STABILITY AND PERFORMANCE ANALYSIS OF NETWORKS SUPPORTING SERVICES WITH RATE CONTROL – COULD THE INTERNET BE UNSTABLE?

GUSTAVO DE VECIANA, TAE-JIN LEE AND TAKIS KONSTANTOPOULOS
Department of Electrical and Computer Engineering
The University of Texas, Austin, TX 78712, USA

Abstract—**We consider the stability and performance of a model for networks supporting services that adapt their transmission to the available bandwidth. Not unlike real networks, in our model connection arrivals are stochastic and have a random amount of data to send, so the number of connections in the system changes over time. In turn the bandwidth allocated to, or throughput achieved by, a given connection, may change during its lifetime due to feedback control mechanisms that react to congestion and thus implicitly to the number of ongoing connections. Ideally, for a fixed number of connections, such mechanisms reach an equilibrium typically characterized in terms of its 'fairness' in allocating bandwidth to users, *e.g.,* max-min fair. In this paper we prove the stability of such networks when the offered load on each link does not exceed its capacity. We use simulation to investigate the performance, in terms of average connection delays, for various network topologies and fairness criteria. Finally we pose an architectural problem in TCP/IP's decoupling of the transport and network layer from the point of view of guaranteeing connection level stability, which we claim may explain congestion phenomena on the Internet.**

I. INTRODUCTION

Future communication networks are likely to support *elastic* applications that permit adaptation of the data transmission rate to the available network bandwidth while achieving a graceful degradation in the perceived quality of service [19]. Transport services that match the flexibility of such applications are already supported on the Internet via TCP wherein end-systems adjust their transmissions in response to delayed or lost packets, *i.e.,* implicit indicators of available bandwidth [11]. Available Bit Rate service, defined for ATM networks, draws on both the end-systems and network elements to implement a similar functionality through adaptive rate control mechanisms that strive to allocate the available bandwidth among ongoing connections [5]. Typically such mechanisms represent an efficient way to carry traffic corresponding to elastic applications, ranging from today's file transfers to future rate adaptive voice/video applications.

Since mechanisms to adapt transmissions typically draw on delayed (implicit or explicit) feedback from the network, much work has been devoted to establishing their stability, particularly for networks supporting a *fixed* number of connections. Stability is usually interpreted as avoiding queue/delay buildups, and/or somewhat loosely as ensuring that transmission rates converge to an equilibrium corresponding to a bandwidth allocation among ongoing connections, see *e.g.,* [2], [3], [6], [20], [15], [1], [13]. Such equilibria are in turn usually characterized in terms of their 'fairness' to users, such as (weighted) max-min fairness or proportional fairness [4], [12]. Thus given a fixed number of users and fixed network capacities, one can typically arrange (through an appropriate control mechanism) to achieve an equilibrium which represents, according to some criterion, an equitable allocation of resources among users.

By contrast very little is known about the network's performance when the number of connections in the network is in constant flux. Previous work along these lines has focused on studying transients, *i.e.,* how quickly will the transmission rates reach a new equilibrium. In this paper we consider a novel model that includes stochastic arrivals and departures. However it abstracts the queueing and rate adaptation that would be taking place in the network by assuming that an equilibrium, and thus appropriate bandwidth allocation is immediately achieved. In essence, this corresponds to assuming a *separation of time scales* between the time scales of connection arrivals and departures and those on which rate control processes converge to equilibria. Our focus is on exploring the stability and performance of this connection-level model for networks using different types of rate control and thus operated under different fairness policies.

Paralleling models used in the circuit switched literature, we assume connection arrivals processes are Poisson and that each connection has a random, exponentially distributed, amount of data to send.[1] In contrast to circuit switched models, the bandwidth allocated to each user will be a function of the global state of the network. Indeed recall that the bandwidth allocated to a user depends on the equilibrium achieved by the rate control mechanisms and the number of ongoing connections.

In general, one expects work conserving systems to be stable when the offered load to each link (queue) in the network does not exceed its capacity. However given the complex network-wide interactions underlying the bandwidth allocation mechanism, a demonstration of this fact was an open question. Note that our model can be said to be 'non-work conserving' in the sense that a link supporting active connections may not be operating at a full capacity because its connections are 'bottlenecked' elsewhere – a typical sign of a potential for instability. In this paper we come to terms with this problem by showing the stability of our model when natural conditions are satisfied.

This work is supported by a National Science Foundation NSF Career Awards NCR 96-24230 and NCR 95-02582, by Grant ARP-224 of the Texas Higher Education Coordinating Board, by Southwestern Bell Co., and by an Intel Technology for Education 2000 equipment grant.

[1] This arrivals model is a typical and a reasonable assumption for connections generated by a large population of independent users. The exponential assumption simplifies our analysis but is likely not to be critical for the stability results in this paper.

Since ours is a higher layer model, it is logical to consider network-level performance, say in terms of average connection delays. This is important because the goals of fairness and low connection delays may not be compatible, and should be examined prior to committing to a particular architecture for large-scale broadband networks. Moreover network designers might want to dimension capacities to achieve a reasonable responsiveness, say for web browsing, when the network is subject to typical loads. Our preliminary simulations suggest that indeed it may be of interest to examine more carefully the impact of a given fairness criterion and topology on the overall network performance.

Based on our model we point out an insidious architectural problem in networks supporting adaptive services of this type. To achieve connection layer stability we must ensure that connection level loads do not exceed link capacities. Clearly this then requires that the routing layer be aware of the connection level offered loads. However, typical routing algorithms draw on short term link averages of utilization or packet delays. Such metrics reflect the connection level offered loads quite poorly, since connections are adapting their transmission rates depending on link congestion. Loosely speaking, the router is indifferent to the fact that a 90 % link utilization may be due to a single traffic source or a thousand sources transmitting at a thousandth of the latter's rate. Herein lies a possible explanation for the congestion currently experienced on the Internet, *i.e.,* connection level instability.

The paper is organized as follows. In §II, we present our model and define the max-min, weighted max-min and proportionally fair bandwidth allocations. Next, in §III we show the stability of the model by constructing appropriate Lyapunov functions. Performance issues are discussed in §IV. In §V we return to our question concerning possible connection level instabilities in current networks and discuss future work.

II. Network model and bandwidth allocation schemes

Our network model consists of a set of links $\mathcal{L}$ with fixed capacities $c = (c_\ell, \ell \in \mathcal{L})$ in bits/sec shared by a collection of routes $\mathcal{R}$. Routes are undirected and may traverse several links in the network.[2] A 0-1 matrix $A = (A_{\ell r}, \ell \in \mathcal{L}, r \in \mathcal{R})$ indicates which links a route traverses. In other words, $A_{\ell r} = 1$ if route r uses link ℓ and zero otherwise.

The dynamics of the model are as follows. New connections are initiated on route $r \in \mathcal{R}$ at random times forming a Poisson process Π_r with rate λ_r connections/sec. The collection of processes $\Pi = \{\Pi_r, r \in \mathcal{R}\}$, with rates $\lambda = (\lambda_r, r \in \mathcal{R})$ are assumed to be independent. Each connection has a volume of data (in bits) to transmit, which is assumed to be an exponentially distributed random variable with mean b bits. The parameter b is the same for all connections, irrespective of route or arrival time. This assumption simplifies the description of the system state and, consequently, its analysis. The random variables representing connection volumes are thus i.i.d. and also independent of Π. We let $\nu_\ell = c_\ell b^{-1}$ denote the capacity of link ℓ expressed in connections/sec, and let $\nu = (\nu_\ell, \ell \in \mathcal{L})$.

The "state" of the network is denoted by $n = (n_r, r \in \mathcal{R})$ where n_r is the number of connections currently on route r. We assume that the bandwidth allocated to each ongoing connection depends only on the current state n of the system. Let $\mu_r(n)$ denote the total bandwidth allocated to connections on route r when the system state is n, expressed as a service rate in connections/sec. The choice of the functions $\mu = (\mu_r : \mathbb{Z}_+^{\mathcal{R}} \to \mathbb{R}_+, r \in \mathcal{R})$ will be described in the sequel. If the state of the system changes during the sojourn of a connection (*e.g.,* due to the establishment of a new connection or the termination of an existing one), then, there may be a corresponding change (speed-up or slow-down) in its service rate. Indeed since no arriving connections are blocked, new connections must accommodated by changing the bandwidth allocation, whereas bandwidth made available by departing connections is reallocated to the remaining ones. We assume that ongoing connections are *greedy* in the sense that they will use whatever network bandwidth is made available to them. Note that in reality a given connection may have a limit on the rate at which it can transmit, *e.g.,* may be limited by the access network or network interface card. Herein we shall assume that such bottlenecks have been explicitly modeled by incorporating limited capacity access links in the network.

Let $\Pi_r(t)$ denote the number of connections arriving on route r on the time interval $(0, t]$. This is a rate-λ_r Poisson counting process. Let $\Phi_r(t)$ be another independent unit rate Poisson process. Letting $\{N_r(t), t \geq 0\}$ be the random process corresponding to the number of connections on route r, we have

$$N_r(t) = N_r(0) + \Pi_r(t) - \Phi_r\left(\int_0^t \mu_r(N(s))\, ds\right), \qquad (1)$$

where $r \in \mathcal{R}$, $t \geq 0$, which captures the state dependent service rates along each route in the network. It should be clear that given an initial state $N_r(0)$, this evolution equation has a unique solution. Moreover, if the initial condition $(N_r(0), r \in \mathcal{R})$ is selected independently of the arrivals and service processes then the $\mathbb{Z}_+^{\mathcal{R}}$–valued process $N(t) = (N_r(t), r \in \mathcal{R})$ is Markovian.

In the sequel, we describe various bandwidth allocation schemes, or, equivalently, various possible functions μ. In particular we will use μ^m, μ^w and μ^p to denote the max-min, weighted max-min and proportionally fair bandwidth allocation functions. Notice that these functions, of the state n, depend on the capacity vector ν, the routing matrix A, and the type of rate control used on the network. By contrast with standard queuing models, which track packets and queues throughout the network, it is through this dependence that the evolution (1) models the dynamics of the network. Also note that we have assumed that connections are not rerouted once they are initiated. One could in principle account for rerouting or split-

[2] Our model is at the connection level, so we can assume undirected routes without loss of generality.

ting of flows across the network but this will not be considered here. Finally, and to avoid possible confusion, bandwidth will be measured in units of connections/sec rather than bits/sec – see above discussion.

A. Max-min fair bandwidth allocation

We first consider max-min fair bandwidth allocation. An allocation is said to be max-min fair if the bandwidth allocated to a connection cannot be increased without also decreasing that of a connection having a less than or equal allocation [4]. For a single link network this translates to giving each connection traversing the link the same amount of bandwidth. In general one first determines what would be the maximum minimum bandwidth one could assign to any connection in the network and allocates it to the most poorly treated connections. One then removes these connections and the allocated bandwidths from the network, and iteratively repeats the process of maximizing the minimum bandwidth allocation for the remaining connections. More formally the max-min fair allocation can be defined in terms of a hierarchy of optimization problems, described in detail in [10], which is easily solved via the above procedure. Below we briefly review how given the state n of the network one determines the max-min fair bandwidth allocations per connection and in turn determines the bandwidth allocations $(\mu_r^m(n), r \in \mathcal{R})$ per route.

Let the vector $a^* = (a_r^*, r \in \mathcal{R})$ be the max-min fair allocation where a_r^* denotes the bandwidth, in connections/sec, allocated to a single connection on route r. Notice that we have suppressed the dependence of a^* on n. All connections on the same route get the same allocation so $\mu_r^m(n) = n_r a_r^*$. We determine a^* as follows. First for all routes $r \in \mathcal{R}$ such that $n_r = 0$ we set $a_r^* = 0$ and thus $\mu_r^m(n) = 0$. Next we solve a hierarchy of optimization problems starting with

$$f^{(1)}(n) := \max_a \{\min_{r\in\mathcal{R}} a_r : \sum_{r\in\mathcal{R}} A_{\ell r} n_r a_r \le \nu_\ell,\ \ell \in \mathcal{L}\}, \quad (2)$$

which corresponds to maximizing the minimum bandwidth per connection subject to the link capacity constraints. It can be shown, see [10], that the solution to this problem is given by

$$f^{(1)}(n) = \min_{\ell\in\mathcal{L}} f_\ell^{(1)}(n) \text{ with } f_\ell^{(1)}(n) := \frac{\nu_\ell}{\sum_{r\in\mathcal{R}} A_{\ell r} n_r}, \quad (3)$$

where $f_\ell^{(1)}(n)$ can be thought of as the *fair share* at link ℓ, *i.e.,* the bandwidth per connection at link ℓ if its capacity were equally divided among the connections traversing the link.

Let $\mathcal{L}^{(1)}$ be the set of links ℓ such that $f_\ell^{(1)}(n) = f^{(1)}(n)$. This is the set of first-level *bottleneck links.* The set of first-level *bottleneck routes* $\mathcal{R}^{(1)}$ is the set of routes traversing a link in $\mathcal{L}^{(1)}$. These two sets make up the first-level of the *bottleneck hierarchy.* Finally, for each route $r \in \mathcal{R}^{(1)}$, let $a_r^* = f^{(1)}(n)$. The remaining, if any, components of a^* are determined by repeating this process on a reduced network as explained next.

In its second step, if it arises, the algorithm replaces the sets $\mathcal{L}$ and $\mathcal{R}$ by $\mathcal{L} \setminus \mathcal{L}^{(1)}$ and $\mathcal{R} \setminus \mathcal{R}^{(1)}$, respectively. The new state of the system is simply the projection $(n_r, r \in \mathcal{R} \setminus \mathcal{R}^{(1)})$, and a new link capacity vector, $\nu^{(1)}$ is defined on $\mathcal{L} \setminus \mathcal{L}^{(1)}$, where ν_ℓ is reduced to

$$\nu_\ell^{(1)} = \nu_\ell - \sum_{r\in\mathcal{R}^{(1)}} A_{\ell r}\mu_r^m(n) = \nu_\ell - f^{(1)}(n) \sum_{r\in\mathcal{R}^{(1)}} A_{\ell r} n_r.$$

From (2) and the definition of $\mathcal{L}^{(1)}$, it is clear that the reduced capacities are non-negative. A new problem paralleling (2) but on the reduced network (with reduced sets or routes and links, reduced state, and reduced capacities–as described above) is then defined and solved to obtain a new value $f^{(2)}(n)$, and second-level bottleneck sets $\mathcal{L}^{(2)}$ and $\mathcal{R}^{(2)}$. Finally for $r \in \mathcal{R}^{(2)}$ we set $a_r^* = f^{(2)}(n)$. If necessary this process is once again repeated, but, since the sets $\mathcal{R}^{(1)}, \mathcal{R}^{(2)}, \ldots$ are nonempty, it terminates in a finite number of steps, uniquely specifying the vector a^* and thus $\mu^m(n)$.

Notice that in the above procedure n need not be integer valued, hence $\mu^m(n)$ can be easily extended for real-valued arguments. We shall use the same notation to denote the extension of μ^m from $\mathbb{Z}_+^{\mathcal{R}}$ to $\mathbb{R}_+^{\mathcal{R}}$. Some straightforward properties of this function are summarized below.

Proposition II.1: The function $\mu^m : \mathbb{R}_+^{\mathcal{R}} \to \mathbb{R}_+^{\mathcal{R}}$ is radially homogeneous, in the sense that

$$\mu^m(\alpha x) = \mu^m(x),\ \ x \in \mathbb{R}_+^{\mathcal{R}},\ \ \alpha > 0.$$

In the interior of the positive orthant $\mathbb{R}_+^{\mathcal{R}}$, the function μ^m is continuous, and has strictly positive components. Finally, μ^m is bounded.

The proof of this proposition can be shown by induction on the bottleneck hierarchy and considering the dependence on x of the max-min fair bandwidth allocation.

Notice that the bandwidth allocation policy reflected in μ^m satisfies the link capacity constraints, is fair in the max-min fair sense, but the performance, *e.g.,* in terms of connection delays, may be poor. In the next section we discuss the weighted max-min fair bandwidth allocation which allows some latitude in controlling performance by giving different priorities to connections based on their routes.

B. Weighted max-min fair bandwidth allocation

Let $w = (w_r, r \in \mathcal{R})$ be positive "weights" associated with each route in the network, and $a^{w*} = (a_r^{w*}, r \in \mathcal{R})$ denote the weighted max-min fair bandwidth allocation vector. For a given state n we determine a^{w*} in a similar fashion to the max-min fair allocation. First for all routes $r \in \mathcal{R}$ such that $n_r = 0$ set $a_r^{w*} = 0$. Next, replace (2) with

$$f^{(1),w}(n) := \max_a \{\min_{r\in\mathcal{R}} \{a_r/w_r\} : \sum_{r\in\mathcal{R}} A_{\ell r} n_r a_r \le \nu_\ell,\ \ell \in \mathcal{L}\},$$

which can again be solved by first defining the *weighted fair share* on link ℓ as

$$f_\ell^{(1),w}(n) := \frac{\nu_\ell}{\sum_{r\in\mathcal{R}} A_{\ell r} w_r n_r} \quad (4)$$

and then setting $f^{(1),w}(n) = \min_{\ell\in\mathcal{L}} f_\ell^{(1),w}(n)$. Paralleling the max-min fair case, the first-level bottleneck links and routes, denoted $\mathcal{L}^{(1),w}$ and $\mathcal{R}^{(1),w}$ respectively, can be defined, and one can proceed iteratively to determine the bandwidth allocation for connections on all routes. We will let $\mu^w(n)$ denote the vector of bandwidths allocated to each route where $\mu_r^w(n) = w_r n_r a_r^{w*}$, and let $\mu^w = (\mu_r^w : \mathbb{Z}_+^{\mathcal{R}} \to \mathbb{R}_+, r \in \mathcal{R})$.

One can again extend μ^w for real-valued arguments *i.e.,* from $\mathbb{Z}_+^{\mathcal{R}}$ to $\mathbb{R}_+^{\mathcal{R}}$, and show that

$$\mu^w(x) = \mu^m(Dx), \tag{5}$$

where μ^m corresponds to the unweighted max-min fair allocation discussed in the previous section, and $D = \text{diag}(w)$, *i.e.,* a square matrix with components $(w_r, r \in \mathcal{R})$ along its diagonal. Thus one way to view the weighted max-min fair allocation is as a max-min fair allocation where the "effective number" of ongoing connections is Dx. Moreover one can easily see that the results in Proposition II.1 also apply to μ^w.

A weighted max-min fair allocation can be used to differentiate among connections following different routes and thus give priority based on geographic, administrative, or service requirements by grouping like connections on a route. However specific criteria for the selection of weights need to be developed – this is discussed in §IV. In principle one can consider control policies which adjust the weights based on the state of the network – a simple example is briefly considered in §III-D

C. Proportionally fair bandwidth allocation

As a final alternative we consider a framework where utility functions $U_r : \mathbb{R}_+ \to \mathbb{R}, r \in \mathcal{R}$ have been associated with connections following various routes. Here $U_r(a_r)$ is the utility to a user/connection on route r of a bandwidth allocation a_r.[3] A bandwidth allocation policy which maximizes the total network utility when the state is n can be obtained by solving the following optimization problem:

$$\max_a \left\{ \sum_{r\in\mathcal{R}} n_r U_r(a_r) : \sum_{r\in\mathcal{R}} A_{\ell r} n_r a_r \le \nu_\ell, \ell \in \mathcal{L}; a \ge 0 \right\}, \tag{6}$$

where we assume that the utility functions are concave and so the optimizer is unique. This approach to allocating bandwidth is pleasing in the sense that it finds an appropriate compromise between the extent to which users value bandwidth and the *overall* user "satisfaction."

In general it is unclear how to select utility functions. However, [12] and others, have considered the case where $U_r(a_r) = \log a_r$ and shown that in this case the maximizer $a^{p*} = (a_r^{p*}, r \in \mathcal{R})$ corresponds to a *proportionally fair* bandwidth allocation in the sense that the vector is feasible, *i.e.,* satisfies the link capacity constraints, and for any other feasible rate $a' = (a'_r, r \in \mathcal{R})$, the aggregate proportional change is negative, *i.e.,*

$$\sum_{r\in\mathcal{R}} n_r \frac{a'_r - a_r^{p*}}{a_r^{p*}} < 0.$$

Determining the maximizer of (6) for log utility functions can be done explicitly for simple networks. Alternatively, as with max-min fairness, one can design rate control mechanisms that converge to the associated bandwidth allocation [13]. We will let $\mu_r^p(n) = n_r a_r^{p*}$ denote the total bandwidth allocated to connections along route $r \in \mathcal{R}$ and $\mu^p(n) = (\mu_r^p(n), r \in \mathcal{R})$ be the bandwidth allocations per route when proportional fairness is used.

III. Stability of the Stochastic Network

In this section we will consider the stability of the stochastic network model defined in §II, for various types of bandwidth allocation. Assuming $\{\Pi_r, \Phi_r, r \in \mathcal{R}\}$ are independent Poisson processes on $[0, \infty)$, where Π_r has rate λ_r and Φ_r has rate 1, the evolution equation (1) defines a Markov chain on $\mathbb{Z}_+^{\mathcal{R}}$ with transition rates

$$q(n,m) = \begin{cases} \lambda_r, & m = n + e^r,\ r \in \mathcal{R} \\ \mu_r(n), & m = n - e^r,\ r \in \mathcal{R} \\ 0, & \text{otherwise} \end{cases}, \tag{7}$$

for $m \ne n$, where $e^r = (e_s^r, s \in \mathcal{R})$, $e_s^r = 1(r = s)$. Thus, when the state is n, route r sees arrivals with rate λ_r and departures with rate $\mu_r(n)$. Note that when $n_r = 0$ we have $\mu_r(n) = 0$, thus $q(n, n - e^r) = 0$, and so the rates are supported on the positive orthant.

We use the notation Q for the infinitesimal generator (viz., rate matrix) of this continuous-time Markov chain. For a function $\varphi : \mathbb{R}_+^{\mathcal{R}} \to \mathbb{R}$, we write[4]

$$Q\varphi(n) := \sum_{m\in\mathbb{Z}_+^{\mathcal{R}}} q(n,m)\varphi(m) = \sum_{m\in\mathbb{Z}_+^{\mathcal{R}}} q(n,m)[\varphi(m) - \varphi(n)], \tag{8}$$

where the latter equality follows from the fact that Q is conservative: $q(n,n) = -\sum_{m\ne n} q(n,m)$. Note that $Q\varphi(n)$ can be interpreted as the expected drift, *i.e.,* the change in $\varphi(N(t))$ when $N(t) = n$.

Clearly the Markov chain $\{N(t),\ t \ge 0\}$ is irreducible, and we say that it is stable, iff it is positive recurrent. We will show positive recurrence by constructing a Lyapunov function [17], [9]. For our system, a Lyapunov function is any function $V : \mathbb{Z}_+^{\mathcal{R}} \to \mathbb{R}_+$ with the sole property that there exists a finite set $K \subseteq \mathbb{Z}_+^{\mathcal{R}}$, such that

$$\sup_{n\notin K} QV(n) < 0, \tag{9}$$

[3] If there exist connections with different utility functions that follow the same path, one can define several routes carrying connections that share the same utility function.

[4] Notice that the sums in (8) have a finite number of terms, since the chain has only local transitions, *i.e.,* arrivals and departures for every route, thus there are no restrictions on the function φ for $Q\varphi(n)$ to be well defined.

where QV is defined as in (8). Using our formula (7) for the transition rates we can rewrite QV as

$$\begin{aligned} QV(n) &= \sum_{r\in\mathcal{R}}\{\lambda_r[V(n+e^r)-V(n)] \\ &\quad +\mu_r(n)[V(n-e^r)-V(n)]\}. \end{aligned} \tag{10}$$

Intuitively (9) means that when the process $N(t)$ lies outside K, it is such that on average $V(N(t))$ is decreasing, *i.e.,* has negative drift.

Searching for such a Lyapunov function can be a tedious procedure, particularly since the transition rates of our Markov chain are defined via the optimization problem associated with the various fairness criteria. Below we consider the stability of networks subject to various types of max-min bandwidth allocation. Stability for the case of proportionally fair bandwidth allocation has been shown via a quadratic function [14], but has been omitted due to space constraints.

A. Stability under max-min fair bandwidth allocation

We first consider the stability of the network when bandwidth is allocated according to the max-min fair criterion and thus the dynamics of the system are captured by (1) with μ replaced by μ^m as defined in §II-A.

We will begin by considering a *candidate* Lyapunov function, related to the max-min fairness criterion. Let $V(n)$ be the reciprocal of $f^{(1)}(n)$ defined in (2) and extend it from $\mathbb{Z}_+^{\mathcal{R}}$ to $\mathbb{R}_+^{\mathcal{R}}$, namely,

$$V(x)=\max_{\ell\in\mathcal{L}}\{\nu_\ell^{-1}\sum_{r\in\mathcal{R}}A_{\ell r}x_r\},\;\; x\in\mathbb{R}_+^{\mathcal{R}}.$$

For convenience we introduce the vectors

$$\xi^\ell=(\xi_r^\ell,\, r\in\mathcal{R}),\;\; \xi_r^\ell:=\nu_\ell^{-1}A_{\ell r},\;\; \ell\in\mathcal{L}. \tag{11}$$

and let $\varphi^\ell(x)=\langle\xi^\ell,x\rangle,\ell\in\mathcal{L}$ where $\langle\cdot,\cdot\rangle$ denotes the standard inner product in $\mathbb{R}^{\mathcal{R}}$. With this notation we have that

$$V(x)=\max_{\ell\in\mathcal{L}}\,\varphi^\ell(x)=\max_{\ell\in\mathcal{L}}\,\langle\xi^\ell,x\rangle. \tag{12}$$

Thus V is a piecewise linear function. Since the vectors ξ^ℓ have non-negative components, the sets $\{x\in\mathbb{R}_+^{\mathcal{R}}\,:\,V(x)\le\alpha\}$ are compact polytopes, for all $\alpha\ge 0$. For a fixed x, one or more of the indices ℓ achieve the maximum in (12)– these are the first-level bottleneck links defined earlier. We will use $\mathcal{L}^{(1)}(x)$ to denote the dependence of the first-level bottleneck links on x. Similarly $\mathcal{R}^{(1)}(x)$ and $\nu^{(1)}(x)$ will be used to indicate such dependencies in the sequel.

Since for first level bottleneck links the link capacity is fully utilized among ongoing connections, we would expect that, on average, the number of connections on such a link $\sum_{r\in\mathcal{R}}A_{\ell r}n_r$ will decrease as long as the average arrival rate does not exceed the link capacity. The following lemma makes this clear.

Lemma III.1: Assume that $A\lambda<\nu$, *i.e.,* $\sum_{r\in\mathcal{R}}A_{\ell r}\lambda_r<\nu_\ell$, for all $\ell\in\mathcal{L}$.[5] Then, there is a constant $c>0$, such that for all $x\in\mathbb{R}_+^{\mathcal{R}}$, and all $\ell^*\in\text{argmax}_{\ell\in\mathcal{L}}\varphi^\ell(x)$, *i.e.,* first-level bottleneck links $\ell^*\in\mathcal{L}^{(1)}(x)$, we have

$$Q\varphi^{\ell^*}(x)=\langle\xi^{\ell^*},\lambda-\mu^m(x)\rangle\le -c. \tag{13}$$

Proof: First, using (10) and the definition (11) of ξ^ℓ we have that

$$\begin{aligned} Q\varphi^{\ell^*}(x) &= \sum_{r\in\mathcal{R}}\xi_r^{\ell^*}(\lambda_r-\mu_r^m(x))=\langle\xi^{\ell^*},\lambda-\mu^m(x)\rangle \\ &= \sum_{r\in\mathcal{R}}\nu_{\ell^*}^{-1}A_{\ell^* r}(\lambda_r-\mu_r^m(x)). \end{aligned}$$

Next, since ℓ^* is a first-level bottleneck link, it follows that for routes r traversing link ℓ^* we have $\mu_r^m(x)=x_r a_r^*$ where a_r^* is given by (2). Thus,

$$\begin{aligned} Q\varphi^{\ell^*}(x) &= \nu_{\ell^*}^{-1}\left(\sum_{r\in\mathcal{R}}A_{\ell^* r}\lambda_r-\sum_{r\in\mathcal{R}}A_{\ell^* r}\frac{\nu_{\ell^*}x_r}{\sum_{s\in\mathcal{R}}A_{\ell^* s}x_s}\right) \\ &= \nu_{\ell^*}^{-1}\left(\sum_{r\in\mathcal{R}}A_{\ell^* r}\lambda_r-\nu_{\ell^*}\right)\le -c, \end{aligned}$$

where $c:=\max_{\ell\in\mathcal{L}}\{\nu_\ell^{-1}(\nu_\ell-\sum_{r\in\mathcal{R}}A_{\ell r}\lambda_r)\}$ is positive by the stability condition. ■

Despite the promise of Lemma III.1 it is unclear whether V is an appropriate Lyapunov function. Indeed the lemma only suggests that as long as the state makes transitions on regions having the *same* first level bottleneck links, $V(N(t))$ will experience a negative drift. To make this more precise we will explicitly identify these regions and for clarity present an example in §III-C. Let $\mathcal{M}$ be a nonempty subset of $\mathcal{L}$ and let

$$C_{\mathcal{M}}=\{x\in\mathbb{R}_+^{\mathcal{R}}\,:\,\mathcal{L}^{(1)}(x)=\mathcal{M}\}. \tag{14}$$

It is clear that if $\alpha>0, x\in C_{\mathcal{M}}\Rightarrow \alpha x\in C_{\mathcal{M}}$, *i.e.,* these sets are cones, and that

$$\bigcup_{\mathcal{M}\subseteq\mathcal{L},\mathcal{M}\neq\emptyset} C_{\mathcal{M}}=\mathbb{R}_+^{\mathcal{R}}.$$

Suppose that $n\in C_{\mathcal{M}}$, for some nonempty $\mathcal{M}$, then the drift $QV(n)$ can easily be computed (see (10)), provided $n+e^r, n-e^r\in C_{\mathcal{M}}$, for all $r\in\mathcal{R}$. In this case, with ℓ any element of $\mathcal{M}$, we have $QV(n)=\langle\xi^\ell,\lambda-\mu^m(n)\rangle\le -c$, by Lemma III.1. However when n and $n+e^r$ or $n-e^r$ belong to different cones an explicit verification of the negative drift requirement becomes difficult. Indeed when this is the case a transition causes a change in the bottleneck links – alternatively we are "crossing of a boundary" of one of the cones. Intuitively we may argue that this effect is negligible, since it occurs at a relatively small fraction of points in the state space.

[5] In the sequel this will be referred to as the stability condition.

To make this intuition into a rigorous statement observe that Lemma III.1 also implies there is a $c > 0$, such that

$$\langle \nabla V(x), \lambda - \mu^m(x) \rangle \leq -c, \tag{15}$$

for all x at which the gradient $\nabla V(x) := (\partial V(x)/\partial x_r,\ r \in \mathcal{R})$ exists. It is easy to see that this gradient exists almost everywhere, and, when it exists, it equals ξ^ℓ, for some ℓ. We will start by showing that there exists a smoothened version W of the function V that satisfies a drift condition in the sense of (15) for all $x \in \mathbb{R}_+^{\mathcal{R}}$.

Lemma III.2 ([8]) If $A\lambda < \nu$, then there is a non-negative function W, defined on $\mathbb{R}_+^{\mathcal{R}} \setminus \{0\}$, that is at least twice-continuously differentiable, has a Hessian[6], $\nabla^2 W(x)$, such that $\nabla^2 W(x) \to 0$, as $|x| \to \infty$, and which satisfies the following drift condition: there is a $d > 0$, such that

$$\langle \nabla W(x), \lambda - \mu^m(x) \rangle \leq -d, \quad \text{for all } x \neq 0.$$

Next we show that the network is indeed positive recurrent.

Theorem III.1: If $A\lambda < \nu$ then the Markov chain $\{N(t),\ t \geq 0\}$ associated with the max-min fair bandwidth allocation is positive recurrent.

Proof: Since W is twice differentiable it follows by the Mean Value Theorem that for $n, m \in \mathbb{Z}_+^{\mathcal{R}}$ there exists a $\theta, 0 \leq \theta \leq 1$ such that

$$\begin{aligned} W(n+m) - W(n) &= \langle \nabla W(n), m \rangle + \frac{1}{2} m^T \nabla^2 W(n + \theta m) m \\ &:= \langle \nabla W(n), m \rangle + \beta(n, m). \end{aligned}$$

Recall that $\nabla^2 W(n) \to 0$ and thus $\beta(n, z) \to 0$ as $|n| \to \infty$. Now, using this approximation to compute QW, as in (10), we have

$$QW(n) = \langle \nabla W(n), \lambda - \mu^m(n) \rangle + \sum_m q(n, m) \beta(n, m - n).$$

It follows by Lemma III.2 that the first term is at most $-d$. The second term, is a sum of a finite number of terms, and can be made smaller than $d/2$ for all $|n|$ sufficiently large. Thus noting that $\sup_{|n| > \gamma} QW(n) < 0$, for sufficiently large γ, and letting $K = \{n : |n| \leq \gamma\}$ we satisfy the drift condition (9) which as discussed earlier implies positive recurrence. ∎

B. Stability under weighted max-min fair bandwidth allocation

While the previous result is intuitive, in that the number of connections on bottleneck links must be decreasing, it is not easily extended it to show the stability of networks under weighted max-min fair bandwidth allocation. Thus, we develop an alternative approach which, instead of focusing links, focuses on the relative states of each route. Suppose that a set of weights w is selected and the network is operated subject to the bandwidth allocation function μ^w defined in §II-B. We will let $\varphi^r(x) = \lambda_r^{-1} w_r x_r, r \in \mathcal{R}$ and consider the candidate function

$$V(x) = \max_{r \in \mathcal{R}} \varphi^r(x) = \max_{r \in \mathcal{R}} \{\lambda_r^{-1} w_r x_r\}. \tag{16}$$

The following lemma shows why this particular function is useful.

Lemma III.3: Assume that $A\lambda < \nu$ then there is a constant $c > 0$, such that for all $x \in \mathbb{R}_+^{\mathcal{R}}$ and for all $r^* \in \text{argmax}_{r \in \mathcal{R}} \varphi^r(x)$ we have

$$Q\varphi^{r^*}(x) = \lambda_{r^*}^{-1} w_{r^*} (\lambda_{r^*} - \mu_{r^*}^w(x)) \leq -c.$$

Proof: Suppose the the network state is x and let $r^* \in \text{argmax}_{r \in \mathcal{R}} \{\lambda_r^{-1} w_r x_r\}$ then for all $r \in \mathcal{R}$, we have that

$$\frac{w_r x_r}{\lambda_r} \leq \frac{w_{r^*} x_{r^*}}{\lambda_{r^*}}, \tag{17}$$

or equivalently that $\lambda_{r^*} w_r x_r \leq \lambda_r w_{r^*} x_{r^*}$. Now summing over all routes traversing a link $\ell \in \mathcal{L}$ we have that

$$\lambda_{r^*} \sum_{r \in \mathcal{R}} A_{\ell r} w_r x_r \leq w_{r^*} x_{r^*} \sum_{r \in \mathcal{R}} A_{\ell r} \lambda_r,$$

which one can rearrange to show that

$$\begin{aligned} \lambda_{r^*} &\leq \frac{w_{r^*} x_{r^*}}{\sum_{r \in \mathcal{R}} A_{\ell r} w_r x_r} \sum_{r \in \mathcal{R}} A_{\ell r} \lambda_r \\ &= \frac{w_{r^*} x_{r^*}}{\sum_{r \in \mathcal{R}} A_{\ell r} w_r x_r} \nu_\ell - \frac{w_{r^*} x_{r^*}}{\sum_{r \in \mathcal{R}} A_{\ell r} w_r x_r} (\nu_\ell - \sum_{r \in \mathcal{R}} A_{\ell r} \lambda_r). \end{aligned}$$

Given (17) and the stability condition one can easily show the existence of a positive lower bound, $\varepsilon > 0$, for the term on the right-hand side :

$$\begin{aligned} &\frac{w_{r^*} x_{r^*}}{\sum_{r \in \mathcal{R}} A_{\ell r} w_r x_r} (\nu_\ell - \sum_{r \in \mathcal{R}} A_{\ell r} \lambda_r) \\ \geq\ &\min_{\ell \in \mathcal{L}} \min_{r \in \mathcal{R}} \left\{ \frac{A_{\ell r} \lambda_r}{\sum_{r \in \mathcal{R}} A_{\ell r} \lambda_r} (\nu_\ell - \sum_{r \in \mathcal{R}} A_{\ell r} \lambda_r) \right\} = \varepsilon. \end{aligned}$$

Thus we have that

$$\lambda_{r^*} \leq w_{r^*} x_{r^*} f_\ell^{(1),w}(x) - \varepsilon$$

where we recognize a term corresponding to the fair share $f_\ell^{(1),w}(x)$ at link ℓ, see (4). Moreover since this is true for all ℓ we have that

$$\lambda_{r^*} \leq w_{r^*} x_{r^*} f^{(1),w}(x) - \varepsilon \tag{18}$$

where $f^{(1),w}(x) = \min_{\ell \in \mathcal{L}} f_\ell^{(1),w}(x)$ is the fair share at first level bottleneck links $\mathcal{L}^{(1),w}(x)$.

Now if r^* is a first level bottleneck route, *i.e.,* $r^* \in \mathcal{R}^{(1),w}(x)$, then $\mu_{r^*}^w(x) = w_{r^*} x_{r^*} f^{(1),w}(x)$, and it follows by (18) that $\lambda_{r^*} - \mu_{r^*}^w(x) \leq -\varepsilon$. If r^* is not a first level bottleneck route, we will show that its bandwidth allocation must

[6]Here $\nabla^2 W(x)$ denotes the $|\mathcal{R}| \times |\mathcal{R}|$ matrix with entries $\{\frac{\partial^2 W}{\partial x_r \partial x_s}(x), r, s \in \mathcal{R}\}$.

exceed $w_{r*} x_{r*} f^{(1),w}(x)$ and so again by (18) we have that $\lambda_{r*} - \mu^w_{r*}(x) \le -\varepsilon$.

We begin by showing that $f^{(2),w}(x) \ge f^{(1),w}(x)$. Suppose $\ell \in \mathcal{L} \setminus \mathcal{L}^{(1),w}(x)$ and note that

$$\sum_{r \in \mathcal{R}} A_{\ell r} w_r x_r f^{(1),w}(x) \le \sum_{r \in \mathcal{R}} A_{\ell r} w_r x_r f^{(1),w}_\ell(x) = \nu_\ell$$

so it follows that

$$\nu_\ell - f^{(1),w}(x) \sum_{r \in \mathcal{R}^{(1),w}(x)} A_{\ell r} w_r x_r \ge f^{(1),w}(x) \sum_{r \in \mathcal{R} \setminus \mathcal{R}^{(1),w}(x)} A_{\ell r} w_r x_r.$$

Rearranging terms and recalling the definition of fair share for the links in the second level of the bottleneck hierarchy we have that

$$\begin{aligned} f^{(2),w}_\ell(x) &= \frac{\nu^{(1),w}(x)}{\sum_{r \in \mathcal{R} \setminus \mathcal{R}^{(1),w}(x)} A_{\ell r} w_r x_r} \\ &= \frac{\nu_\ell - f^{(1),w}(x) \sum_{r \in \mathcal{R}^{(1),w}(x)} A_{\ell r} w_r x_r}{\sum_{r \in \mathcal{R} \setminus \mathcal{R}^{(1),w}(x)} A_{\ell r} w_r x_r} \ge f^{(1),w}(x). \end{aligned}$$

Thus $f^{(2),w}(x) = \min_{\ell \in \mathcal{L}} f^{(2),w}_\ell(x) \ge f^{(1),w}(x)$. Similarly it follows by induction that $f^{(i+1),w}(x) \ge f^{(i),w}(x)$, until the bottleneck hierarchy is exhausted.

Now since $\mu^w_{r*}(x) = w_{r*} x_{r*} f^{(j),w}(x)$ for some level j in the bottleneck hierarchy, it follows that $\mu^w_{r*}(x) \ge w_{r*} x_{r*} f^{(1),w}(x)$ and so $\lambda_{r*} - \mu^w_{r*}(x) \le -\varepsilon$. The lemma follows by selecting $c = \varepsilon \min_{r \in \mathcal{R}} \{\lambda_r^{-1} w_r\}$. ■

Theorem III.2: If $A\lambda < \nu$ then Markov chain $\{N(t),\ t \ge 0\}$ associated with weighted max-min fair bandwidth allocation is positive recurrent.

Proof: Based on Lemma III.3, and the technique used in Lemma III.2, it should be clear that an appropriately smooth Lyapunov function W can be constructed from V in (16). Positive recurrence then follows as in Theorem III.1. ■

Note that since max-min fairness is a special case of weighted max-min fairness, Theorem III.2 establishes the stability of both. The two different Lyapunov functions we have introduced, based on links and routes, may be of interest in further studies of performance. These results establish that $A\lambda < \nu$ is a *sufficient* condition for stability. In fact, it is a *necessary* condition. Say there exists a link ℓ such that $\sum_{r \in \mathcal{R}} A_{\ell r} \lambda_r > \nu_\ell$. Clearly such a link in isolation is unstable, *i.e.,* on average will tend to drift off to infinity. When the link is incorporated within a network, the situation can in fact only get worse, since other links may slow down the departures for connections on ℓ.

C. Example network

In this example we consider max-min fair bandwidth allocation for the network shown in Fig. 1 – it consists of two routes $\mathcal{R} = \{1, 2\}$ three links, $\mathcal{L} = \{1, 2, 3\}$, as shown in the figure. Fig. 2 shows the vector field $\lambda - \mu^m(x)$ corresponding to the case with $\lambda = (1.5, 1.5)$ and $\nu = (5, 6, 4)$. We have shown

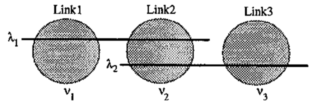

Fig. 1. Example network with three links and two routes.

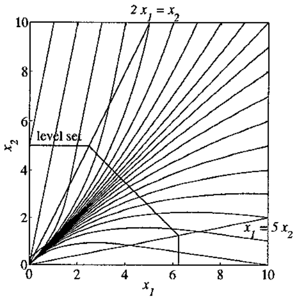

Fig. 2. A vector field of the example network.

the boundaries $x_1 = 5x_2$ and $2x_1 = x_2$ between three cones $C_{\{1\}}$, $C_{\{2\}}$ and $C_{\{3\}}$ corresponding to links 1,2 or 3 being bottlenecks respectively. Also shown on the figure is a level set of the function V. From the figure it is clear that on each cone the network's dynamics push inwards, *i.e.,* have negative drift with respect to V. By smoothing V as in Lemma III.2 we obtain a Lyapunov function W from which the stability of the system follows.

D. Stability under a state dependent weighted max-min fair control policy

In this section we briefly consider a simple extension to our model with weighted max-min fair allocation, wherein a control policy is implemented by letting the weights depend on the network state. Let $w = (w_r : \mathbb{Z}_+^{\mathcal{R}} \to \mathbb{R}_+, r \in \mathcal{R})$ now denote functions where $w(n) = (w_r(n), r \in \mathcal{R})$ are understood to be the weights associated with each route when the network state is n. Assume that when the system is in state n bandwidth is allocated to routes according to a weighted max-min fair allocation with weights $w(n)$. Let $\mu^{w(n)}(n) = (\mu_r^{w(n)}(n), r \in \mathcal{R})$ denote the bandwidths allocated to each route in the network when the state is n. Our interest in this type of model, was motivated by work on stability of Generalized Processor Sharing networks [18]. Without delving into the details of their model, we remind the reader that in such networks a connection is assigned a weight at each node (representing a queue) which determines the fraction of the available capacity it receives at that node. The authors showed the queue/delay stability of non-acyclic networks of this type when connections received a *consistent relative treatment.* By analogy here, we will say that a state dependent weight based control policy gives routes a *uniform relative treatment* if $\forall n \in \mathbb{Z}_+^{\mathbb{R}}$ and $r, s \in \mathcal{R}$,

$$\frac{\lambda_s}{\lambda_r} \ge \frac{n_s w_s(n)}{n_r w_r(n)}. \tag{19}$$

An example one on such control policy would be $w_r(n) = \lambda_r / n_r$ for $n_r \neq 0$. Thus upon admitting or tearing down a connection along a given route the network controller would need to adjust the weight associated with that route. The following lemma shows that subject to the natural stability condition, a weight based control policy that gives routes uniform relative treatment is stable.

Lemma III.4: Assume $A\lambda < \nu$ and a weight based control policy $w(\cdot)$ that gives routes a uniform relative treatment is used to allocate bandwidth in the network. Then $\forall n \in \mathbb{Z}_+^{\mathcal{R}}$ and $\forall r \in \mathcal{R}$ such that $n_r > 0$ we have

$$\lambda_r < \mu_r^{w(n)}(n).$$

It follows that the network is positive recurrent.
The proof of this lemma is almost identical to that of Lemma III.3. Positive recurrence follows since the number of connections on every route has negative drift if it is not empty.

IV. Performance

In this section we use simulation to briefly evaluate network performance, in terms of average connection delays, based on our model. This type of metric might be of interest in dimensioning networks to provide a reasonable call-level quality of service. One might also wish to design network control mechanisms, *e.g.,* assign priorities (weights) to different routes, or spread the call level loads across the network in a manner that improves the individual or overall delays experienced in the network.

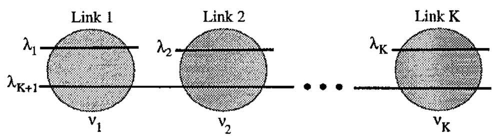

Fig. 3. A network for simulations.

We shall consider a network consisting of K links in series, see Fig. 3. A long route traverses each link in the network, while short, single link routes, model "cross traffic." To investigate the degradation in performance as connections traverse an increasing number of links we simulated several configurations where $K = 2, 3, 4$ and 5. Herein we will only present results for a symmetric, moderate load, scenario where the arrival rate on each route was 2 connections/sec and the capacity of each link was 6 connections/sec. We simulated max-min, weighted max-min, and proportionally fair bandwidth allocation mechanisms in order to assess their impact on connection delays. In the case of weighted max-min fairness, short and long connections were given weights 1 and 2 respectively, *i.e.,* priority was given to connections traversing several links as they are likely to experience the poorest performance. Figures 4 - 6, show the average connection delays on short routes, long routes, and overall, for the various types of fairness and as the number of links K in the network increases. These results suggest the trends discussed below.

We first contrast the performance of max-min fair bandwidth allocation, which strives to maximize the worst case individual performance versus proportional fairness which strives to maximize the overall network utility. The latter tends to give more bandwidth to connections crossing a small number of links, as they are more efficient in terms of their resource requirements. As a result long routes may linger in the network possibly degrading the overall performance. This effect is aggravated as the number of links in the network increases. For example, for $K = 5$, the relative change in delays for proportional versus max-min fair bandwidth allocation is -10 % on short routes, +46 % on long routes, and +5% overall. This suggests that for networks supporting a larger number of longer routes one might find that max-min outperforms proportionally fair allocations in terms of delays on long routes as well as overall delays. Finally note, that as the number of links increase, proportional fairness leads to a surprisingly flat average delay on short routes, while long routes see a linear growth in average delay, see Figs. 4 and 5.[7] This suggests that proportional fairness may provide a clean performance differentiation among routes that have different lengths.

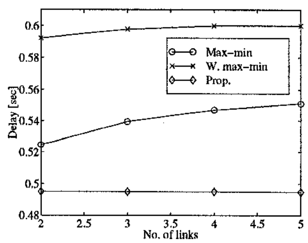

Fig. 4. Average delay on short routes.

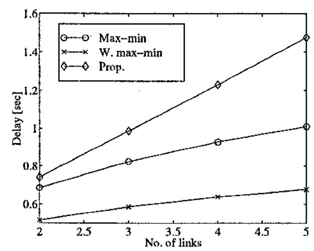

Fig. 5. Average delay on long routes.

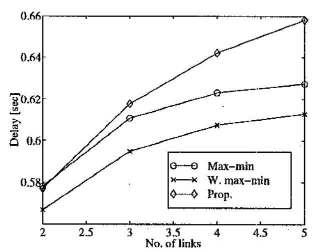

Fig. 6. Average overall delay.

Next, we consider the impact that using a weighted max-min fair bandwidth allocation will have on delays, if weights are selected so as expedite connections on long routes. Continuing with our example, when $K = 5$, the relative change in delays for the weighted versus the max-min fair bandwidth allocation is +9 % on short routes, -33 % on long routes, and -2 % overall. Thus, one can not only dramatically improve the delays experienced on long routes, but also marginally improve the overall performance.

These results exhibit the potential impact that a the fairness criterion selected by designers may have on network performance. However, a better characterization of network perfor-

[7]The overall delay is not linear since it is an average of delays on short and long routes. Since the relative total load on short versus long routes is increasing with K, the overall delay behavior is not linear.

mance and tools to 'optimally' select weights, or route connections, will need to be developed if a call level quality of service such as that considered here is deemed important in future networks. Also note that one could in theory introduce weights on a proportionally fair allocation in order to also enhance the performance seen on long connections. Hence our results do not suggest that a particular mechanism is best, we merely suggest that a consideration of these issues might be warranted.

V. Could the Internet be Unstable ?

In this paper we have considered the stability and performance of an idealized model for a network supporting services that adjust their transmissions to network loads. The model is only a *rough caricature* of the Internet today, in that it assumes TCP operates efficiently by immediately achieving an *average* throughput related to a weighted proportionally fair bandwidth allocation. [8] So a connection's throughput is dictated by a weighted allocation of resources at congested or bottleneck links. Average round-trip time experienced by connections and loss rate can be captured by weights given to connections which in turn impact the equilibrium throughput achieved by TCP connections. This model parallels the one proposed and validated via simulation in [16]. We also assume that packets associated with a given TCP connection typically follow the same route, and connections send data in a greedy manner and depart. Subject to these, perhaps fanciful assumptions, one can show that network stability cannot be guaranteed unless the connection-level offered loads do not exceed the network's link capacities.

While this result is not entirely surprising, it presents an interesting architectural dilemma for future networks. Since routing algorithms on the Internet base their decisions on short term measures, *i.e.,* are not explicitly tracking the long-term averages required to assess the connection level offered loads, there is no reason to believe that the Internet would satisfy a connection level stability requirement. Instability would be perceived by users as an unacceptably low throughput, or inordinate delays, and typically cause them to abandon, thus in some sense solving the problem. To avoid such extremes one might overprovision the network. Unfortunately, this may result in a network which is still unstable, resulting in sporadic long lasting congestion events that are challenging to explain.

Currently we are researching using methodologies similar to those we have used to prove stability, to explore performance issues and consider in more depth the compromises one might make to achieve good performance at the connection level. It would of course be interesting to look at congestion patterns on the Internet today and attempt to explain them in terms of a connection-level instability. However, given the typically non-stationary demands on today's networks and the detailed data that would be required to provide a conclusive answer to this question this appears to be a challenging task.

[8] For a single congested link, weighted max-min fair or weighted proportionally fair allocation model TCP appropriately [16], [7]. We believe that weighted max-min fair allocation can be adopted as a network model if the weights are selected to reflect round-trip delays and TCP dynamics.

References

[1] S. P. Abraham and A. Kumar. A stochastic approximation approach for max-min fair adaptive rate control of abr sessions with MCRs. In *Proc. IEEE INFOCOM*, 1998.

[2] E. Altman, F. Bacelli, and J-C. Bolot. Discrete-time analysis of adaptive rate control mechanism. In *Proc. 5th Int. Conference on Data and Communications*, pages 121–40, 1993.

[3] L. Benmohamed and S. M. Meerkov. Feedback control of congestion in packet switching networks the case of a single congested node. *IEEE/ACM Trans. on Networking*, 1(6):693–709, Dec. 1993.

[4] D. Bertsekas and R. Gallager. *Data Networks*. Prentice Hall, 1992.

[5] F. Bonomi and K.W. Fendick. The rate-based flow control framework for the Available Bit Rate ATM service. *IEEE Network Mag.*, pages 25–39, Mar/Apr. 1995.

[6] F. Bonomi, D. Mitra, and J. B. Seery. Adaptive algorithms for feedback-based flow control in high-speed wide-area ATM networks. *IEEE J. Select. Areas Commun.*, 13(7):1267–83, Sept. 1995.

[7] J. Crowcroft and P. Oechslin. Differentiated end-to-end internet services using a weighted proportional fair sharing TCP. *Computer Communication Review*, 28(3), 1998.

[8] D. Down and S. P. Meyn. Piecewise linear test functions for stability and instability of queueing networks. *Queueing Systems*, 27:205–226, 1997.

[9] G. Fayolle, V. A. Mayshev, and M. V. Menshikov. *Topics in Constructive Theory of Countable Markov Chains*. Cambridge University Press, Cambridge and New York, 1995.

[10] E. M. Gafni and D. Bertsekas. Dynamic control of session input rates in communication networks. *IEEE Trans. Automatic Control*, 29:1009–16, 1984.

[11] V. Jacobson and M. Karels. Congestion avoidance and control. In *Proc. ACM SIGCOM*, pages 314–29, Aug. 1988.

[12] F. Kelly. Charging and rate control for elastic traffic. *European Transactions on Telecommunications*, 8:33–37, 1997.

[13] F. P. Kelly, A. K. Mauloo, and D. K. H. Tan. Rate control in communication networks shadow prices, proportional fairness and stability. *Journal Oper. Res. Soc.*, 15(49):237–55, 1998.

[14] T.-J. Lee. Flow control and network dimensioning in multiservice ATM networks. *Ph.D. Dissertation*, E.C.E. Dept, University of Texas, Austin, 1999.

[15] T.-J. Lee and G. de Veciana. A decentralized framework to achieve max-min fair bandwidth allocation for ATM networks. In *Proc. IEEE GLOBECOM*, 1998.

[16] M. Mathis, J. Semke, J. Mahdavi, and T. Ott. The macroscopic behavior of the TCP congestion avoidance algorithm. *Computer Communication Review*, 27(3):67–82, 1997.

[17] S. P. Meyn and R. L. Tweedie. *Markov chains and stochastic stability*. Springer-Verlag, Berlin, 1993.

[18] A.K. Parekh and R.G. Gallager. A generalized processor sharing approach to flow control in integrated services networks: The multiple node case. *IEEE/ACM Trans. Networking*, July 1994.

[19] S. Shenker. Fundamental design issues for the future internet. *IEEE J. Select. Areas Commun.*, 13(7):1176–88, Sept. 1995.

[20] C.-F. Su, G. de Veciana, and J. Walrand. Explicit rate flow control for ABR services in ATM networks. *IEEE/ACM Trans. Networking*, Jan. 1998. Submitted.