

Network Stability under Alpha Fair Bandwidth Allocation with General File Size Distribution

Fernando Paganini*, Ao Tang[†], Andrés Ferragut* and Lachlan L. H. Andrew[‡]

* Universidad ORT, Montevideo, Uruguay. [†] Cornell University, Ithaca, NY.

[‡] Swinburne University of Technology, Australia.

Abstract—Rate allocation among a fixed set of end-to-end connections in the Internet is carried out by congestion control, which has a well established model: it optimizes a concave network utility, a particular case of which is the alpha-fair bandwidth allocation. This paper studies the slower dynamics of connections themselves, that arrive randomly in the network and are served at the allocated rate. It has been shown that under the condition that the mean offered load at each link is less than its capacity, the resulting queueing system is stochastically stable, for the case of exponentially distributed file-sizes. The conjecture that the result holds for general file-size distributions has remained open, and is very relevant since heavy-tailed distributions are often the best models of Internet file sizes.

In this paper, building on existing fluid models of the system, we use a partial differential equation to characterize the dynamics. The equation keeps track of residual file size and therefore is suitable for general file size distributions. For alpha fair bandwidth allocation, with any positive alpha parameter, a Lyapunov function is constructed with negative drift when the offered load is less than capacity. With this tool we answer the conjecture affirmatively in the fluid sense: we prove asymptotic convergence to zero of the fluid model for general file-size distributions of finite mean, and finite-time convergence for those of finite $p > 1$ moment. In the stochastic sense, we build on recent work that relates fluid and stochastic stability subject to a certain light-tailed restriction. We further provide the supplementary fluid stability argument to establish the conjecture for this class that includes phase-type distributions. Results are supplemented by illustrative network simulations at the packet level.

I. INTRODUCTION AND PRIOR WORK

A fundamental step in the understanding of resource allocation mechanisms in the Internet has been the formulation by Kelly et al. [10] of congestion control in terms of network utility maximization. In a scenario of a fixed number of connections across routes in a network, this approach characterizes an equilibrium and leads to the formulation of dynamic, distributed methods to achieve it. An interesting class of utility functions is the “ α -fair” family of Mo and Walrand [18]; by varying the α parameter it encompasses various notions of flow-level fairness, in particular proportional fairness ($\alpha = 1$) and max-min fairness ($\alpha \rightarrow \infty$). These notions are commonly used to describe various current or proposed network protocols (e.g., [1], [25]).

However, this analysis with fixed numbers of connections does not capture the reality that flows come and go in the network, a situation better modeled through stochastic processes. This issue was identified by Roberts and Massoulié [21], who studied queueing systems with random arrivals and workloads, and a processor sharing discipline where service

rates depend on bandwidth allocation, assumed to occur at a faster time-scale. This leads to a basic *stability* question, first posed by De Veciana et al. [7]: under which connection level demands (job arrival rate and mean workload) is the resulting queueing process stable? The answers given in [7] apply to Poisson arrivals and exponentially distributed job sizes, and max-min fair or proportionally fair bandwidth allocation. In this case the numbers of connections per route form a Markov chain, which is shown to be stable (i.e., ergodic) under the natural stability condition: namely, that the *mean load in each link of the network is strictly less than the link capacity*. In a subsequent paper by Bonald and Massoulié [3], these results were generalized to the α -fair case; other utility functions are considered in [26]. Further extensions include relaxing the time-scale separation [14] and relaxing the model of fixed-capacity links [15].

We note that the natural condition is not sufficient for all allocation policies, such as when the network seeks to maximize instantaneous throughput ($\alpha = 0$, see [3]) or under certain forms of prioritization. For a demonstrative example we refer to Section VI. Verloop et al. [24] show that the form of scheduling that is optimal for a single link (shortest remaining processing time, SRPT) can cause networks to be unstable even if the maximum individual link utilization approaches 0. The intuition is that a flow on a multi-link path can be bottlenecked by flows on any one of its links, leaving other links under-utilized at some instant; if the bottleneck link shifts as flows arrive and depart, the multi-link path could be unstable even though the mean load on each link is strictly less than its capacity.

Therefore the cited stability results establish a non-trivial fact. They are, however, critically restricted by the assumption of exponentially distributed inter-arrival times and file sizes. The latter is particularly unsatisfactory, since it has been observed that file sizes in the Internet follow a heavy tailed Pareto-type distribution [5]. This has motivated recent efforts in extending the stability results for general file-size distributions. Removing the exponential file-size assumption is “*well-known to be a difficult problem*” [14]; without it, the number of ongoing connections is no longer a Markov state. Some existing partial results are the following: [16] has showed that for $\alpha = 1$, the stability result can be generalized to an appropriate Jackson-type routing scheme, thus providing the tool to establish the condition for phase-type file distributions; [12] gives a result for phase-type file distribution and general α , in two particular network topologies through Lyapunov

functions obtained numerically; [27] proves stability for quite general arrival and service times in both stochastic and worst-case settings, for a variant in which at most one flow on each route is allocated capacity at a given time.

A strategy that has proven relevant for this problem is the use of *fluid limits*, already invoked in [3] and further studied in [11], for the exponential case. The extension to general distributions is developed by Gromoll and Williams [8], [9], as briefly reviewed in Section II. Based on these models, [4] obtains a stability result for general file size distributions of bounded support, for a sufficiently small α . In [8] it is also established that the stability condition guarantees the fluid model is stable in the special cases of linear and tree networks. In a recent thesis by Lee [13], the connection between fluid and stochastic stability was investigated, for the models of [9]: it was found that under a certain “light-tailed” restriction on the distribution (which includes phase-type, but not Pareto) fluid stability indeed implies stochastic stability. The main missing piece is then to establish, for these models, that fluid stability holds under the natural condition on the loads. This is the subject of the present paper.

First, in Section II we build on the results of [8], [9] and obtain a fluid model of the process in terms of a partial differential equation (PDE) for the distribution of remaining job workloads. A main step in Section III is to show that the natural stability condition allows for the construction of a Lyapunov function with negative drift along trajectories of the PDE, regardless of the job size distribution and the underlying network topology, and for any value of α . In Section IV we use this tool to prove fluid stability results: asymptotic convergence to zero of the state in the general case, and finite-time convergence under a mild moment condition satisfied with great generality, including the heavy-tailed case. Some of the above material appeared in the conference paper [20], although the finite-time result is new. In Section V we reconsider the stability question in the original, stochastic queueing model. We show that for the distribution class considered in [13] the required fluid stability assumptions are indeed satisfied, and thus the problem is closed for this class. Section VI provides a set of packet-level simulations to demonstrate the validity of the models and the relevance of the results to real networks.

II. PROBLEM FORMULATION AND PARTIAL DIFFERENTIAL EQUATION MODEL

The problem under consideration is the stability of a queueing system in which customers are flows that arrive at various routes, and are served according to rates allocated by a congestion control algorithm. The latter is modeled through the network utility maximization problem

$$\max_{m: z_m > 0} \sum z_m U_m \left(\frac{\varphi_m}{z_m} \right), \text{ subject to } \sum_m R_{lm} \varphi_m \leq c_l. \quad (1)$$

Here m denotes the route, z_m the number of connections in the route, and φ_m the total allocated rate of all connections of route m ¹. Matrix R is the routing matrix ($R_{lm} = 1$ if

route m uses link l and zero otherwise) and $c = (c_l)$ the vector of link capacity constraints. The utility function U_m is assigned to each connection as a function of its per-flow rate $x_m = \varphi_m/z_m$. For background on these models we refer to [22]. In this paper we focus on the “ α -fair” utility functions introduced in [18], where $U'_m(x) = x^{-\alpha}$ for some strictly positive α .

From a modeling perspective, this does not assume any particular scheduling discipline, buffer size or active queue management (AQM) settings. When all flows see the same congestion price, such as under first-in-first-out scheduling, the capacity allocation is determined by the senders, with congestion control algorithms such as Reno [1] and FAST [25] approximating α -fairness with $\alpha = 2$ and $\alpha = 1$ respectively. Similarly, under round-robin scheduling flows achieve approximate max-min ($\alpha = \infty$) fairness for a wide range of congestion control algorithms.

Given $z = (z_m)$, it is assumed the congestion control sets $\varphi = (\varphi_m)$ to the optimum of (1); we assume separation of time-scales, i.e. the mapping $z \mapsto \varphi$ is instantaneous. This assumption is standard in most of the literature, with the exception of [14]; we will require this to maintain the tractability of the model.

Remark 1. *A property of α -fair utility functions (with common α across routes) is that the resource allocation is invariant under scaling: namely, if all z_m are scaled by a common factor r , the resulting φ_m do not change.*

Consider now a network where flows arrive at route m through a stochastic process of mean intensity $\lambda_m > 0$, and a general distribution of the file sizes: let $G_m(\sigma)$ be the complementary cumulative distribution function (CCDF), i.e. the probability that the file size is greater than σ . The mean file-size is given by

$$\frac{1}{\mu_m} = \int_0^\infty \sigma [-dG_m] = \int_0^\infty G_m(\sigma) d\sigma,$$

assumed to be finite. At any given time, active flows are served with the rate φ_m/z_m that results from (1).

The arrival process is independent of the number of flows at all earlier points in time. In particular, this excludes “closed loop” traffic models in which the completion of flow may cause a user to request another. The present “open loop” model is appropriate when each independent user initiates a single flow, or stream of contiguous flows.

The aim is to prove that if the loads $\rho_m := \lambda_m/\mu_m$ strictly satisfy the network capacity constraints,

$$\sum_m R_{lm} \rho_m < c_l \quad \forall l, \quad (2)$$

then the queueing system is stable². This result, if true in generality, implies for practitioners that provisioning a network to accommodate the mean loads suffices for stabilization, provided the underlying resource allocation corresponds to some form of α -fairness.

¹Only routes with a nonzero number of connections participate in the resource allocation, the rest receive $\varphi_m = 0$.

²We emphasize that in this model queue occupancy is the number of flows on a path, not the amount of data in a buffer. The stability issue thus decouples from buffer size considerations.

The classically studied case [7], [3] is when the arrival process is Poisson, and the file distribution exponential. In that case the process is a Markov chain with state $z = (z_m)$, and stability means positive recurrence [2]. For the general case, the system state requires substantially more information, as discussed later on. We now turn our attention to fluid models for this problem, which are the basis of our stability studies.

A. Fluid model

We first recall the fluid model for the exponential case, following [3], [11]. This is obtained in the limit by scaling time and the initial condition of the process, leaving fixed the network capacity and the external load. Let r be the scaling parameter, and define $z^r(0) = rz(0)$, where $\|z(0)\| = 1$ in a suitable vector norm. If $z^r(t)$ is the resulting stochastic process as described above, the fluid limit is defined by

$$Z(t) = \lim_{r \rightarrow \infty} \frac{z^r(rt)}{r}.$$

Invoking the strong law of large numbers, [3] writes the following ordinary differential equation model:

$$\frac{dZ_m}{dt} = \lambda_m - \varphi_m(t)\mu_m \quad (3)$$

for each m . Here $\varphi_m(t)$ corresponds to the service rate with re-scaled time³. Due to the scale invariance of the resource allocation (see Remark 1) the fluid versions of φ and Z are still related by the the analog of (1), i.e. $\varphi(t)$ is the maximizer of

$$\max_{m: Z_m > 0} \sum Z_m U_m \left(\frac{\varphi_m}{Z_m} \right), \text{ s.t. } \sum_m R_{lm} \varphi_m \leq c_l. \quad (4)$$

We now state a basic inequality that will characterize the resource allocation. It follows directly from the fact that at the optimal point of (4), the feasible set must be inside a negative half-space defined by the gradient vector.

Lemma 1. *Let (φ_m) be the vector of rates that optimizes (4), and (ψ_m) another vector of rates satisfying the constraints $\sum_m R_{lm} \psi_m \leq c_l$. Then:*

$$\sum_{m: Z_m > 0} U'_m \left(\frac{\varphi_m}{Z_m} \right) (\psi_m - \varphi_m) \leq 0. \quad (5)$$

B. Behavior at $Z_m = 0$

The model (3) applies while $Z_m > 0$, it must be appropriately modified to include the behavior when $Z_m(t)$ reaches zero.

To understand the behavior in this case, consider the following example: a network with a single link of capacity c , and a single class m , such that $\rho_m = \lambda_m/\mu_m < c$. In this case, whenever $Z_m(t) > 0$ we have $\varphi_m = c$, therefore $Z_m(t)$ decreases linearly to zero following (3). When Z_m reaches zero, then by definition φ_m suddenly becomes zero (a discontinuity); but then the arrivals term λ_m would move Z_m

again into positive values: this ‘‘chattering’’ behavior around $Z_m = 0$ is modeled by replacing (3) with

$$\frac{dZ_m}{dt} = \begin{cases} \lambda_m - \varphi_m(t)\mu_m & \text{if } Z_m(t) > 0; \\ 0 & \text{if } Z_m(t) = 0, \end{cases} \quad (6)$$

which allows the state to remain at $Z_m = 0$.

In the previous example, staying at zero is the steady-state equilibrium behavior. In more complex networks, the trajectory may be such that Z_m spends an interval of time at zero and later departs to $Z_m > 0$; this may happen if the capacity available for route m later moves below ρ_m due to competition from other routes that share a bottleneck with m . In that case, equation (6) exhibits non-uniqueness of solutions, since it still allows the solution $Z_m \equiv 0$. While this solution should be ruled out for physical reasons, the model (6) is less descriptive and allows it. The same issue appears with the PDE model (13) to be introduced below. In terms of stability studies, however, this issue does not create a difficulty: if we show that *all* solutions to this broader model converge to zero, in particular the ‘‘physical’’ one will.

C. PDE Model

In the general distribution case, bandwidth allocation is still a function only of the numbers of flows per route, $Z = (Z_m)$. However, once we remove the memoryless property of the exponential distribution, characterizing the network state requires keeping track of residual file-sizes, not just their number. Furthermore, the resource allocation per route is a processor sharing discipline, where all flows present receive equal service. This complicates the description since we must keep track of residual file sizes of *all* flows. In order to proceed, we look at the problem in more detail by modeling, in a fluid setting, the evolution of the residual file distribution.

Let $F_m(t, \sigma)$ ($t \geq 0$, and $\sigma \geq 0$) be the number of class (route) m files at time t with residual file size larger than σ , in the fluid limit. $F_m(t, \sigma)$ is a finer descriptor of the system than $Z_m(t)$, indeed the definition implies

$$Z_m(t) = F_m(t, 0). \quad (7)$$

We now model the evolution of $F_m(t, \sigma)$ through the following partial differential equation:

$$\frac{\partial F_m(t, \sigma)}{\partial t} = \frac{\partial F_m(t, \sigma)}{\partial \sigma} \frac{\varphi_m(t)}{Z_m(t)} + \lambda_m G_m(\sigma). \quad (8)$$

The above equation holds under the assumption that $Z_m(t) > 0$; again it will be suitably complemented for $Z_m = 0$ in (13) below.

Note that (8) reduces to (3) in the exponential file size distribution case, $G_m(\sigma) = \exp(-\mu_m \sigma)$. This can be readily checked by using $F_m(t, \sigma) = Z_m(t) \exp(-\mu_m \sigma)$ in (8), which reduces it to (3).

D. An intuitive derivation of the PDE

At time $t + dt$, jobs that have residual file size at least σ come from two sources:

³We refer to [11] for more extensive details on this type of scaling.

- New arrivals between t and $t + dt$ of size greater than σ . With arrival rate λ_m , we have $\lambda_m dt G_m(\sigma)$ such jobs, in the fluid limit.
- Files already present at time t , which had at that time a residual size of at least $\sigma + \frac{\varphi_m(t)}{Z_m(t)} dt$. Note each file receives a service rate $\varphi_m(t)/Z_m(t)$.

Therefore

$$F_m(t + dt, \sigma) = F_m\left(t, \sigma + \frac{\varphi_m(t)}{Z_m(t)} dt\right) + \lambda_m G_m(\sigma) dt.$$

Subtracting $F_m(t, \sigma)$ from both sides and dividing by dt gives

$$\frac{F_m(t + dt, \sigma) - F_m(t, \sigma)}{dt} = \frac{F_m\left(t, \sigma + \frac{\varphi_m(t)}{Z_m(t)} dt\right) - F_m(t, \sigma)}{dt} + \lambda_m G_m(\sigma).$$

In the limit when $dt \rightarrow 0$ we obtain (8).

E. A formal justification based on [9]

We now explain how to relate (8) to the rigorous fluid limit set up in Gromoll and Williams [9].

In this formulation, the system state is characterized by a time-dependent, positive *measure* $\zeta_m(t)$ for each class (route) m . The measure is defined over the positive real numbers, representing distribution of residual workload. In particular, in the stochastic model $\zeta_m(t)$ at any given time is a finite sum of Dirac deltas, located at the sizes of remaining workloads for currently active jobs. The integral of this measure is the number of active jobs $Z_m(t)$.

In the fluid limit under appropriate scaling, the limiting measure $\zeta_m(t)$ satisfies (for all t except a set of Lebesgue measure zero) the following:

$$\frac{d}{dt} \langle f, \zeta_m(t) \rangle = \begin{cases} -\frac{\varphi_m(t)}{Z_m(t)} \langle f', \zeta_m(t) \rangle + \lambda_m \langle f, \nu_m \rangle, & \text{for } Z_m > 0; \\ 0, & \text{for } Z_m = 0. \end{cases} \quad (9)$$

This equation coincides with (5.62) in [9], modulo notational changes. Here the measure ν_m represents the probability distribution of arriving jobs; $f(\sigma)$ is an arbitrary bounded and continuously differentiable test function in the class

$$\mathcal{C} = \{f \in C_b^1(\mathbb{R}_+), f(0) = f'(0) = 0\};$$

and $\langle f, \nu \rangle := \int_0^\infty f(\sigma) d\nu$.

In this model, the probability of an arriving job being larger than σ , and the number of jobs at time t with residual workload greater than σ , are represented by

$$G_m(\sigma) := \int_\sigma^\infty d\nu_m; \quad F_m(t, \sigma) := \int_\sigma^\infty d\zeta_m(t).$$

To derive the PDE we assume that the measures ν_m and $\zeta_m(t)$ are absolutely continuous with respect to Lebesgue measure. In particular,

$$d\nu_m = -G'_m(\sigma) d\sigma, \quad d\zeta_m(t) = -\frac{\partial F_m(t, \sigma)}{\partial \sigma} d\sigma.$$

By integration by parts we have the following identities:

$$\begin{aligned} \langle f, \nu_m \rangle &= -\int_0^\infty f(\sigma) G'_m(\sigma) d\sigma \\ &= -f(\sigma) G_m(\sigma) \Big|_{\sigma=0}^\infty + \int_0^\infty f'(\sigma) G_m(\sigma) d\sigma. \\ \langle f, \zeta_m(t) \rangle &= -\int_0^\infty f(\sigma) \frac{\partial F_m(t, \sigma)}{\partial \sigma} d\sigma \\ &= -f(\sigma) F_m(t, \sigma) \Big|_{\sigma=0}^\infty + \int_0^\infty f'(\sigma) F_m(t, \sigma) d\sigma. \end{aligned}$$

Due to the definition of the class \mathcal{C} , the incremental terms above vanish, which turns (9) into

$$\frac{d}{dt} \int_0^\infty f'(\sigma) F_m(t, \sigma) d\sigma = \frac{\varphi_m}{Z_m} \int_0^\infty f'(\sigma) \frac{\partial F_m(t, \sigma)}{\partial \sigma} d\sigma + \lambda_m \int_0^\infty f'(\sigma) G_m(\sigma) d\sigma,$$

for the case $Z_m > 0$. Assuming the differentiation with respect to t on the left can be interchanged with integration, the above yields

$$\int_0^\infty f'(\sigma) \mathcal{D}[F_m(t, \sigma)] d\sigma = 0,$$

where $\mathcal{D}[\cdot]$ is the differential operator given by

$$\mathcal{D}[F_m] := \frac{\partial F_m}{\partial t} - \frac{\partial F_m}{\partial \sigma} \frac{\varphi_m(t)}{Z_m(t)} - \lambda_m G_m(\sigma).$$

Since $f'(\sigma)$ is a free continuous function, we must have $\mathcal{D}[F_m] \equiv 0$, i.e. (8).

Remark 2. [9] contains integral versions of (9), which define the PDE in the distributional sense beyond the above assumptions (absolute continuity, differentiation under the sign). We will not pursue this issue here, and assume a solution $F(t, \sigma)$ to (8) of enough smoothness to allow for a classical treatment.

F. Locally integrating the PDE

It will also be convenient to have an integral form for the PDE solution. We state the following.

Proposition 2. Consider a solution F of (8) such that $Z_m(t) > 0$ in $t \in [t_0, t_1]$. Denote $x_{F,m}(t) = \frac{\varphi_m(t)}{Z_m(t)}$ for this F . Then for any $(t, \sigma) \in [t_0, t_1] \times \mathbb{R}_+$ we have

$$\begin{aligned} F_m(t, \sigma) &= F_m\left(t_0, \sigma + \int_{t_0}^t x_{F,m}(\tau) d\tau\right) \\ &+ \lambda_m \int_{t_0}^t G_m\left(\sigma + \int_T^t x_{F,m}(\tau) d\tau\right) dT. \end{aligned} \quad (10)$$

Proof: For simplicity we omit the subindex m . Note from continuity of $Z(t)$ and boundedness of $\varphi(t)$ that $x_F(t)$ is bounded in the interval $[t_0, t_1]$, so the integrals above are well defined, and finite.

Define $\tilde{F}(t, \sigma)$ to be the right-hand side of (10). It remains to show that $F(t, \sigma) = \tilde{F}(t, \sigma)$ for all $t \geq t_0$ and $\sigma \geq 0$. By

differentiation,

$$\begin{aligned} \frac{\partial \tilde{F}t, \sigma}{\partial t} &= \frac{\partial F}{\partial \sigma} \left(t_0, \sigma + \int_{t_0}^t x_F(\tau) d\tau \right) x_F(t) \\ + \lambda G(\sigma) + \lambda \int_{t_0}^t G' \left(\sigma + \int_T^t x_F(\tau) d\tau \right) dT x_F(t) \\ &= \frac{\partial \tilde{F}(t, \sigma)}{\partial \sigma} x_F(t) + \lambda G(\sigma). \end{aligned}$$

This has a similar form to (8), differing in the use of $x_F(t)$ defined in terms of F , in place of $x_{\tilde{F}}(t)$ defined in terms of \tilde{F} . The similarity suffices to show that indeed $\tilde{F} \equiv F$, as follows. Define

$$\Delta(t, \sigma) := F(t, \sigma) - \tilde{F}(t, \sigma),$$

which is a solution to the homogeneous equation

$$\frac{\partial \Delta(t, \sigma)}{\partial t} = \frac{\partial \Delta(t, \sigma)}{\partial \sigma} x_F(t), \quad \Delta(t_0, \sigma) \equiv 0,$$

where the boundary condition holds since $\tilde{F}(t_0, \sigma) = F(t_0, \sigma)$. Consider now the mapping

$$\Gamma(t, \sigma) = \left(t, \sigma + \int_{t_0}^t x_F(\tau) d\tau \right).$$

Γ is a diffeomorphism between $[t_0, t_1] \times \mathbb{R}_+$ and the set

$$\mathcal{R} := \left\{ (t, \sigma) : t \in [t_0, t_1], \sigma \geq \int_{t_0}^t x_F(\tau) d\tau \right\}.$$

By composition define a function on \mathcal{R} ,

$$\Phi(t, \sigma) := \Delta \circ \Gamma^{-1}(t, \sigma) = \Delta \left(t, \sigma - \int_{t_0}^t x_F(\tau) d\tau \right).$$

By differentiation we obtain for any $(t, \sigma) \in \mathcal{R}$,

$$\frac{\partial \Phi(t, \sigma)}{\partial t} = \frac{\partial \Delta(t, \sigma)}{\partial t} - \frac{\partial \Delta(t, \sigma)}{\partial \sigma} x_F(t) = 0. \quad (11)$$

Also, $\Phi(t_0, \sigma) \equiv 0$. Note that for any $(t, \sigma) \in \mathcal{R}$, the line segment $[t_0, t] \times \{\sigma\}$ is contained in \mathcal{R} , therefore by integration of (11) in this segment we have $\Phi(t, \sigma) \equiv 0$ on \mathcal{R} . Therefore $\Delta = \Phi \circ \Gamma$ is also identically zero, which establishes (10). \square

The solution form (10) also applies to *open* intervals in which $Z_m(t) > 0$, as is now shown.

Corollary 3. Consider a solution F of (8) such that $Z_m(t) > 0$ in $t \in (t_0, t_1)$. Then (10) holds for any $(t, \sigma) \in (t_0, t_1) \times \mathbb{R}_+$. In particular, if $Z_m(t_0) = 0$ it takes the form

$$F_m(t, \sigma) = \lambda_m \int_{t_0}^t G_m \left(\sigma + \int_T^t x_{F,m}(\tau) d\tau \right) dT. \quad (12)$$

Proof: If $Z_m(t_0) > 0$ then this is a special case of Proposition 2. It remains to consider the case $Z_m(t_0) = 0$. For $t \in (t_0, t_1)$, consider a $t_a \in (t_0, t)$. Applying Proposition 2 to $[t_a, t]$ gives

$$\begin{aligned} F_m(t, \sigma) &= F_m \left(t_a, \sigma + \int_{t_a}^t x_{F,m}(\tau) d\tau \right) \\ + \lambda_m \int_{t_a}^t G_m \left(\sigma + \int_T^t x_{F,m}(\tau) d\tau \right) dT. \end{aligned}$$

Now consider the limit as $t_a \downarrow t_0$. The left hand side remains constant. Since G_m is bounded, the second term tends to the right hand side of (12). The first term is bounded below by zero and above by $F_m(t_a, 0)$, which tends to $F_m(t_0, 0) = Z_m(t) = 0$ by continuity of F_m . \square

III. LYAPUNOV BOUNDS

In this section we present a Lyapunov function and bounds on its derivative, which will be used in the later sections to prove that network stability holds when the loads satisfy the capacity constraints (2). Before proceeding we recapitulate the partial differential equation model

$$\frac{\partial F_m(t, \sigma)}{\partial t} = \begin{cases} \frac{\partial F_m(t, \sigma)}{\partial \sigma} \frac{\varphi_m(t)}{Z_m(t)} + \lambda_m G_m(\sigma) & Z_m > 0, \\ 0 & Z_m = 0, \end{cases} \quad (13)$$

and establish some basic facts involving the *residual workload function* $W_m(t)$ for each route m . This measures the total residual workload at time t , in the fluid limit, and can be expressed as

$$W_m(t) = \int_0^\infty \sigma d\zeta_m(\sigma) = \int_0^\infty F_m(t, \sigma) d\sigma. \quad (14)$$

Here the second step follows by integration by parts. We will assume that the initial workload $W_m(0)$ is finite for all m . The following Lemma (analogous to Lemma 3.3 in [9]) describes the evolution of $W_m(t)$.

Lemma 4. Given a solution $F_m(t, \sigma)$ to (8), the workload function $W_m(t)$ defined in (14) satisfies

$$\dot{W}_m = \begin{cases} \rho_m - \varphi_m(t), & Z_m > 0, \\ 0 & Z_m = 0, \end{cases}$$

and therefore the bound

$$W_m(t) \leq W_m(0) + \rho_m t. \quad (15)$$

In particular, $W_m(t)$ remains finite for all time.

Proof: Focusing on the case $Z_m > 0$, integrating the PDE with respect to σ and using (7) yields

$$\begin{aligned} \dot{W}(t) &= \int_0^\infty \frac{\partial F_m(t, \sigma)}{\partial t} d\sigma \\ &= \frac{\varphi_m(t)}{Z_m(t)} \left[F_m(t, \sigma) \right]_0^\infty + \lambda_m \int_0^\infty G_m(\sigma) d\sigma \\ &= -\varphi_m(t) + \rho_m. \end{aligned}$$

\square

A. Lyapunov function

Choose a sufficiently small δ satisfying

Condition 1. For all l , $\tilde{\rho}_m = (1 + \delta)\rho_m$ satisfies $\sum_m R_{lm} \tilde{\rho}_m < c_l$, and $(1 - \delta)(1 + \delta)^{\alpha+1} > 1$ and $\delta > 0$.

Recall that $\alpha > 0$ is the fairness parameter used by the congestion control. The second inequality always holds for

$0 < \delta < \alpha/(2 + \alpha)$; note $\delta \rightarrow 0$ as $\alpha \rightarrow 0$. Introduce the Lyapunov function

$$\begin{aligned} L(t) &= \sum_m L_m(t) \\ &= \sum_m \frac{1}{\tilde{\rho}_m^\alpha} \int_0^\infty [F_m(t, \sigma)]^{\alpha+1} w_m(\sigma) d\sigma. \end{aligned} \quad (16)$$

Here $w_m(\sigma)$ is a ‘‘spatial weight’’ to be selected shortly; we impose that it is non-negative and bounded in $\sigma \geq 0$, and normalized to $w_m(0) = 1$.

As a first remark, note that since $F_m(t, \sigma)$ is by definition monotonically non-increasing in σ , we have

$$F_m(t, \sigma)^{\alpha+1} \leq Z_m(t)^\alpha F_m(t, \sigma),$$

therefore

$$\begin{aligned} L_m(t) &\leq \frac{\|w_m\|_\infty Z_m^\alpha(t)}{\tilde{\rho}_m^\alpha} \int_0^\infty F_m(t, \sigma) d\sigma \\ &= \kappa_m Z_m^\alpha W_m(t), \end{aligned} \quad (17)$$

with $\kappa_m := \|w_m\|_\infty \tilde{\rho}_m^{-\alpha}$. So L_m is finite for all time using Lemma 4, and the Lyapunov function is well-defined. We now compute the time derivative of

$$[\tilde{\rho}_m]^\alpha L_m = \int_0^\infty [F_m(t, \sigma)]^{\alpha+1} w_m(\sigma) d\sigma$$

along the trajectory, for any $m : Z_m > 0$. We have:

$$\begin{aligned} \tilde{\rho}_m^\alpha \dot{L}_m &= \int_0^\infty (\alpha + 1) [F_m(t, \sigma)]^\alpha \frac{\partial F_m(t, \sigma)}{\partial t} w_m(\sigma) d\sigma \\ &= \frac{\varphi_m(t)}{Z_m(t)} \int_0^\infty \frac{\partial F_m^{\alpha+1}(t, \sigma)}{\partial \sigma} w_m(\sigma) d\sigma \\ &+ \int_0^\infty (\alpha + 1) [F_m(t, \sigma)]^\alpha \lambda_m G_m(\sigma) w_m(\sigma) d\sigma. \end{aligned} \quad (18)$$

Integrating by parts in the first term, we have

$$\begin{aligned} &\int_0^\infty \frac{\partial F_m^{\alpha+1}(t, \sigma)}{\partial \sigma} w_m(\sigma) d\sigma \\ &= F_m^{\alpha+1}(t, \sigma) w_m(\sigma) \Big|_{\sigma=0}^\infty - \int_0^\infty F_m(t, \sigma)^{\alpha+1} w'_m(\sigma) d\sigma \\ &= -Z_m(t)^{\alpha+1} - \int_0^\infty F_m(t, \sigma)^{\alpha+1} w'_m(\sigma) d\sigma \end{aligned}$$

Substituting in (18) we obtain

$$\begin{aligned} \tilde{\rho}_m^\alpha \dot{L}_m &= -\varphi_m(t) Z_m(t)^\alpha \\ &+ \int_0^\infty F_m(t, \sigma)^\alpha \left\{ -\frac{\varphi_m(t)}{Z_m(t)} F_m(t, \sigma) w'_m(\sigma) \right. \\ &\left. + (\alpha + 1) \lambda_m G_m(\sigma) w_m(\sigma) \right\} d\sigma. \end{aligned} \quad (19)$$

B. Choice of weight $w_m(\sigma)$

We now specify that $w_m(\sigma)$ satisfies the following differential equation in σ ,

$$w'_m(\sigma) = K \mu_m G_m(\sigma) w_m(\sigma)^{\frac{\alpha+1}{\alpha}}$$

for some $K \in (0, \alpha)$ to be specified later. This equation can be readily solved (for $w_m(0) = 1$) to yield

$$w_m(\sigma) = \left(1 - \frac{K \mu_m}{\alpha} \int_0^\sigma G_m(u) du \right)^{-\alpha}. \quad (20)$$

Note that since $\mu_m \int_0^\infty G_m(u) du = 1$, for $K < \alpha$ the term in brackets is strictly positive, bounded away from zero, so $w_m(\sigma)$ is well-defined, non-negative and bounded. With this choice, (19) becomes

$$\begin{aligned} \tilde{\rho}_m^\alpha \dot{L}_m &= -\varphi_m(t) Z_m(t)^\alpha + \\ &\int_0^\infty F_m(t, \sigma)^\alpha \left\{ -\frac{\varphi_m(t)}{Z_m(t)} F_m(t, \sigma) w_m(\sigma)^{\frac{1}{\alpha}} K \right. \\ &\left. + (\alpha + 1) \rho_m \right\} \mu_m G_m(\sigma) w_m(\sigma) d\sigma. \end{aligned} \quad (21)$$

C. Bounding the Lyapunov derivative.

We wish to upper bound the terms involving $F_m(t, \sigma)$ in the above integral. For this we calculate the maximum of the function

$$\phi(F) := F^\alpha \{(\alpha + 1) \rho_m - bF\},$$

over $F \geq 0$. Here we have denoted

$$b = \frac{\varphi_m(t)}{Z_m(t)} w_m(\sigma)^{\frac{1}{\alpha}} K.$$

By differentiation we have

$$\phi'(F) = F^{\alpha-1} \{ \alpha(\alpha + 1) \rho_m - (\alpha + 1) bF \},$$

which has a root

$$F^* = \frac{\alpha \rho_m}{b} = \frac{\alpha \rho_m Z_m}{K \varphi_m w_m^{1/\alpha}},$$

and yields a maximum

$$\phi(F^*) = \left(\frac{\alpha \rho_m Z_m}{K \varphi_m w_m^{1/\alpha}} \right)^\alpha \rho_m = \frac{\alpha^\alpha \rho_m^{\alpha+1} Z_m^\alpha}{K^\alpha \varphi_m^\alpha w_m}.$$

Returning to (21), we obtain the bound

$$\begin{aligned} \tilde{\rho}_m^\alpha \dot{L}_m &\leq -\varphi_m(t) Z_m(t)^\alpha + \\ &\int_0^\infty \frac{\alpha^\alpha \rho_m^{\alpha+1} Z_m(t)^\alpha}{K^\alpha \varphi_m(t)^\alpha w_m(\sigma)} \mu_m G_m(\sigma) w_m(\sigma) d\sigma \\ &= -\varphi_m(t) Z_m(t)^\alpha \\ &+ \frac{\alpha^\alpha \rho_m^{\alpha+1} Z_m(t)^\alpha}{K^\alpha \varphi_m(t)^\alpha} \int_0^\infty \mu_m G_m(\sigma) d\sigma \\ &= -\varphi_m(t) Z_m(t)^\alpha + \frac{\alpha^\alpha \rho_m^{\alpha+1} Z_m(t)^\alpha}{K^\alpha \varphi_m(t)^\alpha}. \end{aligned} \quad (22)$$

Since K is a free parameter, restricted only by $K < \alpha$, we can now choose it to satisfy

$$\left(\frac{\alpha}{K} \right)^\alpha = (1 - \delta)(1 + \delta)^{\alpha+1} > 1. \quad (23)$$

Then (22) becomes

$$\dot{L}_m \leq Z_m(t)^\alpha \left\{ -\frac{\varphi_m(t)}{\tilde{\rho}_m^\alpha} + \frac{\tilde{\rho}_m(1 - \delta)}{\varphi_m(t)^\alpha} \right\}. \quad (24)$$

Lemma 5. For any positive numbers $\tilde{\rho}$, φ ,

$$-\frac{\varphi}{\tilde{\rho}^\alpha} + \frac{\tilde{\rho}}{\varphi^\alpha} \leq (\alpha + 1) \frac{(\tilde{\rho} - \varphi)}{\varphi^\alpha}. \quad (25)$$

Proof: Bounding the convex function $h(x) = x^{\alpha+1}$ by its tangent around the point $x = \tilde{\rho}$ gives

$$\varphi^{\alpha+1} \geq \tilde{\rho}^{\alpha+1} + (\alpha + 1)\tilde{\rho}^\alpha(\varphi - \tilde{\rho}).$$

Dividing by $\tilde{\rho}^\alpha\varphi^\alpha$ and reordering terms yields (25). \square

We now state the main result of this section, that shows the Lyapunov function has negative drift along trajectories, provided the natural stability condition (2) holds.

Theorem 6. Suppose that the capacity constraints (2) are satisfied, and consider the Lyapunov function L defined in (16) with $w_m(\sigma)$ in (20), K in (23) and δ satisfying Condition 1. Then L satisfies

$$\dot{L} \leq -\delta \sum_{m:Z_m>0} \tilde{\rho}_m \left(\frac{Z_m}{\varphi_m} \right)^\alpha. \quad (26)$$

Proof: We use the bound (25) in (24), and obtain

$$\dot{L}_m \leq (\alpha + 1) \left(\frac{Z_m}{\varphi_m} \right)^\alpha (\tilde{\rho}_m - \varphi_m) - \delta \tilde{\rho}_m \left(\frac{Z_m}{\varphi_m} \right)^\alpha$$

for any m where $Z_m > 0$. Note also that $\dot{L}_m = 0$ when $Z_m = 0$, (refer to Remark in Section II-E). Superimposing all terms we get

$$\begin{aligned} \dot{L} = \sum_m \dot{L}_m &\leq (\alpha + 1) \sum_{m:Z_m>0} \left(\frac{Z_m}{\varphi_m} \right)^\alpha (\tilde{\rho}_m - \varphi_m) \\ &\quad - \delta \sum_{m:Z_m>0} \tilde{\rho}_m \left(\frac{Z_m}{\varphi_m} \right)^\alpha. \end{aligned}$$

Noting that $\left(\frac{Z_m}{\varphi_m} \right)^\alpha = U'_m \left(\frac{\varphi_m}{Z_m} \right)$, we are in a position to apply (5), with $\psi_m = \tilde{\rho}_m$ that satisfy the capacity constraints. This is the only step that relies on the underlying congestion control resource allocation. \square

IV. STABILITY RESULTS

A. Asymptotic convergence

We show first an *asymptotic* stability result under very general conditions: requiring only that loads strictly satisfy the capacity constraints, the solutions to the fluid model (8) asymptotically converge to zero.

Theorem 7. Under the capacity constraints (2), the Lyapunov function L defined in (16) with $w_m(\sigma)$ in (20), K in (23) and δ satisfying Condition 1, satisfies

$$\lim_{t \rightarrow \infty} L(t) = 0.$$

Proof: First note that a bound of the form

$$L_m \leq (A + Bt)Z_m^\alpha \quad (27)$$

follows from from (17) and (15), by defining $A = \max_m \kappa_m W_m(0)$, $B = \max_m \kappa_m \rho_m$. This leads to

$$\tilde{\rho}_m \left(\frac{Z_m}{\varphi_m} \right)^\alpha \geq \frac{L_m \rho_0}{C^\alpha (A + Bt)},$$

where $\rho_0 = \min_m \rho_m$ and $C = \max_l c_l$. Take $\eta := \rho_0 C^{-\alpha}$, and apply Theorem 6 to obtain

$$\dot{L} \leq -\frac{\delta \eta}{(A + Bt)} \sum_m L_m = -\frac{\delta \eta}{(A + Bt)} L.$$

While $L(t) > 0$, we can integrate to obtain

$$\log L(t) \leq \log L(0) - \int_0^t \frac{\delta \eta}{(A + Bu)} du.$$

As $t \rightarrow \infty$ the integral diverges, so $L(t) \rightarrow 0$. \square

Remark 3. Since the definition of $L_m(t)$ involves a weighted norm of the state $F_m(t, \cdot)$ in the function space $\mathcal{L}^{1+\alpha}$, it follows that $F_m(t, \cdot)$ converges to zero in this space. Now, given that the functions $F_m(t, \sigma)$ are monotonically non-increasing in σ , there must also be pointwise convergence to zero for every $\sigma \in (0, \infty)$.

B. Finite time convergence

In order to relate fluid stability with the stochastic models, convergence of the fluid models to zero in *finite time* is often required [8], [13]. We show in this section that such a result follows from a very slight strengthening of the hypothesis: requiring that the file-size distributions ν_m have a finite p moment for some $p > 1$. Specifically,

$$B_m := \int_0^\infty \sigma^p d\nu_m = \int_0^\infty \sigma^p [-dG_m(\sigma)] < \infty. \quad (28)$$

Notice that this still allows for heavy-tailed distributions: for instance, the Pareto distribution with complementary cumulative distribution function

$$G_m(\sigma) = \min\left\{1, \frac{1}{\sigma^\gamma}\right\}$$

has finite mean if and only if $\gamma > 1$. In this case it will also have a finite p -moment for p close enough to 1.

We will impose an analogous restriction to the initial condition $F_m^{(0)}(\sigma) := F_m(0, \sigma)$. This is the complementary CDF of a finite measure $\zeta_m(0)$ (note that $F_m(0, 0)$ is assumed finite, not necessarily unity), and we can write the moment condition

$$A_m := \int_0^\infty \sigma^p d\zeta_m(0) = \int_0^\infty \sigma^p [-dF_m^{(0)}(\sigma)] < \infty. \quad (29)$$

Proposition 8. Suppose G_m and $F_m^{(0)}$ satisfy respectively (28) and (29), then $F_m^{(t)}(\sigma) := F_m(t, \sigma)$ satisfies the finite moment condition

$$\int_0^\infty \sigma^p [-dF_m^{(t)}(\sigma)] \leq A_m + B_m \lambda_m t. \quad (30)$$

Proof: For brevity we drop the subindex m from all variables. Since $\frac{\partial F}{\partial \sigma} \leq 0$, the PDE (13) gives

$$\frac{\partial F(t, \sigma)}{\partial t} \leq \lambda G(\sigma),$$

which can be integrated in time to give

$$F(t, \sigma) \leq F(0, \sigma) + \lambda G(\sigma)t.$$

Now invoke integration by parts to write

$$\begin{aligned}
& \int_0^a \sigma^p [-dF^{(t)}(\sigma)] \\
&= -a^p F^{(t)}(a) + \int_0^a p \sigma^{p-1} F^{(t)}(\sigma) d\sigma \\
&\leq \int_0^a p \sigma^{p-1} [F^{(0)}(\sigma) + \lambda G(\sigma)t] d\sigma \\
&= \int_0^a \sigma^p [-dF^{(0)}(\sigma)] + a^p F^{(0)}(a) + \\
&\quad \lambda t \left[\int_0^a \sigma^p [-dG(\sigma)] + a^p G(a) \right].
\end{aligned}$$

Taking limit with $a \rightarrow \infty$, the ‘‘tail’’ terms vanish due to the finite moment assumption on $F^{(0)}$ and G , and we obtain (30). \square

As a corollary we derive the following bound on the residual workload function. This is a generalization of the inequality in Lemma 4. Indeed, if we use $p = 1$ in (31) below, we obtain (15).

Proposition 9. *Under the conditions of Proposition 8,*

$$W_m(t) \leq (A_m + B_m \lambda_m t)^{\frac{1}{p}} Z_m(t)^{1 - \frac{1}{p}}. \quad (31)$$

Proof: Apply Hölder’s inequality to the integral

$$\begin{aligned}
W_m(t) &= \int_0^\infty \sigma [-dF_m^{(t)}(\sigma)] \\
&\leq \left[\int_0^\infty \sigma^p [-dF_m^{(t)}(\sigma)] \right]^{\frac{1}{p}} \left[\int_0^\infty [-dF_m^{(t)}(\sigma)] \right]^{1 - \frac{1}{p}}.
\end{aligned}$$

The desired bound then follows from (30). \square

Using the above bounds, we are now ready to prove our result about finite time convergence.

Theorem 10. *Suppose G_m and $F_m^{(0)}$ satisfy respectively (28) and (29). Then, under the capacity constraints (2), $F_m^{(t)}$ converges to zero in finite time.*

Proof: It suffices to show that $L(t)$ chosen as in Theorem 6 reaches zero in finite time. Applying (31) to (17) yields

$$L_m \leq (\kappa_m^p A_m + \kappa_m^p B_m \lambda_m t)^{\frac{1}{p}} Z_m^{\alpha + 1 - \frac{1}{p}},$$

which we simplify to

$$L_m \leq (A + Bt)^{1-\beta} Z_m^{\alpha+\beta}, \quad (32)$$

by setting $A = \max_m \kappa_m^p A_m$, $B = \max_m \kappa_m^p B_m \lambda_m$, and $\beta = 1 - 1/p$. Note that $0 < \beta < 1$.

Note the similarity between (32) and (27). Proceeding as in Theorem 7, when $Z_m \neq 0$, whence $\varphi_m \neq 0$, we write

$$\tilde{\rho}_m \left(\frac{Z_m}{\varphi_m} \right)^\alpha \geq \frac{\rho_0}{C^\alpha} \frac{L_m^{\frac{\alpha}{\alpha+\beta}}}{(A+Bt)^{\frac{\alpha(1-\beta)}{\alpha+\beta}}}.$$

Therefore

$$\begin{aligned}
\sum_{m: Z_m > 0} \tilde{\rho}_m \left(\frac{Z_m}{\varphi_m} \right)^\alpha &\geq \frac{\rho_0 C^{-\alpha}}{(A+Bt)^{\frac{\alpha(1-\beta)}{\alpha+\beta}}} \sum_{m: Z_m > 0} (L_m)^{\frac{\alpha}{\alpha+\beta}} \\
&\geq \frac{\rho_0 C^{-\alpha}}{(A+Bt)^{\frac{\alpha(1-\beta)}{\alpha+\beta}}} \max_m (L_m)^{\frac{\alpha}{\alpha+\beta}} \\
&\geq \frac{\rho_0 C^{-\alpha} \left(\frac{1}{M} \sum_m L_m \right)^{\frac{\alpha}{\alpha+\beta}}}{(A+Bt)^{\frac{\alpha(1-\beta)}{\alpha+\beta}}} \\
&= \frac{\eta}{(A+Bt)^{\frac{\alpha(1-\beta)}{\alpha+\beta}}} L^{\frac{\alpha}{\alpha+\beta}}.
\end{aligned}$$

Here M is the number of routes, and η is appropriately defined. Returning now to (26), we obtain

$$\dot{L} \leq - \frac{\delta \eta}{(A+Bt)^{\frac{\alpha(1-\beta)}{\alpha+\beta}}} L^{\frac{\alpha}{\alpha+\beta}}.$$

When $L > 0$, this yields

$$\begin{aligned}
\frac{d}{dt} L^{\frac{\beta}{\alpha+\beta}} &= \frac{\beta}{\alpha+\beta} L^{-\frac{\alpha}{\alpha+\beta}} \dot{L} \\
&\leq - \frac{\beta}{\alpha+\beta} \frac{\delta \eta}{(A+Bt)^{\frac{\alpha(1-\beta)}{\alpha+\beta}}} \\
&= -\epsilon \frac{d}{dt} (A+Bt)^{\frac{\beta(\alpha+1)}{\alpha+\beta}},
\end{aligned}$$

where we define $\epsilon = \frac{\delta \eta}{B(1+\alpha)} > 0$ and use the identity

$$1 - \frac{\alpha(1-\beta)}{\alpha+\beta} = \frac{\beta(\alpha+1)}{\alpha+\beta}.$$

Integrating we have

$$L(t)^{\frac{\beta}{\alpha+\beta}} \leq L(0)^{\frac{\beta}{\alpha+\beta}} - \epsilon (A+Bt)^{\frac{\beta(\alpha+1)}{\alpha+\beta}} + \epsilon A^{\frac{\beta(\alpha+1)}{\alpha+\beta}}. \quad (33)$$

Since the second term on the right of (33) grows without bound in t , $L(t)$ must reach zero in finite time. \square

Note that both here and in Theorem 7, the speed of convergence is controlled by δ , and this parameter goes to zero as $\alpha \rightarrow 0$. This is consistent with the fact that for $\alpha = 0$, the network need not be stable [3, Example 1].

Convergence time.

We will additionally show that the convergence time grows linearly as a function of the ‘‘size’’ of the initial condition $F_m(0, \sigma)$, expressed in terms of $Z_m(0) = F_m(0, 0)$ and the p -moment A_m defined in (29). We have the following result.

Proposition 11. *Under the assumptions of Theorem 10, there exists an constant κ with the following property: for every $r > 0$, if the initial condition satisfies*

$$F_m(0, 0) \leq r, A_m \leq r \text{ for each } m, \quad (34)$$

then $F_m(t, \sigma) = 0$ for every $t \geq \kappa r$ and every σ, m .

Proof: Since by definition $L(t) = 0$ implies $F_m(t, \sigma) \equiv 0$ for all σ, m , we focus on the time t_0 when $L(t)$ reaches zero. From (33), this time satisfies

$$(A+Bt_0)^{\frac{\beta(\alpha+1)}{\alpha+\beta}} = A^{\frac{\beta(\alpha+1)}{\alpha+\beta}} + \frac{1}{\epsilon} L(0)^{\frac{\beta}{\alpha+\beta}}. \quad (35)$$

It is sufficient to show that when (34) holds, the above t_0 can be bounded by κr for a fixed constant κ . We bound the terms on the right-hand side of (35).

The first term is bounded directly since $A = \max_m \kappa_m^p A_m \leq r \max_m \kappa_m^p$. So

$$A^{\frac{\beta(\alpha+1)}{\alpha+\beta}} \leq \bar{\kappa} r^{\frac{\beta(\alpha+1)}{\alpha+\beta}}$$

for an appropriate $\bar{\kappa}$. For the second term, apply (32) at $t = 0$ to get

$$\begin{aligned} L_m(0) &\leq A^{(1-\beta)} Z_m(0)^{\alpha+\beta} \\ &\leq \max_m \kappa_m r^{(1-\beta)r^{\alpha+\beta}} \\ &= r^{\alpha+1} \max_m \kappa_m. \end{aligned}$$

Therefore

$$L(0)^{\frac{\beta}{\alpha+\beta}} \leq \hat{\kappa} r^{\frac{\beta(\alpha+1)}{\alpha+\beta}}$$

for an appropriate $\hat{\kappa}$. Combining now both terms on the right of (35) we write

$$(Bt_0)^{\frac{\beta(\alpha+1)}{\alpha+\beta}} \leq (\bar{\kappa} + \hat{\kappa}/\epsilon) r^{\frac{\beta(\alpha+1)}{\alpha+\beta}},$$

from which the bound $t_0 \leq \kappa r$ follows. The resulting κ depends on a number of fixed constants, including the p -moment of the arriving file-size distributions, but not on the initial condition. \square

V. STOCHASTIC STABILITY

As explained in Section II, the PDE model under consideration is mainly motivated by taking a *fluid limit* on a stochastic queueing model. The question we now pose is whether, having established the stability of the fluid model, can we infer the corresponding result for the original stochastic model.

A first issue is what is meant by stochastic stability. In the exponential file size case, the stochastic process is a Markov chain where the number of connections $z = (z_m)$ per class is the state; in this case stability is usually defined to be positive recurrence of the Markov chain (see [2]). The natural generalization of the Markov model to general arrival times and file sizes (G/G) that form a renewal process is the one used by Dai [6]. There, a Markov process is defined where the state keeps track of residual arrival times and service times of currently active jobs, in addition to queue sizes. Stability is defined as the positive Harris recurrence of such Markov process. Dai [6] also obtains a fluid limit model, and defines a notion of stability for fluid models in terms of convergence to zero in finite time, similar to the one obtained in Theorem 10.

Does fluid stability imply stochastic stability? Dai [6] establishes this for service disciplines where the number of residual times in the state remains bounded. This does not cover processor sharing disciplines, where all jobs present in the system receive service, as is the case for our problem. Although [6] claims that extensions to this case “should be evident”, we share the view of Gromoll and Williams [8], [9] that such extensions are not straightforward.

The recent PhD thesis of Nam Lee [13] has addressed precisely this problem: using the fluid models of [8], [9]

mentioned in Section II, conditions are given under which fluid stability implies stochastic stability (Harris recurrence). We now highlight some aspects of this work, without getting into technical details which are far beyond our scope.

A number of technical assumptions are made in [13] on the processes of inter-arrival times and file-sizes. We focus here on a “light-tails” condition (Assumption 2.2.1 (iii) in [13]) that is imposed on the file size distribution. If X is a random file, the condition can be stated as

$$\limsup_{r \rightarrow \infty} \sup_{a \geq 0} E[(X - a) \mathbf{1}_{\{X - a > r\}} | X > a] = 0. \quad (36)$$

This condition is not too restrictive: in particular, phase-type distributions, which are known to be dense in the space of distributions [2], satisfy (36). Still, it is more restrictive than what we imposed in the previous section; in particular Pareto distributions do not satisfy it. Under the condition (36), [13] proves that the stochastic model is stable provided fluid stability holds in the following sense:

Definition 1 (Def 2.6.3, [13]). *The fluid model is stable if for each r , there exists t_r such that any fluid model solutions with initial condition $F = (F_m^{(0)})$*

$$\mathcal{B}_r := \{F : F_m(0, 0) \leq r, \quad W_m(0) \leq r \quad \forall m\} \quad (37)$$

satisfies $F_m(t, \sigma) \equiv 0$ for $t \geq t_r$.

Note the strong similarity of this statement and the result we obtained in Proposition 11; here as well we are requiring convergence in a finite time that depends on the size of the initial condition. The only slight difference is that the set described by (34) (let us call it \mathcal{B}_r^p) involves the p -moment, $p > 1$, whereas \mathcal{B}_r is in terms of the 1-moment. It is not difficult to see that $\mathcal{B}_r^p \subset \mathcal{B}_r$, so Proposition 11 falls short of establishing fluid stability in the sense of Definition 1.

We are not certain as to whether this issue is significant, or if instead the theory in [13] can be extended to work with the initial set of (34). Nevertheless, we will show that under the light-tailed condition (36), the requirements of Definition 1 can indeed be satisfied, leading to a complete answer in this case.

For simplicity, we will replace (36) with

$$\sup_{a \geq 0} E[(X - a) | X > a] = \Theta < \infty. \quad (38)$$

It is not difficult to see that (36) implies (38). Therefore, we can use this fact when invoking the theory of [13]. We can also express (38) in terms of the complementary cumulative distribution function (CCDF), as follows.

Condition 2. *A finite measure ζ on \mathbb{R}_+ has light tails if there exists $\Theta < \infty$ such that the complementary cumulative distribution $F(\sigma) = \zeta((\sigma, \infty))$ satisfies*

$$\int_a^\infty F(\sigma) d\sigma \leq \Theta F(a) \quad \forall a \geq 0. \quad (39)$$

We will apply Condition 2 to the file-size distribution ν_m (with CCDF $G_m(\sigma)$). To see the equivalence with (38), note that the CCDF of the variable $X - a$, conditioned on $X > a$, is $G_m(a + \sigma)/G_m(a)$. We also impose the bound (39) to the

initial condition $\zeta_m(0)$ (with CCDF $F_m^{(0)}$); in this case the measure is not normalized, but (39) is still meaningful.

The key property of Condition 2 is that it *propagates* through time when one follows the PDE (8).

Proposition 12. *Consider a solution of (8) in a time interval (t_0, t_1) where $Z_m(t) > 0$. Assume $F^{(t_0)}$ and G_m satisfy Condition 2 with a common Θ . Then $F^{(t_1)}$ satisfies Condition 2 with the same Θ .*

Proof: If $Z_m(t_1) = 0$, then $F^{(t_1)} \equiv 0$ and the proposition holds. It remains to consider the case $Z_m(t_1) > 0$. If $Z_m(t_0) > 0$, we can invoke the integral form in Proposition 2:

$$F_m^{(t_1)}(\sigma) = F_m^{(t_0)} \left(\sigma + \int_{t_0}^{t_1} x_m(\tau) d\tau \right) + \lambda_m \int_{t_0}^{t_1} G_m \left(\sigma + \int_T^{t_1} x_m(\tau) d\tau \right) dT.$$

Integrating over $\sigma \in [a, \infty)$ gives

$$\begin{aligned} \int_a^\infty F_m^{(t_1)}(\sigma) d\sigma &= \int_a^\infty F_m^{(t_0)} \left(\sigma + \int_{t_0}^{t_1} x_m(\tau) d\tau \right) d\sigma \\ &+ \lambda_m \int_{t_0}^{t_1} \int_a^\infty G_m \left(\sigma + \int_T^{t_1} x_m(\tau) d\tau \right) d\sigma dT \\ &\leq \Theta F_m^{(t_0)} \left(a + \int_{t_0}^{t_1} x_m(\tau) d\tau \right) \\ &+ \Theta \lambda_m \int_{t_0}^{t_1} G_m \left(a + \int_T^{t_1} x_m(\tau) d\tau \right) dT \\ &= \Theta F_m^{(t_1)}(a). \end{aligned}$$

The change in order of integration is admissible since functions are non-negative. If instead $Z_m(t_0) = 0$, then we start with the form in Corollary 3, and perform the same manipulations with the first term omitted. \square

Note also that (39) is trivially propagated during an interval of time where $F_m(t, \sigma)$ remains at zero. Therefore, if we assume the initial condition and the arriving file-sizes satisfy Condition 2, the solution $F_m^{(t)}$ of the PDE will satisfy it for all time.

An importance consequence is the following bound on the workload, which follows from applying (39) to $F_m^{(t)}$ at $a = 0$:

$$W_m(t) \leq \Theta Z_m(t). \quad (40)$$

This bound is a strengthening of (31), which holds under these narrower (light-tailed) conditions. Indeed, it corresponds to setting $p = \infty$ in (31), or equivalently $\beta = 1$ in the proof of Theorem 10. Referring back to this proof, we have:

- (32) can be now replaced by

$$L_m \leq A Z_m^{\alpha+1},$$

by setting $A = \max_m \kappa_m \Theta$.

- Analogous steps lead to the condition

$$\dot{L} \leq -\frac{\delta\eta}{A^{\frac{1}{\alpha+1}}} L^{\frac{\alpha}{\alpha+1}}.$$

- For $L > 0$, this yields

$$\frac{d}{dt} L^{\frac{1}{\alpha+1}} \leq -\frac{\delta\eta}{(\alpha+1)A^{\frac{1}{\alpha+1}}}.$$

The right hand-side is now a constant, which we denote $-\epsilon$, and leads to

$$L(t)^{\frac{1}{\alpha+1}} \leq L(0)^{\frac{1}{\alpha+1}} - \epsilon t,$$

and thus again finite-time convergence.

We can now state the final result:

Theorem 13. *Assume that the initial conditions and file size-distributions satisfy Condition 2. Then, under the capacity constraints (2), the fluid model is stable in the sense of Definition 1.*

Proof: We start with an initial condition in \mathcal{B}_r as in (37). We bound the convergence time from the previous derivation by

$$t \leq \frac{1}{\epsilon} L(0)^{\frac{1}{\alpha+1}} = \left(\sum_m L_m(0) \right)^{\frac{1}{\alpha+1}} \leq A^{\frac{1}{\alpha+1}} M^{\frac{1}{\alpha+1}} r.$$

Here again M is the number of routes. If we denote by t_r the right-hand expression, we see that $L(t) = 0$ (and thus $F(t, \sigma) \equiv 0$ for $t \geq t_r$). \square

Putting all pieces together we are able to state a result on the stochastic system, invoking the theory in [13].

Corollary 14. *Consider a stochastic model that satisfies the technical Assumptions 2.2.1 and 2.2.2 from [13], in particular the light-tailed condition (36). Suppose that the loads ρ_m satisfy*

$$\sum_l R_{lm} \rho_m < c_l \text{ for each } l,$$

and jobs are served with an α -fair bandwidth allocation, $0 < \alpha < \infty$. Then the stochastic system is stable (Harris recurrent).

VI. PACKET-LEVEL VALIDATION

Over the course of this theoretical paper we have worked with highly idealized models of network protocols. In this section we offer evidence that the results remain relevant to real networks, by matching model predictions with packet-level simulations using the network simulator ns-2 [19].

A. Fluid model validation

Our models have involved *two* fluid limits:

- A lower-layer fluid model that replaces the packet counts of the TCP algorithm by a continuous rate satisfying the α -fair resource allocation of (1).
- A connection-level fluid model that replaces a stochastic queueing system of jobs served by TCP, by the PDE (8) where a continuous state $F(t, \sigma)$ quantifies jobs with residual work larger than σ .

The former has been the subject of many empirical studies (see for example [23]), which have shown that models representing flows as fluids are good approximations for systems in which each flow has a large bandwidth-delay product (BDP).

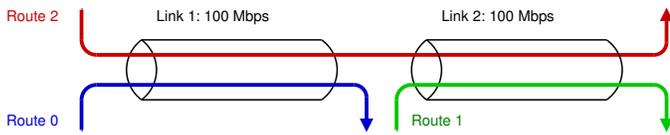


Fig. 1. Simulated “parking-lot” topology.

In particular, (weighted) α -fairness is a useful model for TCP Reno with $\alpha = 2$ and for TCP-Vegas and TCP-FAST with $\alpha = 1$ (see [22], [25]), in steady state. One remaining issue here is the validity of neglecting the congestion control transient in connection-level studies.

The latter fluid limit is supported by the theory in [11], [9], which shows convergence to the fluid model when time, and initial conditions, are re-scaled as described in Section II. This is sufficient to answer the question of stability, which is defined in terms of boundedness of the number of flows as time goes to infinity. However, a natural question is its predictive power beyond this asymptotic.

Another issue that may compromise applicability is that for large numbers of flows and fixed capacity, as stipulated in the second fluid limit, the BDP of each flow will be low, calling into question the accuracy of the first. The following simulation shows that applying both fluid models simultaneously can retain predictive power on the packet network dynamics.

We use the network of two links and three routes of Figure 1. Links have 100 Mbps capacity. Flows use TCP Reno (which approximates α -fairness with $\alpha = 2$), and each link has a round trip delay of 20 ms. Over each route we generate a Poisson process of TCP connections, and transmitted files with sizes independently drawn from a Pareto distribution with mean 5 Mbit and shape parameter 1.5. The mean load on each path was 45 Mbps, giving a load on each link of 90%. Since the fluid model proposed here describes the return from excursions in the number of flows, each path initially had 100 flows.

Figure 2 plots the number of ongoing flows for the ns2 simulation, in comparison with a numerical integration of the PDE. As expected, the number of flows in the packet-level simulation varied randomly around the value predicted by the fluid model, in particular the model captures the main characteristics such as the time to return to the steady state.

B. Stability under different TCP resource allocation models

Figure 2 demonstrates the stability of the network under less than 100% load, for the fairness model of TCP-Reno that corresponds to $\alpha = 2$. We now present simulations that illustrate the sort of instability that occurs when the load limit is exceeded, for two different fairness models.

Figure 3 compares the situations of 90% and 110% load, for zero initial conditions, still for the $\alpha = 2$ case. We see the instability of the latter case, in which the number of pending jobs starts to accumulate in the network.

In Figure 4 we replace the underlying transport protocol with TCP-FAST, which can be modelled by a proportionally fair allocation ($\alpha = 1$). We use the same loads as in the

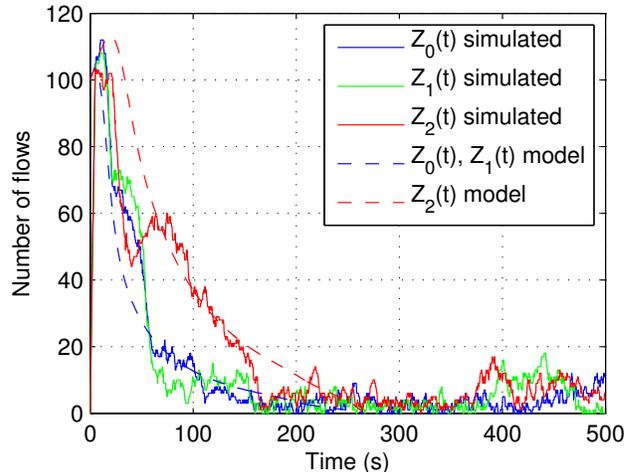
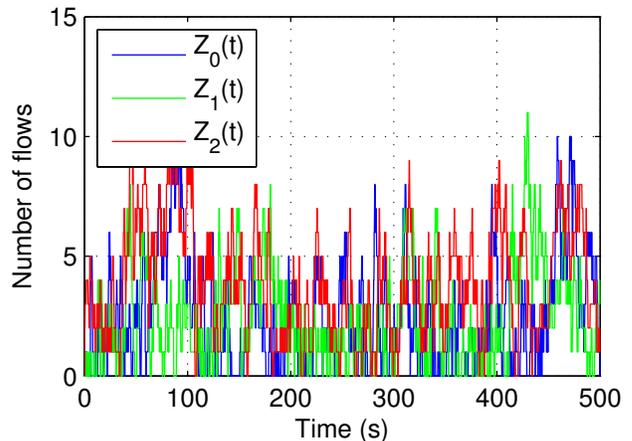
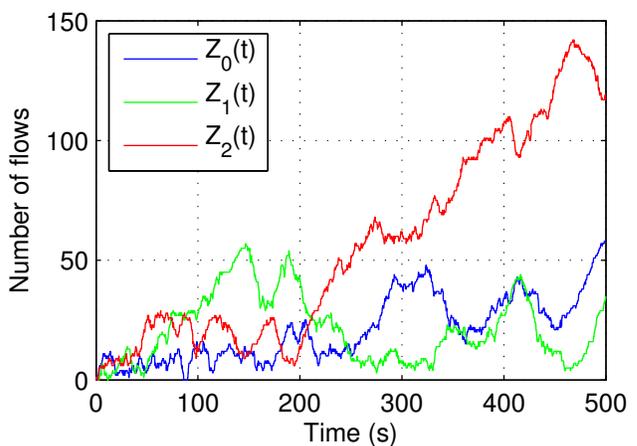


Fig. 2. Number of ongoing connections on each path of a two-hop network, in ns2 simulation (solid lines) and according to the fluid model (dotted lines). In the fluid model, routes 0 and 1 have identical numbers of flows. TCP Reno ($\alpha = 2$), with 90% load of Pareto-1.5 flows.



(a) $\rho = 0.9$



(b) $\rho = 1.1$

Fig. 3. Simulation example of a parking lot network with TCP/Reno ($\alpha = 2$), with Pareto-1.5 flow sizes.

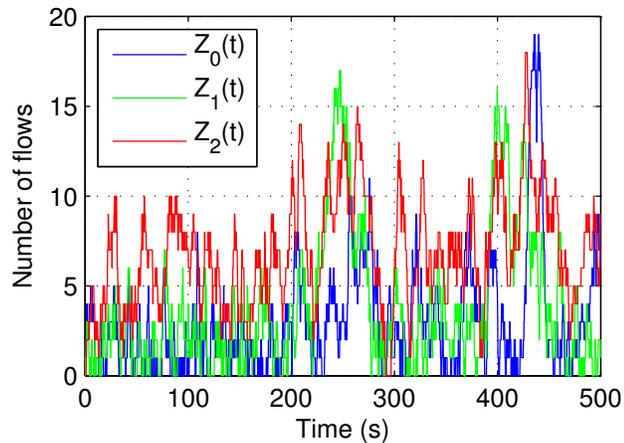
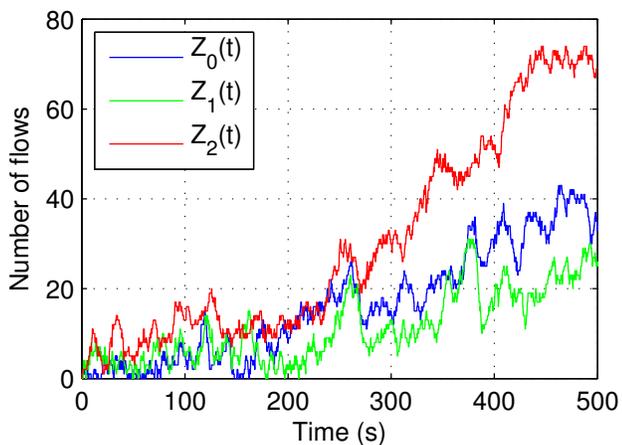
(a) $\rho = 0.9$ (b) $\rho = 1.1$

Fig. 4. Simulation example of a parking lot network with TCP/Fast ($\alpha = 1$), with Pareto-1.5 flow sizes.

previous example, and obtain similar results in regard to stability.

C. Example of instability when $\alpha = 0$

As mentioned in [3], an α -fair network with $\alpha = 0$ (i.e., the maximum throughput resource allocation) can be unstable even if the average load on each link is less than the capacity. In a symmetric parking-lot topology such as Figure 1, 0-fairness must give strict priority to the single-hop flows when they are present in both links. When only one link has single-hop flows, the 0-fair solution is indeterminate, but one possible choice is to still give priority to the single-hop flow present. In that case the set of stable arrival rates is strictly smaller than the capacity region of the network. This is illustrated in Figure 5, which shows the number of flows for this topology under 80% load, but with scheduling that gives priority to single-hop packets. In this case, the number of single-hop flows remains small, but the number of two-hop flows grows without bound.

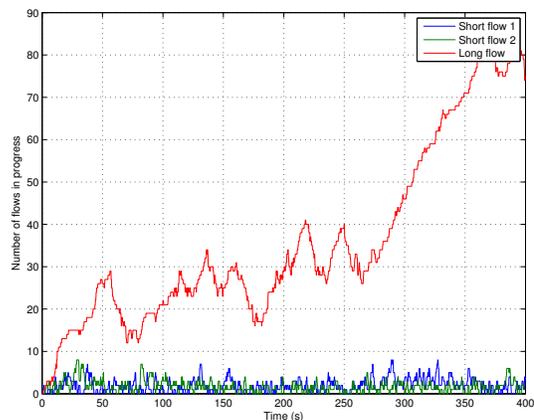


Fig. 5. Instability with each link underloaded (load 80%), with $\alpha = 0$. The number of flows on each path is plotted against time.

VII. CONCLUSION

We have considered the conjecture that the natural condition (all mean link loads strictly below capacity) suffices for the stability of a network with randomly arriving files of *general* size distributions, when jobs are served with α -fairness. Building on recent fluid limit studies [9], we formulated a partial differential equation model for the problem, where the state $F_m(t, \sigma)$ represents the residual workload distributions per route. We constructed a Lyapunov function a suitably weighted spatial $\alpha + 1$ -norm of F_m , that is shown to have negative drift under the natural stability condition. With these tools, we provided three main stability results:

- For general file-size distributions, only assumed to have finite mean, we prove the asymptotic convergence to zero of the fluid model.
- For file-size distributions of finite p -moment, $p > 1$, we prove convergence to zero of the fluid model in finite time, and bound this time as a function of the initial state. Note that the above restriction is very mild, in particular it includes “heavy-tailed” distributions such as Pareto, commonly invoked for Internet traffic models.
- For file-size distributions satisfying a “light-tailed” condition introduced in [13], we are able to sharpen our finite time convergence bound. In this way, we can directly invoke the theory in [13] to claim convergence of the associated stochastic model for this problem. The light-tailed condition is somewhat restrictive, but it still includes the class of phase-type distributions, which can approximate any other distribution.

From the above results, we regard the conjecture as fundamentally answered in an affirmative way. In terms of fluid models, the answer is always affirmative, and with a very mild restriction stability can be strengthened to finite time convergence. For stochastic models, the conjecture now has an affirmative answer in all cases where there is a proof mapping fluid stability to stochastic stability; this includes phase-type distributions. The only potential room for improvement in this latter case would be to map fluid to stochastic stability for a larger class. In this regard, a candidate class would be

distributions with finite $p > 1$ finite moment.

ACKNOWLEDGEMENTS

This work was supported by AFOSR-US under grant FA9550-09-1-0504, by NSF under CCF-0835706, by ANII-Uruguay under grant FCE 2007_265, and by the Australian Research Council under grants DP0985322 and FT0991594.

REFERENCES

- [1] M. Allman, V. Paxson and W. Stevens. *TCP Congestion Control*. RFC 2581, 1999.
- [2] S. Asmussen, *Applied Probability and Queues (2nd. ed)*, Springer, 2003.
- [3] T. Bonald and L. Massoulié. Impact of fairness on Internet performance. In *Proc. ACM Sigmetrics*, 2001.
- [4] M. Chiang, D. Shah and A. Tang. Stochastic stability under fair bandwidth allocation: General file size distribution. In *Proc. Allerton Conference*, 2006.
- [5] M. Crovella, A. Bestavros. Self-similarity in World Wide Web Traffic: Evidence and possible causes *IEEE/ACM Trans. Networking*, 5(6):835–846, December 1997.
- [6] J.G. Dai, “On positive Harris recurrence of multiclass queueing networks: A unified approach via fluid limit models”, *Ann. Appl. Probab.* 5:49–77, 1995.
- [7] G. de Veciana, T.-J. Lee, and T. Konstantopoulos. Stability and performance analysis of networks supporting services with rate control — Could the Internet be unstable? In *Proc. IEEE INFOCOM*, volume 2, pp. 802–810, 1999.
- [8] H. Gromoll and R. Williams. Fluid model for a data network with α -fair bandwidth sharing and general document size distributions: two examples of stability. In *IMS Collections — Markov Processes and Related Topics: A Festschrift for Thomas G. Kurtz*, vol.4, pp. 253–265, 2008.
- [9] H. Gromoll and R. Williams. Fluid limits for networks with bandwidth sharing and general document size distributions. *Ann. Appl. Probab.*, 19:243–280, 2009
- [10] F. Kelly, A. Maulloo and D. Tan Rate control in communication networks: shadow prices, proportional fairness and stability, *J. Op. Res. Soc.*, 39:237–252, 1998.
- [11] F. P. Kelly and R. J. Williams Fluid model for a network operating under a fair bandwidth-sharing policy. *Ann. Appl. Probab.* 14:1055–1083, 2004.
- [12] A. Lakshminantha, C. Beck and R. Srikant. On the use of SoS methods for analysis of connection-level stability in the Internet. In *Proc. American Control Conference*, Portland, OR, 2005.
- [13] Nam H. Lee. A sufficient condition for stochastic stability of an Internet congestion control model in terms of fluid model stability. PhD thesis, University of California at San Diego, 2008.
- [14] X. Lin, N. Shroff and R. Srikant. On the connection-level stability of congestion-controlled communication networks. *IEEE Trans. Information Theory*, 54(5):2317–2338, May 2008
- [15] J. Liu, A. Proutière, Y. Yi, M. Chiang, H.V. Poor. Flow-level stability of data networks with non-convex and time-varying rate regions. In *Proc. ACM Sigmetrics*, 2007, pp. 239-250.
- [16] L. Massoulié. Structural properties of proportional fairness: stability and insensitivity. *Ann. Appl. Probab.*, 17(3):809–839, 2007.
- [17] S. P. Meyn and R. L. Tweedie. Stability of Markovian Processes III: Foster-Lyapunov criteria for continuous time processes. *Adv. in Appl. Probab.* 25:518–548, 1993.
- [18] J. Mo and J. Walrand Fair end-to-end window based congestion control. *IEEE/ACM Trans. Networking*, 8(5):556–567, Oct. 2000.
- [19] “The Network Simulator NS-2,” <http://www.isi.edu/nsnam/ns/>.
- [20] F. Paganini, A. Tang, A. Ferragut, L. Andrew. Stability of Networks under General File Size Distribution with Alpha Fair Rate Allocation *Proc. Allerton Conference*, Monticello, IL, Oct 2009.
- [21] J. Roberts, L. Massoulié. Bandwidth Sharing and Admission Control for Elastic Traffic, In *Proc. ITC Specialist Seminar*, Yokohama, October 1998.
- [22] R. Srikant, *The Mathematics of Internet Congestion Control*, Birkhauser, 2004.
- [23] A. Tang, J. Wang, S. Hegde and S. H. Low. Equilibrium and Fairness of Networks Shared by TCP Reno and Vegas/FAST, *Telecommunication Systems* 30(4):417439, 2005.

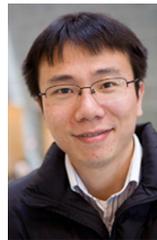
- [24] M. Verloop, S. Borst, R. Núñez-Queija. Stability of size-based scheduling disciplines in resource-sharing networks *Perf. Eval.* 62:247–262, 2005.
- [25] D. X. Wei, C. Jin and Steven H. Low. FAST TCP: Motivation, architecture, algorithms, performance, *IEEE/ACM Trans. Networking*, 14(6):1246–1259, Dec. 2006.
- [26] H.-Q. Ye. Stability of data networks under an optimization-based bandwidth allocation, *IEEE Trans. Automatic Control* 48(7):1238–1242, Jul. 2003.
- [27] H. Ye, J. Ou, and X. Yuan. Stability of data networks: Stationary and bursty models. *Operations Research*, 53:107–125, 2005.



Fernando Paganini Fernando Paganini (M’90–SM’05) received his Ingeniero Electricista and Licenciado en Matemática degrees from Universidad de la Republica, Montevideo, Uruguay, in 1990, and his M.S. and PhD degrees in Electrical Engineering from the California Institute of Technology, Pasadena, in 1992 and 1996 respectively. His PhD thesis received the 1996 Wilts Prize and the 1996 Clauser Prize at Caltech.

From 1996 to 1997 he was a postdoctoral associate at MIT. Between 1997 and 2005 he was on the faculty the Electrical Engineering Department at UCLA, reaching the rank of Associate Professor. Since 2005 he is Professor of Electrical and Telecommunications Engineering at Universidad ORT, Uruguay.

Dr. Paganini has received the 1995 O. Hugo Schuck Best Paper Award, the 1999 NSF CAREER Award, the 1999 Packard Fellowship, the 2004 George S. Axelby Best Paper Award, and the 2010 Elsevier Scopus Prize. He is a member of the Uruguayan National Academy of Sciences. His research interests are control and networks.



Ao Tang Ao Tang (S’01–M’07) received the B.E. in electronics engineering from Tsinghua University, Beijing, China, and the Ph.D. in electrical engineering with a minor in applied and computational mathematics from the California Institute of Technology (Caltech), Pasadena, CA, in 1999 and 2006, respectively.

He is currently an Assistant Professor in the School of Electrical and Computer Engineering at the Cornell University, where his current research interests include communication networks, sparse signal recovery and distributed dynamical systems.

Dr. Tang was the recipient of the 2006 George B. Dantzig Best Dissertation Award, the 2007 Charles Wilts best dissertation Prize, a 2009 IBM Faculty Award, and a 2010 DTRA Young Investigator Award.



Andrés Ferragut Andrés Ferragut obtained his degree in Electrical Engineering (Telecommunications) at Universidad de la Republica, Uruguay (2004). He has held teaching and research positions at the Mathematics and Electrical Engineering Depts. at Univ. de la Republica (2000 to 2009) and since 2007 at the School of Engineering, Universidad ORT, Uruguay where he is currently affiliated with the Mathematics Applied to Telecommunications (MATE) research group. He is a PhD Candidate in Electrical Engineering, having spent two years

(2004–2006) of graduate work at Ecole Nationale Supérieure de Télécommunications, Paris, France. His research interests are in mathematical modeling applied to telecommunication networks.



Lachlan L.H. Andrew Lachlan Andrew (M'97-SM'05) received the B.Sc., B.E. and Ph.D. degrees in 1992, 1993, and 1997, from the University of Melbourne, Australia. Since 2008, he has been an associate professor at Swinburne University of Technology, Australia, and since 2010 he has been an ARC Future Fellow. From 2005 to 2008, he was a senior research engineer in the Department of Computer Science at Caltech. Prior to that, he was a senior research fellow at the University of Melbourne and a lecturer at RMIT, Australia. His research interests include energy-efficient networking and performance analysis of resource allocation algorithms. He was co-recipient of the best paper award at IEEE INFOCOM 2011 and IEEE MASS07. He is a member of the ACM.