

Fluid model for a data network with α -fair bandwidth sharing and general document size distributions: two examples of stability

H. C. Gromoll^{1,*} and R. J. Williams^{2,†}

University of Virginia and University of California, San Diego

Abstract: The design and analysis of congestion control mechanisms for modern data networks such as the Internet is a challenging problem. Mathematical models at various levels have been introduced in an effort to provide insight to some aspects of this problem. A model introduced and studied by Roberts and Massoulié [13] aims to capture the dynamics of document arrivals and departures in a network where bandwidth is shared fairly amongst flows that correspond to continuous transfers of individual elastic documents. Here we consider this model under a family of bandwidth sharing policies introduced by Mo and Walrand [14]. With generally distributed interarrival times and document sizes, except for a few special cases, it is an open problem to establish stability of this stochastic flow level model under the nominal condition that the average load on each resource is less than its capacity. As a step towards the study of this model, in a separate work [8], we introduced a measure valued process to describe the dynamic evolution of the residual document sizes and proved a fluid limit result: under mild assumptions, rescaled measure valued processes corresponding to a sequence of flow level models (with fixed network structure) are tight, and any weak limit point of the sequence is almost surely a solution of a certain fluid model. The invariant states for the fluid model were also characterized in [8]. In this paper, we review the structure of the stochastic flow level model, describe our fluid model approximation and then give two interesting examples of network topologies for which stability of the fluid model can be established under a nominal condition. The two types of networks are linear networks and tree networks.

¹Department of Mathematics, University of Virginia, Charlottesville, VA 22903; e-mail: gromoll@virginia.edu

²Department of Mathematics, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0112; e-mail: williams@math.ucsd.edu

*Research supported in part by an NSF Mathematical Sciences Postdoctoral Research Fellowship, NSF Grant DMS FRG 0244323, a European Union Marie Curie Postdoctoral Research Fellowship, and EURANDOM.

[†]Research supported in part by NSF grants DMS 0305272 and DMS 0604537.

AMS 2000 subject classifications: Primary 60K30; secondary 60F17, 90B15.

Keywords and phrases: Bandwidth sharing, α -fair, flow level Internet model, connection level model, congestion control, measure valued process, fluid model, workload, Lyapunov function, simultaneous resource possession, Lagrange multipliers, stability, sensitivity.

1. Introduction

The design and analysis of congestion control mechanisms for modern data networks such as the Internet is a challenging problem. Mathematical models at various levels have been introduced in an effort to provide insight to some aspects of this problem. Roberts and Massoulié [13] have introduced and studied a flow level model of congestion control that represents the randomly varying number of flows present in a data network where bandwidth is shared fairly between flows that correspond to continuous transfers of individual elastic documents. This model assumes a “separation of time scales” such that the time scale of the flow dynamics (i.e., of document arrivals and departures) is much longer than the time scale of the packet level dynamics on which rate control schemes such as TCP converge to equilibrium.

Subsequent to the work of Roberts and Massoulié, assuming Poisson arrivals and exponentially distributed document sizes, de Veciana, Lee and Konstantopoulos [7] and Bonald and Massoulié [1] studied the stability of the flow level model operating under various bandwidth sharing policies. Lyapunov functions constructed in [7] for weighted max-min fair and proportionally fair policies, and in [1] for weighted α -fair policies ($\alpha \in (0, \infty)$) [14], imply positive recurrence of the Markov chain associated with the model when the average load on each resource is less than its capacity. Lin, Shroff and Srikant [11] have recently given sufficient conditions for stability of a Markov model where the assumption of time scale separation is relaxed.

Here we consider the model of Roberts and Massoulié with generally distributed document sizes and interarrival times. We are interested in the stability and heavy traffic behavior of this flow level model operating under a weighted α -fair bandwidth sharing policy ($\alpha \in (0, \infty)$) [14]. (Despite the claim in [1], the proof of sufficient conditions for stability given there does not apply when document sizes are other than exponentially distributed. The reason for this is that the method of Dai [5] quoted there implicitly assumes (through the form of the model equations) that the service discipline is a head-of-the-line discipline and consequently the method does not apply in general to processor sharing types of disciplines, such as the bandwidth sharing policy considered here.)

There are a few results on sufficient conditions for stability of the flow level model with general document size distributions. With Poisson arrivals and document sizes having a phase type distribution, for $\alpha = 1$, Lakshminantha et al. [10] have established stability of some two resource linear networks and a 2×2 grid network when the average load on each resource is less than its capacity. For generally distributed interarrival and document sizes, Bramson [3] has shown sufficiency of such a condition for stability under a max-min fair policy (corresponding to $\alpha \rightarrow \infty$). Under proportional fair sharing, Massoulié [12] has recently established stability of a fluid model for the flow level model with exponential interarrival and document sizes, and additional routing. From this he infers stability when documents have phase type distributions. In contemporaneous work described in a very recent preprint, Chiang, Shah and Tang [4] have developed a fluid approximation for the flow level model when the arrival rate and capacity are allowed to grow proportionally but the bandwidth per flow stays uniformly bounded. Using their fluid model, they derive some conclusions concerning stability for general document size distributions when $\alpha \in (0, \infty)$ is sufficiently small. However, in general, it remains an open question whether, with renewal arrivals and arbitrarily (rather than exponentially) distributed document sizes, the flow level model is stable under an α -fair bandwidth sharing policy when the nominal load placed on each resource is less than its capacity.

This paper reports on some first steps in our study of the flow level model operating under a weighted α -fair bandwidth sharing policy with general interarrival and document size distributions. Here we review the definition of a measure valued process that keeps track of the residual sizes of all documents in the system at any given time. We describe a fluid model (or formal functional law of large numbers approximation) for the flow level model. In a separate work [8], we showed that under mild conditions, appropriately rescaled measure valued processes corresponding to a sequence of flow level models (with fixed network structure) are tight, and any weak limit point of the sequence is almost surely a fluid model solution. The invariant states for the fluid model were also characterized in [8]. Here, as an illustration of sufficient conditions for stability of the fluid model, we establish stability of fluid model solutions with finite initial workload for linear networks and tree networks, under the nominal condition that the average load placed on each resource is less than its capacity. The result for tree networks is particularly interesting as there the distribution of the number of documents process in steady state is expected to be sensitive to the (non-exponential) document size distribution [2]. Future work will be aimed at further analysis of the fluid model and at using it for studying stability and heavy traffic behavior of the flow level model.

The paper is organized as follows. In Section 2, we define the network structure, the weighted α -fair bandwidth sharing policy, the stochastic model, and we introduce the measure valued processes used to describe the evolution of the system. The notion of a fluid model solution is defined in Section 3. In Section 4 we give sufficient conditions for stability of fluid model solutions with finite initial workload for linear networks and tree networks.

1.1. Notation

Let $\mathbb{N} = \{1, 2, \dots\}$, $\mathbb{R} = (-\infty, \infty)$, and let \mathbb{R}^d denote d -dimensional Euclidean space for any $d \geq 1$. For $x, y \in \mathbb{R}$, $x \wedge y$ is the minimum of x and y , and x^+ is the positive part of x . For $x, y \in \mathbb{R}^d$, let $\|x\| = \max_{i=1}^d |x_i|$, and interpret vector inequalities componentwise: $x \leq y$ means $x_i \leq y_i$ for all $i = 1, \dots, d$. The positive d -dimensional orthant is denoted $\mathbb{R}_+^d = \{x \in \mathbb{R}^d : x \geq 0\}$. To ease notation throughout the paper, define $c/0$ to be zero for any real constant c , and define a sum over an empty set of indices or of the form $\sum_{k=j}^l$ with $j > l$ to be zero.

For two functions f and g with the same domain, $f \equiv g$ means $f(x) = g(x)$ for all x in the domain. For a bounded function $f : \mathbb{R}_+ \rightarrow \mathbb{R}$, let $\|f\|_\infty = \sup_{x \in \mathbb{R}_+} |f(x)|$. Let $\mathbf{C}_b(\mathbb{R}_+)$ be the set of bounded continuous functions $f : \mathbb{R}_+ \rightarrow \mathbb{R}$, let $\mathbf{C}^1(\mathbb{R}_+)$ be the set of once continuously differentiable functions $f : \mathbb{R}_+ \rightarrow \mathbb{R}$, and let $\mathbf{C}_b^1(\mathbb{R}_+)$ be the set of functions f in $\mathbf{C}^1(\mathbb{R}_+)$ that together with the first derivative f' are bounded on \mathbb{R}_+ . If $w \in \mathbf{C}^1(\mathbb{R}_+)$ is considered as a function of time, its first derivative will be denoted by \dot{w} . For a Polish space (i.e., a complete separable metrizable space) \mathcal{S} , let $\mathbf{D}([0, \infty), \mathcal{S})$ denote the space of right continuous functions from $[0, \infty)$ into \mathcal{S} that have left limits in \mathcal{S} on $(0, \infty)$. We endow this space with the Skorohod J_1 -topology. All stochastic processes used in this paper will be assumed to have paths in $\mathbf{D}([0, \infty), \mathcal{S})$ for a suitable Polish space \mathcal{S} . For a finite non-negative Borel measure ξ on \mathbb{R}_+ and a ξ -integrable function $f : \mathbb{R}_+ \rightarrow \mathbb{R}$, define

$$\langle f, \xi \rangle = \int_{\mathbb{R}_+} f d\xi.$$

If $\xi = (\xi_1, \dots, \xi_d)$ is a vector of such measures, then we use $\langle f, \xi \rangle$ to denote the

vector $(\langle f, \xi_1 \rangle, \dots, \langle f, \xi_d \rangle)$. All functions $f : \mathbb{R}_+ \rightarrow \mathbb{R}$ are extended to be identically zero on $(-\infty, 0)$ so that $f(\cdot - x)$ is well defined on \mathbb{R}_+ for all $x > 0$. Let $\chi : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ denote the identity function $\chi(x) = x$.

Let \mathbf{M} be the set of finite non-negative Borel measures on \mathbb{R}_+ , endowed with the weak topology: $\xi^k \xrightarrow{\mathbf{w}} \xi$ in \mathbf{M} if and only if $\langle f, \xi^k \rangle \rightarrow \langle f, \xi \rangle$ as $k \rightarrow \infty$, for all $f \in \mathbf{C}_b(\mathbb{R}_+)$. For $\mathbf{I} \in \mathbb{N}$, let

$$\mathbf{M}^{\mathbf{I}} = \{(\xi_1, \dots, \xi_{\mathbf{I}}) : \xi_i \in \mathbf{M} \text{ for all } i \leq \mathbf{I}\}.$$

The spaces \mathbf{M} and $\mathbf{M}^{\mathbf{I}}$ are Polish spaces. Convergence in $\mathbf{M}^{\mathbf{I}}$ is also denoted $\xi^k \xrightarrow{\mathbf{w}} \xi$. The zero measure in \mathbf{M} is denoted $\mathbf{0}$.

2. Flow level model

2.1. Network structure

Consider a network with finitely many *resources* labelled by $j = 1, \dots, \mathbf{J}$, and a finite set of *routes* labelled by $i = 1, \dots, \mathbf{I}$. A route i is a non-empty subset of $\{1, \dots, \mathbf{J}\}$, interpreted as the set of resources used by the route. Let A be the $\mathbf{J} \times \mathbf{I}$ incidence matrix satisfying $A_{ji} = 1$ if resource j is used by route i , and $A_{ji} = 0$ otherwise. Since each route is a non-empty subset of $\{1, \dots, \mathbf{J}\}$, no column of A is identically zero.

A *flow* on route i is the continuous transfer of a document through the resources used by the route. Assume that, while being transferred, a flow takes simultaneous possession of all resources on its route. The *processing rate* allocated to a flow is the rate at which the document associated with the flow is being transferred. There may be multiple flows on a route, and the *bandwidth* Λ_i allocated to route i is the sum of the processing rates allocated to flows on route i . The *bandwidth allocated through resource* j is the sum of the bandwidths allocated to routes using resource j . Assume that each resource $j \leq \mathbf{J}$ has finite *capacity* $C_j > 0$, interpreted as the maximum bandwidth that can be allocated through it. Let $C = (C_1, \dots, C_{\mathbf{J}})$ be the vector of capacities in $\mathbb{R}_+^{\mathbf{J}}$. Then any vector $\Lambda = (\Lambda_1, \dots, \Lambda_{\mathbf{I}})$ of bandwidth allocations must satisfy

$$A\Lambda \leq C.$$

2.2. Bandwidth sharing policy

We consider the network operating under a bandwidth sharing policy first introduced by Mo and Walrand [14]. Bandwidth is dynamically allocated to routes as a function of the number of flows on all routes, and the resulting allocation is shared equally among individual flows on each route.

Let $Z_i(t)$ denote the number of flows on route $i \leq \mathbf{I}$ at time t , and let $Z(t) = (Z_1(t), \dots, Z_{\mathbf{I}}(t))$ be the corresponding vector in $\mathbb{R}_+^{\mathbf{I}}$. The bandwidth allocated to route i at time t is a function of the vector $Z(t)$ and is denoted $\Lambda_i(Z(t))$. The corresponding vector of bandwidth allocations at time t is given by $\Lambda(Z(t)) = (\Lambda_1(Z(t)), \dots, \Lambda_{\mathbf{I}}(Z(t)))$. Although the coordinates of $Z(\cdot)$ are non-negative and integer valued, the function Λ is defined on the entire orthant $\mathbb{R}_+^{\mathbf{I}}$ to accommodate fluid analogues of $Z(\cdot)$ later.

Fix a parameter $\alpha \in (0, \infty)$ and a vector of strictly positive weights $\kappa = (\kappa_1, \dots, \kappa_{\mathbf{I}})$. For $z \in \mathbb{R}_+^{\mathbf{I}}$, let $\mathcal{I}_0(z) = \{i \leq \mathbf{I} : z_i = 0\}$ and $\mathcal{I}_+(z) = \{i \leq \mathbf{I} : z_i > 0\}$.

Let $\mathbb{O}(z) = \{\lambda \in \mathbb{R}_+^{\mathbf{I}} : \lambda_i = 0 \text{ for all } i \in \mathcal{I}_0(z)\}$. Define a function $G_z : \mathbb{R}_+^{\mathbf{I}} \rightarrow [-\infty, \infty)$ by

$$(2.1) \quad G_z(\lambda) = \begin{cases} \sum_{i \in \mathcal{I}_+(z)} \kappa_i z_i^\alpha \frac{\lambda_i^{1-\alpha}}{1-\alpha}, & \alpha \in (0, \infty) \setminus \{1\}, \\ \sum_{i \in \mathcal{I}_+(z)} \kappa_i z_i \log \lambda_i, & \alpha = 1, \end{cases}$$

where the value of $G_z(\lambda)$ is taken to be $-\infty$ if $\alpha \in [1, \infty)$ and $\lambda_i = 0$ for some $i \in \mathcal{I}_+(z)$. For each $z \in \mathbb{R}_+^{\mathbf{I}}$, define $\Lambda(z)$ as the unique vector $\lambda \in \mathbb{R}_+^{\mathbf{I}}$ that solves the optimization problem:

$$(2.2) \quad \text{maximize} \quad G_z(\lambda)$$

$$(2.3) \quad \text{subject to} \quad A\lambda \leq C$$

$$(2.4) \quad \text{over} \quad \mathbb{O}(z).$$

For existence, uniqueness and other properties of the solution $\Lambda(z)$ of this convex optimization problem, see for example Appendix A in [9]. (Although A is assumed to have full row rank in [9], that property is not needed for the results proved in Appendix A there.) The resulting allocation is called a *weighted α -fair allocation*, and the function $\Lambda : \mathbb{R}_+^{\mathbf{I}} \rightarrow \mathbb{R}_+^{\mathbf{I}}$ is called a *weighted α -fair bandwidth sharing policy*. Note that by (2.3),

$$(2.5) \quad \sup_{z \in \mathbb{R}_+^{\mathbf{I}}} \|\Lambda(z)\| \leq \|C\|.$$

Note also that for any $z \in \mathbb{R}_+^{\mathbf{I}}$, $\Lambda_i(z) = 0$ for all $i \in \mathcal{I}_0(z)$. This implies that no bandwidth is allocated to routes with no flows. The bandwidth $\Lambda_i(Z(t))$ allocated to route i at time t is shared equally by all flows on the route. That is, if there are $Z_i(t) > 0$ flows on route i at time t , then each flow on route i is allocated a processing rate of $\Lambda_i(Z(t))/Z_i(t)$ at time t .

When $\kappa_i = 1$ for all $i \leq \mathbf{I}$, the cases $\alpha \rightarrow 0$, $\alpha \rightarrow 1$, and $\alpha \rightarrow \infty$ correspond respectively to a bandwidth allocation which achieves *maximum throughput*, is *proportionally fair* or is *max-min fair* [1, 14]. Weighted α -fair allocations provide a tractable theoretical abstraction of decentralized packet-based congestion control algorithms such as TCP, the transmission control protocol of the Internet, particularly when $\alpha = 2$ and κ_i is the reciprocal of the square of the round trip time on route i .

2.3. Stochastic model

Fix a network structure (A, C) and a weighted α -fair bandwidth sharing policy Λ with parameters (α, κ) . Our stochastic model of document flows consists of the following: a collection of stochastic primitives $E_1, \dots, E_{\mathbf{I}}$ and $\{v_{1k}\}_{k=1}^\infty, \dots, \{v_{\mathbf{I}k}\}_{k=1}^\infty$ describing the arrivals of document flows (including their sizes) to the network, a random initial condition $\mathcal{Z}(0) \in \mathbf{M}^{\mathbf{I}}$ specifying the state of the system at time zero, and a collection of performance processes describing the time evolution of the system state. The performance processes are defined in terms of the primitives and initial condition through a set of descriptive equations.

The stochastic primitives consist of an *exogenous arrival process* E_i and a sequence of *document sizes* $\{v_{ik}\}_{k=1}^\infty$ for each route $i \leq \mathbf{I}$. The arrival process E_i is

a rate $\nu_i > 0$ delayed renewal process with k th jump time U_{ik} . For $t \geq 0$, $E_i(t)$ represents the number of flows that have arrived to route i during the time interval $(0, t]$. The k th such arrival is called flow k for route i and arrives at time U_{ik} ; flows already on route i at time zero are called *initial flows* for route i .

For each $i \leq \mathbf{I}$ and $k \geq 1$, the random variable v_{ik} represents the initial size of the document associated with flow k for route i . This is the cumulative amount of processing that must be allocated to the flow to complete its transfer through the network. The flow is considered to depart or to become inactive once it receives this total amount of processing. Assume that the random variables $\{v_{ik}\}_{k=1}^{\infty}$ are strictly positive and form an independent and identically distributed sequence with common distribution ϑ_i on \mathbb{R}_+ . Assume that the mean $\langle \chi, \vartheta_i \rangle \in (0, \infty)$ and let $\mu_i = \langle \chi, \vartheta_i \rangle^{-1}$. Define the *traffic intensity* on route i by $\rho_i = \nu_i / \mu_i$.

The initial condition specifies $Z(0) = (Z_1(0), \dots, Z_{\mathbf{I}}(0))$, the number of initial flows on each route at time zero, as well as the initial sizes of the documents on these flows at time zero. Assume that the components of $Z(0)$ are nonnegative, integer valued random variables. The initial document sizes of the initial flows on route $i \leq \mathbf{I}$ are the first $Z_i(0)$ elements of a sequence $\{\tilde{v}_{il}\}_{l=1}^{\infty}$ of strictly positive random variables.

The *performance processes* consist of a measure valued process \mathcal{Z} , taking values in $\mathbf{D}([0, \infty), \mathbf{M}^{\mathbf{I}})$, and a collection of auxiliary processes (Z, T, U, W) . The process $Z = (Z_1, \dots, Z_{\mathbf{I}})$ takes values in $\mathbf{D}([0, \infty), \mathbb{R}_+^{\mathbf{I}})$. For $i \leq \mathbf{I}$ and $t \geq 0$, $Z_i(t)$ is the number of (active) flows on route i at time t . Recall that at time t , the bandwidth allocated to route i is $\Lambda_i(Z(t))$, and this bandwidth is shared equally by all $Z_i(t)$ flows on route i ; each such flow receives a processing rate of $\Lambda_i(Z(t))/Z_i(t)$, which equals zero by convention if $Z_i(t) = 0$. Thus, a flow that is active on route i during a time interval $[s, t] \subset [0, \infty)$ receives *cumulative service during* $[s, t]$ equal to

$$(2.6) \quad S_i(s, t) = \int_s^t \frac{\Lambda_i(Z(u))}{Z_i(u)} du.$$

Consider the k th flow for route i . This flow arrives at time U_{ik} and has initial document size v_{ik} . At time $t \geq U_{ik}$, the cumulative service received by this flow during $[U_{ik}, t]$ equals $S_i(U_{ik}, t) \wedge v_{ik}$. The amount of service still required therefore equals $(v_{ik} - S_i(U_{ik}, t))^+$. (Once this latter quantity becomes zero, the flow becomes inactive, i.e., it departs from the system.) For $t \geq 0$, $k \leq E_i(t)$, and $l \leq Z_i(0)$, define the *residual document size at time t* of the k th flow for route i , and the l th initial flow for route i , by

$$(2.7) \quad \begin{aligned} v_{ik}(t) &= (v_{ik} - S_i(U_{ik}, t))^+, \\ \tilde{v}_{il}(t) &= (\tilde{v}_{il} - S_i(0, t))^+. \end{aligned}$$

The measure valued process $\mathcal{Z} = (\mathcal{Z}_1, \dots, \mathcal{Z}_{\mathbf{I}})$ is called the *state descriptor*; it tracks the residual document sizes of the flows for all routes at any given time. Let $\delta_x^+ \in \mathbf{M}$ denote the Dirac measure at x if $x \in (0, \infty)$ and let $\delta_0^+ = \mathbf{0}$. For $t \geq 0$ and $i \leq \mathbf{I}$,

$$(2.8) \quad \mathcal{Z}_i(t) = \sum_{l=1}^{Z_i(0)} \delta_{\tilde{v}_{il}(t)}^+ + \sum_{k=1}^{E_i(t)} \delta_{v_{ik}(t)}^+.$$

We can recover Z from \mathcal{Z} by

$$(2.9) \quad Z_i(t) = \langle 1, \mathcal{Z}_i(t) \rangle, \quad \text{for all } t \geq 0, i \in \mathbf{I}.$$

The process T takes values in $\mathbf{D}([0, \infty), \mathbb{R}_+^{\mathbf{I}})$ and tracks the *cumulative bandwidth allocated* to each route. For $t \geq 0$ and $i \leq \mathbf{I}$,

$$(2.10) \quad T_i(t) = \int_0^t \Lambda_i(Z(s)) ds.$$

The process U takes values in $\mathbf{D}([0, \infty), \mathbb{R}_+^{\mathbf{J}})$ and tracks the *cumulative unused bandwidth capacity* of each resource. For $t \geq 0$,

$$(2.11) \quad U(t) = Ct - AT(t).$$

The process W takes values in $\mathbf{D}([0, \infty), \mathbb{R}_+^{\mathbf{I}})$ and tracks the immediate amount of work still to be transferred on each route. For $t \geq 0$,

$$(2.12) \quad W(t) = \langle \chi, \mathcal{Z}(t) \rangle.$$

Recall that $\chi(x) = x$ and that integration against the vector of measures $\mathcal{Z}(t)$ is interpreted componentwise.

3. Fluid model

Fix a network structure (A, C) and a weighted α -fair bandwidth sharing policy Λ with parameters (α, κ) . This section defines a fluid analogue of the stochastic model introduced in Section 2.3. In [8], under mild assumptions, it was shown that this fluid model is a first order approximation (under functional law of large numbers scaling) to the stochastic model. For details of when this approximation holds, we refer the reader to [8].

As in the stochastic model, fix a vector of positive arrival rates $\nu = (\nu_1, \dots, \nu_{\mathbf{I}})$ and a vector of probability measures $\vartheta = (\vartheta_1, \dots, \vartheta_{\mathbf{I}})$ in $\mathbf{M}^{\mathbf{I}}$, satisfying the assumptions of Section 2. Recall that $\mu_i = \langle \chi, \vartheta_i \rangle^{-1}$ and $\rho_i = \nu_i / \mu_i$ for each $i \leq \mathbf{I}$. The fluid model consists of a deterministic measure valued function of time, called the *fluid model solution*, and a collection of auxiliary functions of time defined below.

Definition 3.1 *Given a continuous function $\zeta : [0, \infty) \rightarrow \mathbf{M}^{\mathbf{I}}$, define the auxiliary functions (z, τ, u, w) of ζ , with respect to the data $(A, C, \alpha, \kappa, \nu, \vartheta)$, by*

$$\begin{aligned} z(t) &= \langle 1, \zeta(t) \rangle, \\ \tau_i(t) &= \int_0^t \left(\Lambda_i(z(s)) 1_{(0, \infty)}(z_i(s)) + \rho_i 1_{\{0\}}(z_i(s)) \right) ds, \quad i \leq \mathbf{I}, \\ u(t) &= Ct - A\tau(t), \\ w(t) &= \langle \chi, \zeta(t) \rangle, \end{aligned}$$

for all $t \geq 0$.

Note that $z(t)$ and $\tau(t)$ take values in $\mathbb{R}_+^{\mathbf{I}}$ and $u(t)$ takes values in $\mathbb{R}_+^{\mathbf{J}}$. On the other hand, $w(t)$ takes values in $[0, \infty]^{\mathbf{I}}$, as the fluid model solution need not have finite first moments. The function ζ is the fluid analogue of the measure valued process \mathcal{Z} . The functions z, τ, u , and w , are fluid analogues of the processes Z, T, U , and W , which keep track of queue length, cumulative bandwidth allocation, unused capacity and workload, respectively. The equation satisfied by τ_i may seem counterintuitive at first. However, the presence of the term involving ρ_i is accounted for by the fact that in passing to a fluid limit of the stochastic model, bandwidth allocations made

when a queue length is near zero in the stochastic model are averaged with the zero bandwidth allocations made when a queue length is zero. The fact that ρ_i is the correct form here is related to the fact that when the fluid workload function w_i is real-valued, at a positive time where it is differentiable (which occurs a.e.) and at which the value of w_i is zero, the derivative of the workload function must be zero (cf. (3.2) below).

The notion of a fluid model solution is defined via projections against test functions in the class

$$\mathcal{C} = \{f \in \mathbf{C}_b^1(\mathbb{R}_+) : f(0) = f'(0) = 0\}.$$

Definition 3.2 *A fluid model solution for the data $(A, C, \alpha, \kappa, \nu, \vartheta)$ is a continuous function $\zeta : [0, \infty) \rightarrow \mathbf{M}^{\mathbf{I}}$ that, together with its first three auxiliary functions (z, τ, u) , satisfies*

- (i) $\|\langle 1_{\{0\}}, \zeta(t) \rangle\| = 0$ for all $t \geq 0$,
- (ii) u_j is non-decreasing for all $j \leq \mathbf{J}$,
- (iii) for each $f \in \mathcal{C}$, $i \leq \mathbf{I}$, and $t \geq 0$,

$$(3.1) \quad \langle f, \zeta_i(t) \rangle = \langle f, \zeta_i(0) \rangle - \int_0^t \langle f', \zeta_i(s) \rangle \frac{\Lambda_i(z(s))}{z_i(s)} ds \\ + \nu_i \langle f, \vartheta_i \rangle \int_0^t 1_{(0, \infty)}(z_i(s)) ds.$$

Recall that in (3.1), the integrand in the first integral term is defined to be zero when its denominator is zero. When there is mass present in the system, the first integral term in (3.1) relates to the movement to the left of the random measure ζ_i at the processing rate of $\Lambda_i(z(s))/z_i(s)$, and the second integral term corresponds to new infusion of mass due to new arrivals coming at a rate of ν_i with a distribution of ϑ_i for route i . The appearance of the indicator function in the last term may seem counterintuitive. The correct form for this term is discerned using the fact that at a time $t > 0$ for which $z_i(t) = 0$ and $\langle f, \zeta_i(\cdot) \rangle$ is differentiable, we must have that the time derivative of $\langle f, \zeta_i(\cdot) \rangle$ is zero. Since the integrand in the first integral term is zero by definition at such times, the same must be true for the second integral term.

When the initial fluid workload is finite, we have the following result which is proved in [8].

Lemma 3.3 *Suppose ζ is a fluid model solution with finite initial workload, i.e., $w_i(0) = \langle \chi, \zeta_i(0) \rangle < \infty$ for all $i \leq \mathbf{I}$. Then, for each $i \leq \mathbf{I}$ and $t \geq 0$,*

$$(3.2) \quad w_i(t) = w_i(0) + \int_0^t (\rho_i - \Lambda_i(z(s))) 1_{(0, \infty)}(z_i(s)) ds \\ = w_i(0) + \rho_i t - \tau_i(t).$$

In particular, the fluid workload $w_i(t)$ is finite for all $t \geq 0$ and $i \leq \mathbf{I}$.

For later use, when $\zeta(\cdot)$ is a fluid model solution with finite initial workload and fluid workload function w , we define $v : [0, \infty) \rightarrow \mathbb{R}_+^{\mathbf{J}}$ by

$$(3.3) \quad v(t) = Aw(t), \quad t \geq 0,$$

so that the j th component of $v(t)$ defines the fluid workload at resource j at time t . In other words, v is a resource level workload, whereas w is a route level workload.

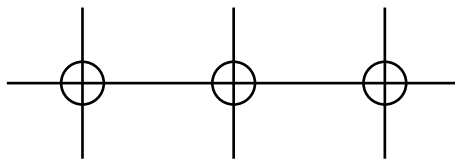


FIG 1. A linear network with 3 resources (denoted by circles) and 4 routes (denoted by line segments)

4. Fluid stability for some network topologies

In this section, we use Lyapunov functions to show stability of fluid model solutions with finite initial workload for two types of network topologies, linear networks and tree networks, under the nominal condition:

$$(4.1) \quad \sum_{i \leq \mathbf{I}} A_{ji} \rho_i < C_j \quad \text{for all } j \leq \mathbf{J},$$

i.e., the average load placed on each resource is less than its capacity. (We note that it follows from the characterization of invariant states for the fluid model given in [8] that under this nominal condition, the only invariant state is the zero state.) We assume that (4.1) holds henceforth. Let

$$(4.2) \quad \varepsilon = \min_{j \leq \mathbf{J}} \left(C_j - \sum_{i \leq \mathbf{I}} A_{ji} \rho_i \right),$$

so that $\varepsilon > 0$.

4.1. Linear Network

A linear network consists of \mathbf{J} resources and $\mathbf{I} = \mathbf{J} + 1$ routes where route j consists of resource j alone for $j = 1, \dots, \mathbf{J}$ and route $\mathbf{J} + 1$ consists of all of the \mathbf{J} resources. A schematic of such a network is shown in Figure 1 for $\mathbf{J} = 3$.

Consider a fluid model solution ζ with finite initial workload $w(0) = \langle \chi, \zeta(0) \rangle$ and associated resource level workload function v as defined in (3.3). Consider the Lyapunov function $H : \mathbb{R}_+^{\mathbf{J}} \rightarrow \mathbb{R}_+$ defined by

$$(4.3) \quad H(v) = \max_{j \leq \mathbf{J}} v_j.$$

A Lipschitz continuous function $x : [0, \infty) \rightarrow \mathbb{R}$ is absolutely continuous, hence it is differentiable almost everywhere and it can be recovered by integration from its a.e. defined derivative. We call a point at which such a Lipschitz continuous function is differentiable a *regular point* for the function.

The auxiliary functions $\tau_i : [0, \infty) \rightarrow \mathbb{R}_+$, $i \leq \mathbf{I}$, are Lipschitz continuous, and hence so too are u_j , $j \leq \mathbf{J}$, w_i , $i \leq \mathbf{I}$ and v_j , $j \leq \mathbf{J}$. The function $H(\cdot)$ is Lipschitz continuous and hence so too is $H(v(\cdot))$. Let $t > 0$ be a regular point for $H(v(\cdot))$, $\tau_i, w_i : i \leq \mathbf{I}$, $u_j, v_j : j \leq \mathbf{J}$, such that for all $i \leq \mathbf{I}$,

$$(4.4) \quad \dot{\tau}_i(t) = \Lambda_i(z(t))1_{(0, \infty)}(z_i(t)) + \rho_i 1_{\{0\}}(z_i(t)),$$

(such points occur a.e.). Suppose that $H(v(t)) > 0$ and let

$$\mathcal{J}_t = \{j \leq \mathbf{J} : H(v(t)) = v_j(t)\}.$$

Then,

$$\begin{aligned} H(v(t)) &= v_j(t) && \text{for } j \in \mathcal{J}_t, \\ H(v(t)) &> v_j(t) && \text{for } j \notin \mathcal{J}_t, \end{aligned}$$

and by the fact that $t > 0$ is a regular point for $H(v(\cdot))$ and $v_j, j \in \mathcal{J}_t$, we have (cf. [6], Section 3),

$$(4.5) \quad \frac{d}{dt}H(v(t)) = \dot{v}_j(t) \quad \text{for all } j \in \mathcal{J}_t.$$

Now, by Lemma 3.3 and (4.4),

$$(4.6) \quad \dot{v}_j(t) = \sum_{i \leq \mathbf{I}} A_{ji} \dot{w}_i(t) = \sum_{i \leq \mathbf{I}} A_{ji} (\rho_i - \Lambda_i(z(t))) 1_{(0, \infty)}(z_i(t)).$$

We consider two cases.

Case (a). Suppose $z_j(t) > 0$ for some $j \in \mathcal{J}_t$. Then by the definition of $\Lambda(\cdot)$ and the fact that route j just contains resource j , it follows that the full capacity of resource j will be used by $\Lambda(z(t))$, i.e.,

$$\sum_{i \leq \mathbf{I}} A_{ji} \Lambda_i(z(t)) 1_{(0, \infty)}(z_i(t)) = C_j.$$

Thus, for this $j \in \mathcal{J}_t$, (4.6) becomes

$$\begin{aligned} \dot{v}_j(t) &= \sum_{i \leq \mathbf{I}} A_{ji} \rho_i 1_{(0, \infty)}(z_i(t)) - C_j \\ &\leq \sum_{i \leq \mathbf{I}} A_{ji} \rho_i - C_j \\ &\leq -\varepsilon < 0, \end{aligned}$$

by the assumption (4.1). It follows that in Case (a),

$$\frac{d}{dt}H(v(t)) \leq -\varepsilon.$$

Case (b). Suppose $z_j(t) = 0$ for all $j \in \mathcal{J}_t$. Then $w_j(t) = 0$ for all $j \in \mathcal{J}_t$ and since

$$v_j(t) = w_j(t) + w_{\mathbf{J}+1}(t), \quad j \leq \mathbf{J},$$

we have

$$v_j(t) = w_{\mathbf{J}+1}(t) \quad \text{for all } j \in \mathcal{J}_t.$$

Since

$$w_{\mathbf{J}+1}(t) \leq v_l(t) < v_j(t) \quad \text{for all } l \notin \mathcal{J}_t, j \in \mathcal{J}_t,$$

it follows that $\mathcal{J}_t = \{1, \dots, \mathbf{J}\}$, and

$$v_j(t) = w_{\mathbf{J}+1}(t) \quad \text{for all } j \leq \mathbf{J}.$$

By Lemma 3.3 and (4.4), since $H(v(t)) = w_{\mathbf{J}+1}(t) > 0$ and hence $z_{\mathbf{J}+1}(t) > 0$, we have

$$(4.7) \quad \dot{w}_{\mathbf{J}+1}(t) = \rho_{\mathbf{J}+1} - \Lambda_{\mathbf{J}+1}(z(t)).$$

Since $z_j(t) = 0$ for all $j \leq \mathbf{J}$, $\Lambda_j(z(t)) = 0$ for all $j \leq \mathbf{J}$, and it follows from the definition of $\Lambda(z(t))$ as the solution of an optimization problem where at least one constraint must be binding, that there is at least one $j \leq \mathbf{J}$ such that

$$\Lambda_{\mathbf{J}+1}(z(t)) = C_j.$$

Here $C_j > \rho_j + \rho_{\mathbf{J}+1}$ by (4.1). It follows that, for this j ,

$$\dot{w}_{\mathbf{J}+1}(t) = \rho_{\mathbf{J}+1} - C_j < \rho_j + \rho_{\mathbf{J}+1} - C_j \leq -\varepsilon < 0.$$

Hence in Case (b),

$$\frac{d}{dt}H(v(t)) = \dot{w}_{\mathbf{J}+1}(t) \leq -\varepsilon.$$

Thus, in either Case (a) or (b), at the regular point $t > 0$,

$$\frac{d}{dt}H(v(t)) \leq -\varepsilon \quad \text{when } H(v(t)) > 0.$$

Since $H(v(\cdot))$ is non-negative, it follows from Lemma 2.2 of Dai and Weiss [6] that

$$H(v(t)) = 0 \quad \text{for all } t \geq \delta = H(v(0))/\varepsilon.$$

We summarize the above analysis as follows.

Lemma 4.1 *Consider a linear network satisfying the condition (4.1) and let $\varepsilon > 0$ be as defined in (4.2). Suppose that ζ is a fluid model solution with finite initial workload $w(0) = \langle \chi, \zeta(0) \rangle$. Then*

$$\zeta(t) = 0 \quad \text{for all } t \geq \delta,$$

where $\delta = \max_{j \leq \mathbf{J}} v_j(0)/\varepsilon$.

In the above sense, the fluid model for any linear network is stable under the natural condition (4.1).

4.2. Tree network

As pointed out by Bonald and Proutière [2], tree networks, as illustrated in Figure 2, are practically interesting as they may represent an access network consisting of several multiplexing stages. Furthermore [2], they typically exhibit sensitivity to document size distributions.

A tree network consists of $\mathbf{J} \geq 2$ resources and $\mathbf{I} = \mathbf{J} - 1$ routes such that a single resource (labeled \mathbf{J} and referred to as the trunk) belongs to all routes and each of the other resources (labeled by $1, \dots, \mathbf{J} - 1$) belongs to a single route.

Proceeding in a similar manner to that for the linear network, consider a fluid model solution ζ with finite initial workload $\langle \chi, \zeta(0) \rangle$. We use the total workload function $H : \mathbb{R}_+^{\mathbf{J}-1} \rightarrow \mathbb{R}_+$ defined by

$$(4.8) \quad H(w) = \sum_{i=1}^{\mathbf{J}-1} w_i$$

as a Lyapunov function. Note that $H(w(\cdot)) = v_{\mathbf{J}}(\cdot)$, the resource level workload for the trunk resource \mathbf{J} . Suppose $t > 0$ is a regular point for $\tau_i, i \leq \mathbf{J} - 1$, such that for all $i \leq \mathbf{J} - 1$,

$$(4.9) \quad \dot{\tau}_i(t) = \Lambda_i(z(t))1_{(0,\infty)}(z_i(t)) + \rho_i 1_{\{0\}}(z_i(t)),$$

(such points t occur a.e.) Then t is a regular point for all $w_i, i \leq \mathbf{J} - 1$. Suppose $H(w(t)) > 0$. Then by Lemma 3.3 and (4.9) we have,

$$(4.10) \quad \frac{d}{dt}H(w(t)) = \sum_{i \leq \mathbf{J}-1} (\rho_i - \Lambda_i(z(t)))1_{(0,\infty)}(z_i(t)).$$

We consider two cases.

Case (a). Suppose

$$\sum_{i \leq \mathbf{J}-1} \Lambda_i(z(t))1_{(0,\infty)}(z_i(t)) = C_{\mathbf{J}}.$$

Then by (4.10) and (4.1) with $j = \mathbf{J}$, we have

$$\frac{d}{dt}H(w(t)) \leq \sum_{i \leq \mathbf{J}-1} \rho_i - C_{\mathbf{J}} \leq -\varepsilon.$$

Case (b). Suppose

$$\sum_{i \leq \mathbf{J}-1} \Lambda_i(z(t))1_{(0,\infty)}(z_i(t)) < C_{\mathbf{J}}.$$

Then, by the definition of $\Lambda(z(t))$, we must have

$$(4.11) \quad \Lambda_i(z(t)) = C_i \quad \text{for those } i \leq \mathbf{J} - 1 \text{ satisfying } z_i(t) > 0.$$

For if not, the value of $\Lambda_i(z(t))$ could be increased on some non-empty route i without exceeding the capacity of the resources i and \mathbf{J} on that route. From (4.10) and (4.11), it follows that

$$\frac{d}{dt}H(w(t)) = \sum_{i \leq \mathbf{J}-1} (\rho_i - C_i)1_{(0,\infty)}(z_i(t)) \leq -\varepsilon < 0,$$

since $\rho_i < C_i$ for all $i \leq \mathbf{J} - 1$ by (4.1), and since $z_i(t) > 0$ for some $i \leq \mathbf{J} - 1$ as $H(w(t)) > 0$.

Thus, in either Case (a) or (b),

$$\frac{d}{dt}H(w(t)) \leq -\varepsilon < 0, \quad \text{when } H(w(t)) > 0.$$

Since $H(w(\cdot))$ is non-negative, it follows from Lemma 2.2 of [6] that

$$H(w(t)) = 0 \text{ for all } t \geq \delta = H(w(0))/\varepsilon.$$

We summarize the above analysis as follows.

Lemma 4.2 *Consider a tree network satisfying the condition (4.1) and let $\varepsilon > 0$ be as defined in (4.2). Suppose that ζ is a fluid model solution with finite initial workload $w(0) = \langle \chi, \zeta(0) \rangle$. Then*

$$\zeta(t) = 0 \text{ for all } t \geq \delta,$$

where $\delta = \sum_{i \leq \mathbf{J}-1} w_i(0)/\varepsilon$.

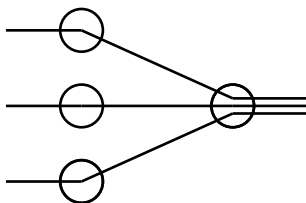


FIG 2. A tree network with 4 resources and 3 routes

References

- [1] BONALD, T. and MASSOULIÉ, L. (2001). Impact of fairness on Internet performance. In *Proceedings of ACM Sigmetrics 2001*.
- [2] BONALD, T. and PROUTIERE, A. (2003). Insensitive bandwidth sharing in data networks. *Queueing Systems* **44** 69–100.
- [3] BRAMSON, M. (2005). Stability of networks for max-min fair routing. Presentation at the 13th INFORMS Applied Probability Conference, Ottawa.
- [4] CHIANG, M., SHAH, D. and TANG, A. K. (2006). Stochastic stability under network utility maximization: general file size distribution. In *Proceedings of Allerton*.
- [5] DAI, J. G. (1995). On positive Harris recurrence of multiclass queueing networks: a unified approach via fluid limit models. *Annals of Applied Probability* **5** 49–77.
- [6] DAI, J. G. and WEISS, G. (1996). Stability and instability of fluid models for re-entrant lines. *Mathematics of Operations Research* **21** 115–134.
- [7] DE VECIANA, G., KONSTANTOPOULOS, T. and LEE, T.-J. (2001). Stability and performance analysis of networks supporting elastic services. *IEEE/ACM Transactions on Networking* **9** 2–14.
- [8] GROMOLL, H. C. and WILLIAMS, R. J. (2006). Fluid limit of a network with fair bandwidth sharing and general document size distributions. Submitted.
- [9] KELLY, F. P. and WILLIAMS, R. J. (2004). Fluid model for a network operating under a fair bandwidth sharing policy. *Annals of Applied Probability* **14** 1055–1083.
- [10] LAKSHMIKANTHA, A., BECK, C. L. and SRIKANT, R. (2004). Connection level stability analysis of the internet using the sum of squares (sos) techniques. In *Conference on Information Sciences and Systems, Princeton*.
- [11] LIN, X., SHROFF, N. and SRIKANT, R. (2006). On the connection-level stability of congestion-controlled communication networks. *IEEE Transactions on Information Theory* To appear.
- [12] MASSOULIÉ, L. (2007). Structural properties of proportional fairness: stability and insensitivity. *Annals of Applied Probability* **17** 809–839.
- [13] MASSOULIÉ, L. and ROBERTS, J. (2000). Bandwidth sharing and admission control for elastic traffic. *Telecommunication Systems* **15** 185–201.
- [14] MO, J. and WALRAND, J. (2000). Fair end-to-end window-based congestion control. *IEEE/ACM Transactions on Networking* **8** 556–567.