos\theta - 5\theta\sin\theta.$$

But

$$h_2'(\theta) > 10\theta^2 - 5\theta\sin\theta > 5\theta(\theta - \sin\theta) > 0,$$

and by $h_2(0) = 0$, one obtains $h_2(\theta) = h_2(0) + \int_0^\theta h_2'(u)du > 0$. Therefore we have shown that the curve $\tilde{g}$ lies below the line $\mathcal{L}$ when $\tau > 1$.

Combining all these results we have confirmed that for any fixed $\omega$, the curve $g(i\omega,\tau)$ lies below the line $\mathcal{L}$ for all $\tau \geq 0$ and therefore the convex hull of $g(i\omega,\tau)$ and 0 cannot contain the point -1 on the Nyquist plane.

## REFERENCES


[1] V. Jacobson, R. Braden, and D. Borman, "TCP extenstions for high performance," RFC 1323, May 1993.

[2] F. Kelly, A. Maulloo, and D. Tan, "Rate control for communication networks: Shadow prices, proportional fairness and stability," *J. Oper. Res. Soc.*, vol. 49, pp. 237–252, Mar. 1998.

[3] S. Low and D. Lapsley, "Optimization flow control I: Basic algorithm and convergence," *IEEE/ACM Trans. Networking*, vol. 7, pp. 861–874, Dec. 1999.

[4] J. Mo and J. Walrand, "Fair end-to-end window-based congestion control," *IEEE/ACM Trans. on Networking*, vol. 8, no. 5, pp. 556–567, 2000.

[5] D. Bertsekas, *Constrained Optimization and Lagrange Multiplier Methods*. Academic Press, 1982.

[6] S. Low, F. Paganini, J. Wang, and J. Doyle, "Linear stability of TCP/RED and a scalable control," *Computer Networks Journal*, vol. 43, no. 5, pp. 633–647, Dec. 2003.

[7] F. Paganini, Z. Wang, J. Doyle, and S. Low, "Congestion control for high performance, stability and fairness in general networks," *IEEE/ACM Trans. Networking*, 2004.

[8] G. Vinnicombe, "On the stability of networks operating TCP-like congestion control," *Proc. IFAC*, 2002.

[9] Z. Wang and F. Paganini, "Global stability with time-delay in network congestion control," *IEEE Conference on Decision and Control*, Dec. 2002.

[10] J. Hale and S. V. Lunel, *Introduction to Functional Differential Equations*. Springer-Verlag New York, Inc., 1993.

[11] X. Fan, M. Arcak, and J. Wen, "Robustness of network flow control against disturbances and time-delay," *System and Control Letters*, vol. 53, no. 1, pp. 13–29, 2004.

[12] R. Ranjan, R. La, and E. Abed, "Global stability conditions for rate control with arbitrary communication delays," *IEEE Conference on Decision and Control*, Dec. 2004.

[13] W. Rugh, *Linear System Theory*, 2nd ed. Prentice Hall, 1999.

[14] K. Narendra and J. Taylor, *Frequency Domain Criteria for Absolute Stability*. Academic Press, 1973.

[15] A. Halanay, *Differential Equations: Stability, Oscillations, Time Lags*. Academic Press, 1966.

[16] X. Li, "On the absolute stability of systems with time lags," *Chinese Mathematics*, vol. 4, pp. 609–626, 1963.

[17] C. Desoer and Y. Yang, "On the generalized Nyquist stability criterion," *IEEE Transcations on Automatic Control*, vol. 25, pp. 187–196, 1980.