PDE models for population and residual work applied to peer-to-peer networks

Fernando Paganini and Andrés Ferragut
Universidad ORT Uruguay

***Abstract*—This paper studies partial differential equations that have recently been proposed as fluid models for queueing networks, where both populations and residual workloads must be accounted for. After reviewing these models in general, we focus on an application to peer-to-peer networks, where the dynamics must keep track of the download progress of a population of peers as content propagates among them through file sharing. Applying control-theoretic methods to this PDE yields a series of analytical results, in particular: local stability analysis of the equilibrium is proved through a small-gain argument on an appropriate feedback loop; variability around this equilibrium in the presence of random noise is analyzed through the frequency domain; and transient studies are performed to compute completion times.**

I. Introduction

Research on data networks is relying increasingly on fluid models (continuous variables, differential equations) as compared to the classical discrete models of queueing theory. A first reason for this shift is that at the microscopic level of packets, the assumptions of queueing theory become dubious given the complex mechanisms in place (e.g. the window control mechanisms of TCP), but a macroscopic model can yield satisfactory predictions of rates (see e.g. [15]).

The queueing model remains compelling to study the statistical multiplexing of *jobs* arriving and being processed by the network. Indeed, researchers since [14] have considered models that combine a discrete random process for the population of flows in the network, and a continuous variable description of the rate allocation among them. Of particular interest has been the study of stochastic stability of such queues for the resource allocation models of the Internet [1], [4]. Nevertheless, fluid models appear as well in this higher layer dynamics for a second, mathematical reason: queues are often difficult to solve analytically, but statements can be made about *scaling limits* of such queues as the number of flows grows large [8], replacing discrete counts with real variables. This opens the door for differential equation techniques for these problems, such as Lyapunov methods for the aforementioned stability question [1]. Most recently, fluid models have been extended to cover another aspect of queue dynamics: characterizing residual workloads of the files in progress, a requirement when the job sizes are not exponential [7]. In this regard, a key observation of [11] was to describe the evolution of the system state through a partial differential equation (PDE), which opened the door for a Lyapunov proof of stability.

Recent research on peer-to-peer (P2P) networks has followed an analogous path to the one described above. Initial models based on discrete queues were proposed in [16] to model peer populations, but such queues were difficult to solve explicitly. In [12], a fluid model based on differential equations was proposed, leading to characterizations of the equilibrium achieved by the swarm population and its stability [12], [13]. Noting that this model does not provide detail of the propagation of content within the P2P swarm, the authors proposed in a recent paper [5] a PDE model for these dynamics, studying various file sharing policies. Equilibrium conditions and some partial stability results were given.

In this paper we review this modeling technique and extend its application to the P2P problem through a series of new results on the dynamics of peer populations and downloaded content, invoking methods of control theory. First, in Section II we discuss general conditions (namely, processor sharing disciplines) to which such PDE models apply, focusing then in Section III on the P2P problem, reviewing the equilibrium from [5]. Section IV provides a proof of local stability of the equilibrium, through a small-gain argument. Section V shows how variability can be analyzed through transfer function methods. Section VI gives results on transient performance. Conclusions are given in Section VII.

II. PDE models for processor sharing networks

Consider a communication network that receives jobs (e.g. download requests) at rate λ arrivals/second. The fraction of jobs greater than size σ is denoted by $H(\sigma)$, i.e. this is the complementary cumulative distribution function (CCDF) of the workload distribution. We normalize file-size units to have a unit mean, i.e. $\int_0^\infty H(\sigma)d\sigma = 1$. A processor sharing (PS) discipline allocates a transmission rate r, simultaneously to each job present in the network. r will depend on the network state, and may possibly depend on time and the residual work σ the job possesses. This assumes a single class of jobs, below we give an example of the multi-class generalization.

A convenient way to describe the state of such system is by a function $F(t,\sigma)$ that counts the number of jobs present at time t that have residual workload larger than σ. This function is monotonically decreasing, satisfying $F(t,\infty) = 0$ and $F(t,0) = x(t)$, the total number of jobs in the network. In a discrete model, $F(t,\cdot)$ would decrease by unit steps at the locations of current residual workloads.

This work was supported by AFOSR-US under grant FA9550-09-1-0504, and ANII-Uruguay, grant PR_FCE_2009_1_2158.
E-mail:{paganini,ferragut}@ort.edu.uy.

In this paper we will work with a *fluid* version of this state, in which $F(t,\sigma)$ is taken to be real-valued and smooth, and postulate a differential equation for the state evolution, motivated by the following incremental analysis:

$$F(t+dt,\sigma) = \lambda H(\sigma)dt + F(t,\sigma + rdt). \quad (1)$$

Given the current state $F(t,\sigma)$, its value after a small time interval dt is determined in (1) by two components:

- The number of new arrivals in $(t, t+dt]$ with workload larger than σ. Of the λdt total arrivals, a fraction $H(\sigma)$ has workload larger than σ.
- The number of jobs present at time t with residual work larger than $\sigma + rdt$. These jobs receive rate of service r and hence process rdt units of work, thus remaining above σ at time $t+dt$.

By subtracting $F(t,\sigma)$, dividing by dt and letting $dt \to 0$ we have the following evolution for the state, in the form of a transport partial differential equation:

$$\frac{\partial F}{\partial t} = \lambda H(\sigma) + r(F,t,\sigma)\frac{\partial F}{\partial \sigma}. \quad (2)$$

In the above dynamics, we have made explicit the dependence of the service rate on the current state of the network, as well as possibly on (t,σ). The precise form of this dependence is determined by the actual network providing the service. Some possibilities are indicated below.

1) Fixed per-flow capacity: Here the network assigns to each connection a rate $r = c$, independent of the network state. An example of this situation would be a client-server network where clients have a download capacity limit c (e.g. given by their domestic Internet access), but servers are greatly overprovisioned, so they can satisfy this demand to all downloaders present in the network.

2) Fixed service capacity with PS discipline: Here we have service capacity c shared equally among jobs present, i.e. $r = c/x = c/F(t,0)$. This could correspond to the situation of homogeneous TCP flows sharing a single bottleneck.

3) Multi-class network with α-fair resource allocation: This models TCP resource allocation over an arbitrary topology. The network is a set of links, indexed by l and of capacity c_l. For each route m across the network, introduce a class of jobs, with their corresponding arrival rate λ_m and file-size CCDF $H_m(\sigma)$. Routing is specified by the incidence matrix R ($R_{lm} = 1$ if route m uses link l and zero otherwise).

The system state is the vector indexed by m with components $F_m(t,\sigma)$, each satisfying an equation of the form (2), with a corresponding service rate r_m. These service rates are obtained as the solution to the Network Utility Maximization (NUM) problem

$$\max \sum_{m:x_m>0} x_m U_m(r_m), \text{ subject to} \sum_m R_{lm} x_m r_m \le c_l.$$

Here $x_m(t) = F_m(t,0)$, and $U_m(\cdot)$ is a utility function defined for flows of class m, which is taken to belong to the "α-fair" family of [10], namely with derivative $U_m'(r) = \kappa_m r^{-\alpha}$ for some strictly positive α. For background on these models for TCP we refer to [15].

Connection to M/G//PS queues*

In the domain of queueing theory, the system under consideration would be modeled by a random process of arrivals (e.g. Poisson of intensity λ), each job bringing an independent random workload of general file-size distribution, characterized by its CCDF $H(\sigma)$. Jobs are served simultaneously with rate r depending on the full network state, and possibly the job's residual workload σ.

One method to model this system as a Markov process is to introduce the state

$$\Phi_t = \sum_{i=1}^{x(t)} \delta_{\sigma_i(t)},$$

a measure with point masses at the locations of the residual workloads σ_i of the jobs $i = 1,\ldots,x(t)$ currently present. We refer to [6], [7] and references therein for the analysis of such measure-valued stochastic processes. If instead of Φ we write its complementary CDF

$$F(t,\sigma) := \Phi_t\big((\sigma,\infty)\big),$$

we would have a random process whose state is a decreasing step function in σ, as discussed before. Relating such random process with our deterministic smooth version of $F(t,\sigma)$ can be done through *scaling*; i.e. considering the limit of a family of such processes with re-scaled time and initial conditions. This is beyond the scope of this paper, we refer to [7] for extensive details, and to [11] for additional comments.

III. Application to peer-to-peer networks

PDE models of the type discussed here for P2P networks were considered in our previous paper [5], building on the previous ODE models of [12]. We summarize in this section some relevant results of [5], with minor notational changes.

In a P2P network, content is divided into small pieces so that peers downloading (termed *leechers*) can themselves contribute by uploading pieces they already have to others. There are also *seeders* who already own content and contribute to the upload without downloading. Let $x(t)$ denote the leecher population, $y(t)$ the number of seeders. All peers are assumed to have an upload bandwidth of μ, thus the maximum total upload bandwidth available is $R_{up} = \mu(x+y)$. The download rate r available to each leecher is constrained by this overall upload "budget", as well as by an individual download capacity limit $c > \mu$. In [5], we studied different alternatives for the function $r(F,t,\sigma)$ with these restrictions; we called the bandwidth sharing *efficient* when these are the only operative constraints, i.e. no spare bandwidth goes to waste. This is a reasonable assumption when there is enough piece diversity, as argued in [12].

The simplest among efficient bandwidth sharing policies is the processor sharing discipline, given by

$$r(F,y,\sigma) = \min\left\{\mu\frac{x+y}{x}, c\right\}. \quad (3)$$

This means all leechers present in the network receive equal service rate, constrained either by the maximum upload capacity, or by the download limit. This appears to be a good model for current BitTorrent systems [2], as supported by simulation-based studies [9].

With this assumption, our dynamic equation (2) becomes

$$\frac{\partial F}{\partial t} = \lambda H(\sigma) + \min\left\{\mu \frac{x+y}{x}, c\right\}\frac{\partial F}{\partial \sigma}.$$

We have not discussed yet the workload distribution $H(\sigma)$. Let us consider momentarily that we chose an exponential distribution of mean 1, $H(\sigma) = e^{-\sigma}$. Then it is easily shown that the preceding equation admits a solution with separation of variables, $F(t,\sigma) = x(t)e^{-\sigma}$, where $x(t)$ satisfies the ordinary differential equation

$$\dot{x} = \lambda - \min\left\{\mu(x+y), cx\right\}. \tag{4}$$

This is precisely the model proposed in [12] for the leecher dynamics (restricted to the efficient case, and without leechers abandoning prematurely). So we see that our PDE dynamics subsumes this earlier model. Now, is it natural to assume leechers have exponential workloads? In the case of a common content file of interest to all peers, it is more natural to assume the workload is deterministic, of unit size, corresponding to $H(\sigma) = \mathbf{1}_{[0,1)}(\sigma)$. To study such case, it will be essential to use the full PDE dynamics, there is no ODE simplification. We focus on this case from now on[1].

In this case no peers will have workloads larger than unity, so we can set $F(t,\sigma) = 0$ for $\sigma > 1$ and restrict the PDE to the interval $[0,1]$, as follows:

$$\frac{\partial F}{\partial t} = \lambda + \underbrace{\min\left\{\mu \frac{x+y}{x}, c\right\}}_{r}\frac{\partial F}{\partial \sigma}, \quad \sigma \in [0,1]. \tag{5}$$

The above assumes that $y(t)$ is exogenously given. The simplest alternative here is to take $y(t) \equiv y_0$, i.e. the seeder population is a set of fixed servers. This assumption, while restrictive, leads to some valuable results and insights, as shown in Section VI. The other choice is to select a dynamic equation for $y(t)$. Following [12], we will consider the dynamics by which leechers turn into seeders upon completion of their download (at $\sigma = 0$) and seeders leave the system with rate γ. This gives the following differential equation:

$$\dot{y} = -r\left.\frac{\partial F}{\partial \sigma}\right|_{\sigma=0} - \gamma y. \tag{6}$$

The P2P system is thus represented by the combined dynamics (5)-(6) in the state variables $F(t,\sigma)$, $y(t)$. Alternatively, one can replace $y(t)$ by the variable

$$z(t) = x(t) + y(t) = F(t,0) + y(t),$$

the total population of the network. By evaluating (5) at $\sigma = 0$ and adding it to (6), we deduce the following dynamic equation

$$\dot{z} = \lambda - \gamma z + \gamma x, \tag{7}$$

that can be used together with (5) to yield a complete model of the system.

[1]A more general $H(\sigma)$ may still be of interest to study peers arriving with partial content, or who are interested in only part of the content.

Equilibrium points

At an equilibrium point of (5)-(6), the number of seeders must be $y^* = \lambda/\gamma$. Also, since r^* does not depend on σ, the equilibrium of (5) has constant $\frac{\partial F^*}{\partial \sigma}$, i.e. we must have

$$F^*(\sigma) = x^*(1-\sigma), \tag{8}$$

a uniform distribution over download states. There are two cases, depending on which of the two constraints is active in (3). These depend on the parameter $\gamma_{cr} := \mu\frac{c}{c-\mu}$, as follows:

(a) If $\gamma < \gamma_{cr}$, the equilibrium satisfies

$$x^* = \frac{\lambda}{c} \text{ and } r^* = c,$$

the system operates *saturated by download capacity*.

(b) If $\gamma > \gamma_{cr}$, the equilibrium satisfies

$$x^* = \lambda\left(\frac{1}{\mu} - \frac{1}{\gamma}\right), \text{ and } r^* = \mu\frac{x^*+y^*}{x^*},$$

the system operates *saturated by upload capacity*.

Note that in case (a), locally around equilibrium the system dispenses a constant rate c to each downloading peer, as in Example II-1. This makes the local dynamics autonomous, and readily shown to be stable (see [5]). Case (b) is dynamically more challenging since the download rate varies (even locally) with the state, introducing feedback in the system. Hence we focus on (b) in our dynamic studies to follow[2].

IV. Local stability analysis

In this section we prove local stability for the processor sharing model around equilibrium, in the case (b), saturated by upload capacity. For simplicity, we will assume $\mu = 1$, which amounts to choosing time units. Note that in this case we have $\gamma > \gamma_{cr} > 1$. The resulting equilibrium conditions are

$$x^* = \lambda\frac{\gamma-1}{\gamma}, \quad z^* = \lambda, \quad r^* = \frac{z^*}{x^*} = \frac{\gamma}{\gamma-1}, \tag{9}$$

and $F^*(\sigma)$ from (8).

We first linearize the dynamics (5)-(7) around this equilibrium, writing the incremental variables $\tilde{x} = x - x^*$, $\tilde{z} = z - z^*$, $\tilde{r} = r - r^*$, $f(t,\sigma) = F(t,\sigma) - F^*(\sigma)$. Note that $f(t,1) \equiv 0$.

$$\frac{\partial f}{\partial t} = r^*\frac{\partial f}{\partial \sigma} - x^*\tilde{r} \tag{10a}$$

$$\dot{\tilde{z}} = -\gamma\tilde{z} + \gamma\tilde{x}. \tag{10b}$$

The aforementioned feedback mechanism results from the download rate $r = \frac{z}{x}$, which gives by linearization the expression in incremental variables

$$x^*\tilde{r} = \tilde{z} - r^*\tilde{x}.$$

Substitution in (10a) yields the linearized model

$$\frac{\partial f}{\partial t} = \frac{1}{\tau}\frac{\partial f}{\partial \sigma} - \tilde{z} + \frac{1}{\tau}\tilde{x},$$

[2](b) is also the most interesting from a P2P perspective; here the leecher contribution is essential to sustain the load.

where for future convenience we have introduced the notation

$$\tau := \frac{\gamma - 1}{\gamma} = \frac{1}{r^*},$$

representing the equilibrium download time for each leecher. Now define a new variable

$$u = -\tau \tilde{z} + \tilde{x};$$

we can recast the local linearized dynamics as the feedback interconnection of two blocks:

- G_1, with input u and output x is the infinite-dimensional system

$$\begin{aligned} \frac{\partial f}{\partial t} &= \frac{1}{\tau}\frac{\partial f}{\partial \sigma} + \frac{1}{\tau} u, && (11a) \\ \tilde{x} &= f(t, 0). && (11b) \end{aligned}$$

- G_2, with input $\tilde{x}$ and output u is the first-order system

$$\begin{aligned} \dot{\tilde{z}} &= -\gamma \tilde{z} + \gamma \tilde{x}, && (12a) \\ u &= -\tau \tilde{z} + \tilde{x}. && (12b) \end{aligned}$$

We study this feedback loop through the transfer functions $\hat{G}_1(s)$, $\hat{G}_2(s)$, which are elements of the $\mathcal{H}_\infty$ space of bounded analytic functions on $Re(s) > 0$, see e.g. [3] for background.

Theorem 1: The feedback interconnection of (11)-(12) has the small-gain property $\|\hat{G}_1 \hat{G}_2\|_\infty < 1$, and therefore the closed loop satisfies $[1 - \hat{G}_1(s)\hat{G}_2(s)]^{-1} \in \mathcal{H}_\infty$.

Proof: The first-order system G_2 has transfer function

$$\hat{G}_2(s) = 1 - \frac{\tau\gamma}{s + \gamma} = \frac{s+1}{s+\gamma}.$$

Since $\gamma > 1$, the above lead-lag system has $\mathcal{H}_\infty$ norm

$$\|\hat{G}_2\|_\infty = \sup_{\omega \in \mathbb{R}} |\hat{G}_2(j\omega)| = 1, \quad \text{achieved as } \omega \to \infty.$$

We find the transfer function of G_1. Let $\hat{f}(s, \sigma)$ be the Laplace transform in the time variable of $f(t, \sigma)$. For zero initial conditions, (11a) yields

$$s\hat{f}(s, \sigma) = \frac{1}{\tau}\frac{\partial \hat{f}}{\partial \sigma} + \frac{1}{\tau}\hat{u}(s);$$

this is now an ordinary differential equation in σ, with constant coefficients. Noting that $\hat{f}(s, 1) = 0$, we have the solution

$$\hat{f}(s, \sigma) = \frac{1}{\tau s}(1 - e^{\tau s(\sigma - 1)})\hat{u}(s).$$

Evaluating at $\sigma = 0$ gives $\hat{\tilde{x}}(s)$, the Laplace transform of the output. Therefore we obtain the transfer function

$$\hat{G}_1(s) = \frac{1 - e^{-\tau s}}{\tau s}.$$

It is easily checked that

$$\|\hat{G}_1(s)\|_\infty = \sup_{\omega \in \mathbb{R}} |\hat{G}_1(j\omega)| = 1,$$

achieved at $\omega = 0$. It follows that the feedback loop has gain

$$\|\hat{G}_1(s)\hat{G}_2(s)\|_\infty = \sup_{\omega \in \mathbb{R}} |\hat{G}_1(j\omega)||\hat{G}_2(j\omega)| < 1,$$

where we use the fact that the two terms achieve their maximum at opposite ends of the spectrum. The small-gain theorem (see [3]) implies that $[1 - \hat{G}_1(s)\hat{G}_2(s)]^{-1} \in \mathcal{H}_\infty$. ■

The above result implies local stability from the input-output point of view used in control theory: in particular, injected disturbances in the feedback will have a bounded impact. If the blocks were finite dimensional, this would be equivalent to asymptotic stability of the autonomous dynamics; here, however, we have an infinite dimensional block $\hat{G}_1(s)$, and the relationship between the various stability notions is not immediate (see [3]): we leave this analysis for future work.

V. Variability analysis

In the fluid model analyzed so far, peer arrivals, departures and internal transitions are assumed to follow a perfectly deterministic pattern. For a more refined analysis we would like to analyze variability around the equilibrium values caused by random variations in these patterns.

In the fluid context, one way to account for this randomness is to add noise to the respective differential equations. This approach was outlined in [12] for the two-state (x, y) differential equation models used there, by adding Brownian noise of variance λ to arrivals, departures, and also for the transition between leechers and seeders. We believe, however, that it is questionable that an independent noise should be added for the latter, given that progress in the system is determined endogenously by the system state, which sets download rates. For this reason, in what follows we will include only noise in arrival and departure terms, keeping the transport portion of the system deterministic. Returning momentarily to non-incremental variables, we write

$$\begin{aligned} \frac{\partial F}{\partial t} &= \lambda + n_1(t) + r(F, y, \sigma)\frac{\partial F}{\partial \sigma}, \\ \dot{y} &= -\, r(F, y, 0)\frac{\partial F}{\partial \sigma}\bigg|_{\sigma=0} - \gamma y + n_2(t). \end{aligned}$$

Above, $n_1(t)$ models deviations from the mean arrival rate λ, and $n_2(t)$ deviations from the seeder departure rate. We characterize their influence on the local dynamics around equilibrium, by evaluating the transfer function from these noise terms to the relevant system variables.

Including the above noise in the linearization around equilibrium, (10) becomes

$$\begin{aligned} \frac{\partial f}{\partial t} &= r^* \frac{\partial f}{\partial \sigma} - x^* \tilde{r} + n_1(t), && (13a) \\ \dot{\tilde{z}} &= -\gamma \tilde{z} + \gamma \tilde{x} + n_1(t) + n_2(t), && (13b) \end{aligned}$$

back in incremental variables. Note that the state $z = x + y$ is influenced by both noise sources.

Replicating the analysis of the previous section, with $\mu = 1$ and the equilibrium values from (9), we obtain from (13a) the equation

$$\frac{\partial f}{\partial t} = \frac{1}{\tau}\frac{\partial f}{\partial \sigma} - \tilde{z} + \frac{1}{\tau}\tilde{x} + n_1(t),$$

where as before $\tau := \frac{\gamma-1}{\gamma}$. Defining now

$$u := -\tau\tilde{z} + \tilde{x} + \tau n_1,$$

we can represent the dynamics as the feedback interconnection of a noise-less system G_1 as in (11), in feedback with

$$\begin{aligned} \dot{\tilde{z}} &= -\gamma\tilde{z} + \gamma\tilde{x} + n_1 + n_2 \\ u &= -\tau\tilde{z} + \tilde{x} + \tau n_1. \end{aligned}$$

In transfer function form, this second system is readily found to be

$$\begin{aligned} \hat{u}(s) &= \frac{s+1}{s+\gamma}\hat{\tilde{x}}(s) + \frac{\tau(s+\gamma-1)}{s+\gamma}\hat{n}_1(s) - \frac{\tau}{s+\gamma}\hat{n}_2(s) \\ &= \hat{G}_2(s)\hat{\tilde{x}}(s) + \hat{H}_1(s)\hat{n}_1(s) + \hat{H}_2(s)\hat{n}_2(s), \end{aligned}$$

where $\hat{G}_2(s)$ coincides with our previous noise-free analysis, for $\hat{H}_1(s)$, $\hat{H}_2(s)$ appropriately defined. The overall feedback diagram is depicted in Fig. 1.

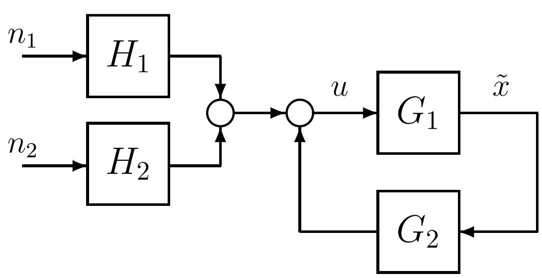

Fig. 1. Linearized dynamics with noise.

The stability of this feedback loop was already established in Theorem 1. To evaluate the variance of the output x, we compute the closed loop relationship

$$\hat{x}(s) = \frac{\hat{G}_1(s)}{1-\hat{G}_1(s)\hat{G}_2(s)}[\hat{H}_1(s)\hat{n}_1(s) + \hat{H}_2(s)\hat{n}_2(s)].$$

For independent white noise inputs of variance λ, we obtain the following expression for the variance of x:

$$E[x^2] = \lambda\int_{-\infty}^{\infty} \frac{|\hat{G}_1(j\omega)|^2\left[|\hat{H}_1(j\omega)^2| + |\hat{H}_2(j\omega)^2|\right]}{|1-\hat{G}_1(j\omega)\hat{G}_2(j\omega)|^2}\frac{d\omega}{2\pi}. \tag{14}$$

The above integral can be evaluated numerically. Figure 2 represents the power spectral density (integrand in (14)) for the case $\gamma = 3$.

For comparison purposes, we include in the Figure the corresponding spectral density that results from using the ODE model (4), in linearized form (for $\mu = 1$):

$$\dot{\tilde{x}} = -(\tilde{x}+\tilde{y}) + n_1, \tag{15a}$$
$$\dot{\tilde{y}} = (\tilde{x}+\tilde{y}) - \gamma\tilde{y} + n_2. \tag{15b}$$

The resulting transfer function expression is

$$\hat{\tilde{x}}(s) = \frac{(s+\gamma-1)\hat{n}_1(s) - \hat{n}_2(s)}{s^2+\gamma s+\gamma}.$$

We see that both models coincide at low frequency, and the PDE model predictably adds many high order modes.

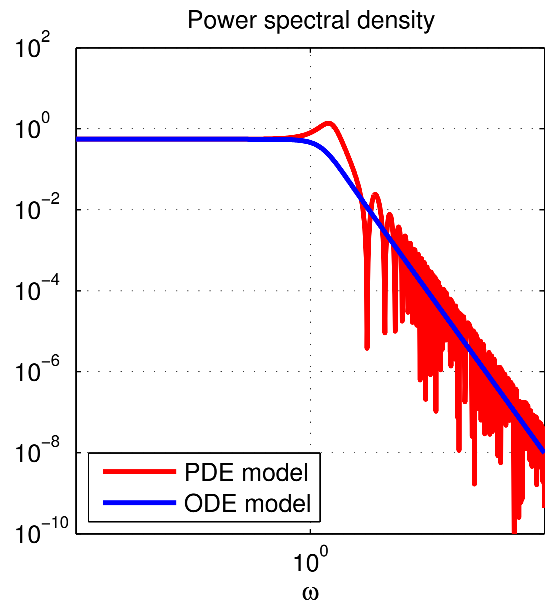

Fig. 2. Bode plot of the power spectral density of x; comparison of PDE and ODE models.

More significantly, however, around the cutoff frequency both models differ substantially. In simulation studies we have found the predictions of variance resulting from the PDE model to be more accurate.

VI. Transient analysis

We consider here a different P2P scenario, where a given number of initial seeders y_0 would like to propagate some content, and are willing to stay permanently in the system. Moreover, an initial distribution of leechers is given, and these leechers download the content and leave immediately after they finish. This is a typical situation in torrents nowadays, with the main relevant performance metric being the completion time, i.e. the time needed to finish to serve all the initial leechers.

Assuming that the bandwidth allocation is processor-sharing, and that the system is not bottlenecked by downlink capacity ($c \to \infty$) the dynamics are given by:

$$\frac{\partial F}{\partial t} = \mu\frac{x+y_0}{x}\frac{\partial F}{\partial \sigma}. \tag{16}$$

Note that we have returned to non-incremental variables. As for the initial condition, we assume that $F(0,\sigma) = \phi(\sigma)$, a strictly decreasing differentiable function of σ, with $x_0 := \phi(0)$, the initial total number of leechers, and $\phi(1) = 0$.

Proposition 2: The time needed to empty a processor-sharing P2P system with y_0 servers and starting from an initial condition $\phi(\sigma)$ is given by:

$$T = \frac{1}{\mu}\int_0^1 \frac{\phi(\sigma)}{\phi(\sigma)+y_0}d\sigma. \tag{17}$$

Proof: For this simplified system with no input, the PDE model (16) with initial condition $\phi(\sigma)$ can be written in the integral form

$$F(t,\sigma) = \phi\left(\sigma + \int_0^t \mu\left(1 + \frac{y_0}{x(\tau)}\right)d\tau\right).$$

We note that the expression remains valid while $x(t) > 0$, and that to interpret it we take $\phi(\sigma) = 0$ for $\sigma \geq 1$. That this is a solution can be readily verified by substituting in (16).

Evaluating the preceding equation in $\sigma = 0$ gives us the following integral equation for the number of leechers $x(t)$:

$$x(t) = F(t,0) = \phi\left(\int_0^t \mu\left(1 + \frac{y_0}{x(\tau)}\right)d\tau\right).$$

This is valid while $x(t) > 0$, and since ϕ is strictly decreasing in this range we can solve for

$$\phi^{-1}(x(t)) = \int_0^t \mu\left(1 + \frac{y_0}{x(\tau)}\right)d\tau.$$

Differentiating in t we get the following autonomous differential equation for x:

$$(\phi^{-1})'(x)\dot{x} = \mu\left(1 + \frac{y_0}{x}\right),$$
$$x(0) = x_0.$$

Applying separation of variables and integrating in $[0, T]$ gives

$$\frac{1}{\mu}\int_{x_0}^{x(T)} \frac{x}{x + y_0}(\phi^{-1})'(x)dx = T.$$

When $x(T)$ tends to zero we obtain the expression for the completion time:

$$T = \frac{1}{\mu}\int_{x_0}^{0} \frac{x}{x + y_0}(\phi^{-1})'(x)dx.$$

Finally, the change of variables $\sigma = \phi^{-1}(x)$ in the above integral leads to (17). ■

Furthermore, noting that the function $\frac{\xi}{\xi + y_0}$ is increasing in $\xi > 0$, and that $\phi(\sigma) \leqslant x_0\ \forall\sigma$, we have the following:

Corollary 3: The time T obtained above satisfies:

$$T \leqslant \frac{1}{\mu}\frac{x_0}{x_0 + y_0}. \tag{18}$$

In fact, the equality in the above expression is achieved when the initial condition ϕ approaches the function $x_0\mathbf{1}_{[0,1)}(\sigma)$, i.e. when all the leechers start empty.

Note that in particular, T is bounded above by $1/\mu$, i.e. the time for completion is finite, and is at most $1/\mu$, the time to upload a copy of the file. This uniform bound holds regardless of the initial number of leechers! This emphasizes the scalability of P2P file exchange mechanisms: when the demand is large, so is the available supply.

Again, let us compare the previous results with the predictions of previous models. In particular, the ODE model of (4) for this situation is

$$\dot{x} = -\mu\left(x + y_0\right),$$
$$x(0) = x_0.$$

With analogous (simpler) calculations, the completion time for this model can be readily calculated as:

$$T' = \int_0^{x_0} \frac{1}{\mu}\frac{1}{x + y_0}dx = \frac{1}{\mu}\log\left(1 + \frac{x_0}{y_0}\right).$$

Note in particular that the ODE model predicts $T' \to \infty$ as $x_0 \to \infty$, albeit logarithmically. Again, our simulation evidence suggests that this is pessimistic, the bounded time we found in (18) gives closer predictions.

VII. Conclusion

PDE models offer an attractive method to describe the dynamics of population and residual work in networks for which the job distribution is not exponential. We have investigated its use for a P2P setting, in which a deterministic job size is a natural model. For this case, we have shown how control-theoretic tools can be used to analyze local stability, random variability, and transient completion times.

One natural direction of future research is to consider other distribution functions $H(\sigma)$; these may arise when peers arrive with partial content or have only interest in some files among a larger set. Another direction is the extension to discriminatory file-sharing policies $r(F, y, \sigma)$ considered in [5], which may offer advantages as suggested by empirical studies [9]. In such cases the equilibrium will depart from the uniform distribution, complicating the local analysis pursued here.

References

[1] T. Bonald and L. Massoulié. "Impact of fairness on Internet performance." *Proc. ACM Sigmetrics*, 2001.

[2] B. Cohen, "Incentives build robustness in BitTorrent," in *1st Workshop on the Economics of Peer-2-Peer Systems, Berkeley*, 2003.

[3] R.F. Curtain and H.J. Zwart, *An Introduction to Infinite-Dimensional Linear Systems Theory*, Springer-Verlag, New York 1995.

[4] G. De Veciana, T. Lee and G. Konstantopoulos. "Stability and performance analysis of networks supporting elastic services." *IEEE/ACM Transactions on Networking*, 9(1):2–14, February 2001.

[5] A. Ferragut, F. Kozynski, and F. Paganini, "Dynamics of content propagation in BitTorrent-like P2P file exchange systems," in *Proc. 50th Conference on Decision and Control, Orlando, FL*, 2011.

[6] H. Gromoll, A. Puha, and R. Williams, "The fluid limit of a heavily loaded processor sharing queue," *Annals of Applied Probability*, vol. 12, pp. 797–859, 2002.

[7] H.C. Gromoll and R. Williams, "Fluid limits for networks with bandwidth sharing and general document size distributions," *Annals of Applied Probability*, vol. 19, pp. 243–280, 2009.

[8] F. P. Kelly and R. J. Williams, "Fluid model for a network operating under a fair bandwidth-sharing policy." *Annals of Applied Probability*, 14:1055–1083, 2004.

[9] F. Kozynski, A. Ferragut, and F. Paganini, "Reduction of oscilations in BitTorrent by preferential unchoking mechanisms", *Informatica na Educacao: Teoría & Prática*, v.14, n.1, 2011.

[10] J. Mo and J. Walrand "Fair end-to-end window based congestion control." *IEEE/ACM Trans. Networking*, 8(5):556–567, Oct. 2000.

[11] F. Paganini, A. Tang, A. Ferragut, L. Andrew. "Network Stability under Alpha Fair Bandwidth Allocation with General File Size Distribution". *IEEE Trans. on Automatic Control*, to appear March 2012.

[12] D. Qiu and R. Srikant, "Modeling and performance analysis of BitTorrent-like peer-to-peer networks," *ACM SIGCOMM Computer Communication Review*, vol. 34, no. 4, pp. 367–378, 2004.

[13] D. Qiu and W. Sang, "Global stability of peer-to-peer file sharing systems," *Computer Communications*, vol. 31, no. 2, pp. 212–219, 2008.

[14] J. Roberts, L. Massoulié. "Bandwidth Sharing and Admission Control for Elastic Traffic," *Proc. ITC Specialist Seminar*, Yokohama, Oct. 1998.

[15] R. Srikant, *The Mathematics of Internet Congestion Control,* Birkhauser, 2004.

[16] X. Yang and G. De Veciana, "Service capacity of peer-to-peer networks," in *Proceedings of IEEE/Infocom*, 2004.