Stochastic Systems
2014, Vol. 4, No. 0, 1–30
DOI: 10.1214/11-SSY53

BANDWIDTH SHARING NETWORKS WITH MULTISCALE TRAFFIC

By Mathieu Feuillet, Matthieu Jonckheere and Balakrishna Prabhu

INRIA, CONICET, and LAAS-CNRS

In multi-class communication networks, traffic surges due to one class of users can significantly degrade the performance for other classes. During these transient periods, it is thus of crucial importance to implement priority mechanisms that conserve the quality of service experienced by the affected classes, while ensuring that the temporarily unstable class is not entirely neglected. In this paper, we examine the complex interaction occurring between several classes of traffic when classes obtain bandwidth proportionally to their incoming traffic. We characterize the evolution of the performance measures of the network from the moment the initial surge takes place until the system reaches its equilibrium. Using a time-space-transition-scaling, we show that the trajectories of the temporarily unstable class can be described by a differential equation, while those of the stable classes retain their stochastic nature. In particular, we show that the temporarily unstable class evolves at a time-scale which is much slower than that of the stable classes. Although the time-scales decouple, the dynamics of the temporarily unstable and the stable classes continue to influence one another. We further proceed to characterize the obtained differential equations for several simple network examples. In particular, the macroscopic asymptotic behavior of the unstable class allows us to gain important qualitative insights on how the bandwidth allocation affects performance. We illustrate these results on several toy examples and we finally build a penalization rule using these results for a network integrating streaming and surging elastic traffic.

1. Introduction.

Communication networks are dealing with very heterogeneous sources of traffic having drastically different behaviors in terms of volume of data and agressivity. Ideally, the network should respond to the different demands in the fairest possible way, i.e. by avoiding a significant degradation of quality of service of a given class of traffic when another class undergoes a major traffic surge.

The impact of large-scale traffic surges, also known as slash-dot-crowds or flash-crowds, on web servers and content distribution networks has been

Received December 2011.
AMS 2000 subject classifications: Primary 90B18; secondary 60K35; tertiary 90B36.
Keywords and phrases: Scaling methods, bnadwidth sharing networks, stochastic averaging.

the subject of several studies (Stavrou, Rubenstein and Sahu, 2004; Kandula et al., 2005; Deshpande et al., 2007). These mainly focus on designing mechanisms to make the content providers resilient to surges of a given type of traffic. However, in addition to overloading the content providers, a traffic surge can also negatively impact the performance of other concurrent flows in the network. The temporarily unstable class can potentially starve the other classes from network capacity thereby subjecting them to unreasonable delays and packet losses. In such circumstances, in addition to protection mechanisms in web servers, it is crucial to implement bandwidth-sharing mechanisms inside the network that would protect the stable classes from the adversarial effects of the surge. It seems natural that such mechanisms should penalize the temporarily unstable class more when the level of congestion it creates is larger, without actually throttling it. (Thus, the more significant the surge is, the smaller the bandwidth each flow in this class gets.) To the best of our knowledge, the consequences of traffic surges on the performance of the different classes in the presence of such bandwidth sharing mechanisms have not been explored.

In this paper, we take a global view at the effects of a traffic surge in a multi-class communication network: our aim is to present an analytic treatment of the complex interaction that takes place between the temporarily unstable class and the stable class during a traffic surge when the temporarily unstable class is penalized proportionally to its level of congestion.

Towards this end, we consider stochastic networks describing the evolution of the number of flows (or calls) in a communication network where different classes of traffic compete for the bandwidth. Bandwidth-sharing network models (Massoulié and Roberts, 2002; Bonald and Proutière, 2003; Gromoll and Williams, 2009) have become quite a standard modeling tool over the past decade for modeling communication networks. In particular, they have been used extensively to represent the flow level dynamics of data traffic in wired or wireless networks (Bonald et al., 2006), as well as for the integration of voice and data traffic (Bonald and Proutière, 2004), hence generalizing more traditional voice traffic models, e.g. Kelly (1979).

To obtain structural results, we introduce a scaling when possibly only a subset of classes have initial conditions converging to infinity, thus generalizing the classical notion of fluid limit. Those classes shall be the ones undergoing a traffic surge (surging classes or temporarily unstable classes). We consider a situation where the allocation of bandwidth shall be meanwhile weighted such that the other classes of traffic (stable classes) are not led to starvation, i.e. the priority weight is very small compared to the offered traffic. Accelerating time together with re-scaling the state of the surging classes allow to "zoom out" the process, just as for usual fluid limits and obtain a bird's-eye view of the large scale dynamics for these classes. For example, such a situation might be the consequence of an inappropriately small level of priority. Alternatively, if the surge is externally caused, possibly due to a network attack, the network may be reacting to it by penalizing surging classes according to its level of congestion, so that other classes do not starve, see for instance (Peng, Lecki and Ramamohanarao, 2007) for practical considerations on the matter.

In order to obtain a classical fluid limit for Jackson networks (Robert, 2003) or for more complex bandwidth-sharing networks (Gromoll and Williams, 2009), all the classes are jointly scaled in time and in space. This yields a set of differential equations that govern the dynamics of all the classes. Under additional assumptions on the drift δ of the considered Markov process, the differential equation is simply of the form $\dot{x}(t) = \delta(x(t))$ (see the considerable amount of work on fluid limits and ODE methods both for Markov processes and for communications networks (Robert, 2003; Dai, 1995; Darling and Norris, 2008; Gromoll and Williams, 2009; Meyn, 2008)).

In our case, the situation differs as the transitions of surging classes are also scaled to model that the priority weight of surging classes is inversely proportional to the level of congestion. This has far-reaching consequences for the structure of the limiting process. Under this scaling, we will show that the dynamics of the temporarily unstable classes can be described by a deterministic differential equation, while the stable classes retain their stochastic nature. Hence, a time-scale separation occurs: the temporarily unstable classes evolve on a much slower time-scale compared to the stable classes. However even with this separation of time-scales, a strong coupling in the dynamics of the temporarily unstable and the stable classes remains. The dynamics of the temporarily unstable class is influenced by the stable classes through their conditional stationary distribution (conditional on the level of congestion of surging classes flows being fixed to its present macroscopic value), which in turn depends on the temporarily unstable classes. Hence, for surging classes the differential equation obtained is of the form $\dot{x}(t) = \bar{\delta}^t(x(t))$, where $\bar{\delta}^t$ is an average of the first coordinate drift according to the conditional distribution of the other classes, given the state of the surging classes. This phenomenon is usually known in the probability literature as averaging principle and has been studied by many authors in different contexts. We follow in particular the methodology introduced in the seminal paper of Kurtz (1992). In the analysis of the fluid limit of bandwidth sharing networks a time-scale separation between classes usually occurs when one class of traffic reaches equilibrium faster than the others, and hence when the fluid limit hits an hyperplane of the state state space. Simple examples of this phenomenon can be found in Robert (2003). A more complex example can be found in Feuillet (2012). The interesting feature in our scaling is the appearance of the *stochastic averaging* principle in the whole state space. Similar *stochastic* averaging phenomena have also been studied in statistical physics (C. Kipnis, 1991) as well as chemistry and biochemistry (Segel and Slemrod, 1989) where the kinetics of chemical reactions can be described by systems of ordinary differential equations. Usually these works assume that one of the dependent variables is in steady state with respect to the instantaneous values of the other dependent variables. Taking this time-scale separation as an assumption, an efficient approximation method called the quasi-steady-state is commonly used in that context. This is however in contrast with our situation where we show that the time-scale decoupling occurs as a consequence of the scaling of the parameters of the transitions of the stochastic processes considered.

Contribution. Our contribution first consists in establishing the convergence in L^1 uniformly on compact sets for stochastic processes commonly adopted in the modeling and analysis of communications networks under the scaling considered. Since the slow part of the processes (surging classes) remains coupled (at a macroscopic scale) to the fast part (the remaining classes), such a proof is not standard and has to be decomposed in several steps. While preliminary results for monotone networks were presented in Jonckheere, Núñez-Queija and Prabhu (2010), a general proof for general bandwidth sharing networks was still missing.

Second, we characterize the responses (evolutions of queue length) of different networks to the surge of traffic. We introduce the notion of robust stability, which describes a situation when the network can absorb a surge of traffic by:

- keeping the usual (non-surging) classes stable,
- reducing the macroscopic state of surging classes to 0.

We call the set of traffic parameters that lead to these two conditions, the robust stability region. We characterize the robust stability region for work conserving allocations and for monotone allocations. We first show that for work conserving allocations, the unstable classes, at their macroscopic time scale, see the other classes as having full priority, while the effect of the surging classes on the other classes gradually vanishes (again at a macroscopic time scale). Hence surging classes tend macroscopically to 0 under the (usual) stability condition of the system $\sum_{j=1}^{N} \rho_j < 1$. The situation is more complex for non work-conserving networks, where the behavior of the surging classes depends in an intricate manner upon that of the other classes. In particular, under the usual stability conditions of the network, the macroscopic state of surging classes might converge to 0 or to a strictly positive number, depending on the conditional distribution of the other classes. We prove for monotone networks that only when the allocation giving full priority to the stable classes, that is the allocation in a network where surging classes have 0 arrival rate, is stable, then surging classes converge to 0 on the macroscopic time scale. This hence demonstrates that the robust stability region might be strictly included in the (usual) stability region. We illustrate these concepts on several simple network topologies.

Finally, we use our analytical results to build an implementable penalization rule allowing to adapt the level of priority of streaming traffic in a network integrating streaming and elastic traffic, such as to target a given loss probability threshold/quality of service.

The rest of the paper is organized as follows. The model and notations are described in the next section. In Section 3, we present the convergence theorem for the considered scaling. In Section 4, we analyze the qualitative behavior of networks after a traffic surge in different cases and give numerical examples of applications of the main result to bandwidth sharing on some simple network topologies. In Section 5, we construct a practical penalization rule for elastic traffic when it shares resources with streaming traffic. Finally, we conclude in Section 6.

2. Model.

Notation. In the sequel, for $x \in \mathbb{Z}^N$, $|\cdot|$ denotes the l_1-norm:

$$|x| = \sum_{i=1}^{N} |x_i|.$$

For $x, y \in \mathbb{Z}^N$, we also use the notation $x \leq y$ to denote the partial order $x_i \leq y_i$ for all $i = 1, \dots, N$, and $x \cdot y$ denotes the usual scalar product.

2.1. *Networks with traffic-weighted allocations.* We consider a bandwidth-sharing network with N traffic classes. Within each of the N traffic classes, resources are shared according to a processor-sharing service discipline. The service rates are state-dependent: they may depend on the number of flows within the same class, as well as on the numbers of flows in all other classes. The service rates of the N traffic classes will be denoted by $\tilde{\phi} = (\tilde{\phi}_i(\cdot))_{i=1}^N$. Several examples are considered in the next section. Note that the service rate function $\tilde{\phi}$ captures the allocation of bandwidth which is determined by the specific network topology and congestion control mechanisms. Special allocation functions that have received much attention in literature include the celebrated max-min fair allocation and the proportional fair allocation.

We assume that class-i customers arrive subject to a Poisson process of intensity λ_i and require exponentially distributed[1] service times of mean μ_i^{-1} for class-i. The arrival processes of all classes are mutually independent. Our main results allow for time-varying arrival rates for the class exhibiting a traffic surge. When applicable, we reflect this dependence in the notation by adding the time parameter to the arrival rates and then $\lambda_1(t)$ is the arrival rate of class 1 at time t. For ease of exposition, however, we restrict ourselves to constant arrival rates for all classes in this section and will formulate our results with time-varying arrival rates in Section 3.

Let X be the stochastic process describing the number of flows (or calls) in progress. In the absence of priority mechanisms, and under the assumptions of Poisson arrivals and exponential flow sizes, X is a multi-dimensional birth and death process with transition rates:

$$\begin{aligned} q(x, x - e_i) &= \mu_i \tilde{\phi}_i(x), \\ q(x, x + e_i) &= \lambda_i, \end{aligned}$$

with $x \in {\mathbb{Z}_+}^N$. Assume now that priority mechanisms are employed in the network such that the actual bandwidth allocation depends on the variables $r_i x_i,\ i = 1, \ldots, N$, rather than simply on $x_i,\ i = 1, \ldots, N$. Hence, if x_i is thought of as a measure of the level of congestion of class i, a differentiation between classes can be enforced by giving different weights to the different classes. (Such a differentiation can be enforced at lower time-scales by packet schedulers like weighted deficit round robin.)

It can also be the case that each class of traffic has a limited peak rate (because of access constraints for instance). It could then be advantageous for providers, in order to meet the demand, to share capacity as a function of the demanded rates $r_i x_i$ rather than as a function of the number of flows of each class in the network. In both configurations, X can now be described as multi-dimensional birth and death process with transition rates:

$$\begin{aligned} q(x, x - e_i) &= \mu_i \phi_i(r.x), \\ q(x, x + e_i) &= \lambda_i, \end{aligned}$$

where $r \cdot x = (r_i x_i)_{i=1,\ldots,N}$ for some $r \in \mathbb{R}_+^N$. To avoid confusion, we emphasize once more that reflecting the dependence on the control parameters

[1]Such assumptions are certainly not necessary to obtain the results we are aiming at; however, a rigorous generalization would be technically very involved and is beyond the scope of the present paper.

r_i in our notation will be more convenient for the purposes in this paper, rather than making this dependence implicit, i.e. through the allocation $\tilde{\phi}(x) = \phi(r \cdot x)$. The load of class i is given by

$$\rho_i = \frac{\lambda_i}{\mu_i}.$$

We shall further suppose that the weights r are chosen for each class proportionally to the size of the traffic surge, i.e.

$$r_i = r_i(|x|) = \frac{\omega_i}{|x|}.$$

3. Fluid limits with time scales decoupling. We model a traffic surge by a large number of initial flows and large arrival rates for a subset of classes being (temporarily at least) unstable. Let $S \leq N \in \mathbb{Z}_+$ denote the number of classes that undergo a surge. To get structural results on the process X, we study the case where:

1. the number of initial class-i, $i = 1, \ldots, S$, flows is of order $K = |x|$.
2. we scale (accelerate) time by a factor K,
3. we scale class i, $i = 1, \ldots, S$, states by a factor $1/K$,
4. the prioritization weight r_i of class i, $i \leq S$, is of order $1/K$.

We now consider a network with several classes of traffic and with class i, $i = 1, \ldots, S$, going through a temporary surge of traffic. Recall that we focus on a regime where $r_i \equiv \frac{\omega_i}{K}$ and $K \to \infty$. We further let Y^K denote the (scaled) process:

$$Y^K(t) = \left(\left(\frac{X_i^K(Kt)}{K} \right)_{i=1,\ldots,S}, \left(X_i^K(Kt) \right)_{i=S+1,\ldots,N} \right). \tag{1}$$

In the following we show that, as $K \to \infty$, Y^K converges to a stochastic process whose first S coordinates are deterministic, which are a solution of a differential equation that can be described in terms of an averaged rate $\bar{\phi}$. In the limit, the result implies a time-scale separation between the surging classes and the other ones.

Define U^z to be a $N-S$ dimensional Markov birth-and-death process with arrival rates λ_i and death rates $\phi_i(z, \cdot)$, $i = S+1, \ldots, N$, where $z \in \mathbb{R}_+^S$. Denote by $\pi^z(\cdot)$ its stationary probability (when it exists). When we do not use a time index, we implicitly suppose that we consider stationary versions of the processes.

To establish our main result, we shall make the following assumptions:

(A_1): $\phi_i(\cdot, x_{S+1}, \ldots, x_N)$ can be extended to a Lipschitz continuous functions from $\mathbb{R}_+^S \setminus \{0\}$ to $\mathbb{R}_+$.

(A_2): for all fixed z, the process U^z is ergodic. We can thus define $\mathbb{E}U^z$ the mean under the stationary distribution of the process U^z.

(A_3): $\frac{1}{K}\int_0^{Kt}\lambda_i(s)\,\mathrm{d}s \to a_i(t),\ i = 1,\dots,S$. It shall be assumed that $a_i(t)$ is differentiable for all t and i.

We note that although the model was formulated for a constant arrival rate for ease of exposition, our results will be proved for time-dependent arrival rates that satisfy assumption (A_3) stated above.

Let $u(t) \in \mathbb{R}^S$ be the solution (assuming it exists and it is unique) of the differential equation:

$$\forall i = 1,\dots,S, \quad \dot{u}_i(t) = \begin{cases} \dot{a}_i(t) - \bar{\phi}_i(u(t)), & \text{if } u_i(t) > 0, \\ 0, & \text{if } u_1(t) = 0, \end{cases} \tag{2}$$

with $\bar{\phi}_i(z) = \sum_{y\in\mathbb{Z}_+{}^{N-S}} \phi_i(z,y)\pi^z(y)$.

We can now proceed to state our main result:

Theorem 3.1. *Under the assumptions (A_1), (A_2) and (A_3), the process $Y_i^K(t)_{i=1,\dots,S}$ converges in L^1, uniformly on compact intervals, to the deterministic trajectory $u(t)$, i.e.*

$$\mathbb{E}\left[\sup_{0\le s\le t}\left|Y_i^K(s) - u_i(s)\right|\right] \to 0, \quad K\to\infty,\ \forall i = 1,\dots,S. \tag{3}$$

Moreover, for all times t, and for all bounded continuous functions f:

$$\begin{aligned} \lim_{K\to\infty} \mathbb{E}\Bigg(\sup_{0\le s\le t}\Bigg|\int_0^s f\left(Y^K(\theta)\right) \\ - \mathbb{E}^{U^{Z(\theta)}}\left(f\left(Z(\theta), U^{Z(\theta)}(\theta)\right) \,\Big|\, Z(\theta) = u(\theta)\right)\mathrm{d}\theta\Bigg|\Bigg) = 0. \end{aligned} \tag{4}$$

Remark 3.1. The processes U^z represent networks where classes $i = 1,\dots,S$, have been frozen. Our main theorem hence simply states that a decoupling of time scales occurs (in the limit) between surging and still classes: the latter see (at their time scale) the surging classes as frozen and hence get stationary (i.e. stable) if and only if the processes U^z are stable. On the contrary, the surging classes see only the stationary mean of the other classes which evolve infinitely faster.

The details of the proof, assuming $S = 1$ to simplify the exposition, are given in the next Section. We underline here the main steps:

- We first prove tightness of the laws of the scaled process, and show that the limit-points of Y_1^K are continuous processes.

- Supposing the convergence in distribution of the first class we characterize the limit of the functional $\int_0^t \mathbb{1}_{\{(X_2^K(Ks),\ldots,X_N^K(Ks))\in\Gamma\}}\,\mathrm{d}s,\ \forall\Gamma\subset\bar{\mathbb{Z}}_+^{N-1}$ and prove the limits are unique (and deterministic given the value of the first class). A key step is the useful characterization of bimeasures.
- Finally, we show that Y_1^K converges in distribution towards a deterministic process which allows to prove, using the previous step, the convergence in L^1, uniformly on compact sets.

3.1. *Proof of Theorem 3.1.*

Step 1. As mentioned previously, for the ease of exposition, we suppose that only class 1 undergoes a surge of traffic, i.e. $S = 1$. The proof then extends directly to the general case. We also remind the reader that the prioritization weight of class 1 is inversely proportional to K.

We thus consider the process $Y^K(t) = (\frac{X_1^K(Kt)}{K}, (X_i^K(Kt))_{i=2,\ldots,N})$ as defined by (1). We define $\bar{\mathbb{Z}}_+ = \mathbb{Z}_+ \cup \{+\infty\}$ topologized by taking as open sets all the open subsets U of $\bar{\mathbb{Z}}_+$ together with all subsets V which contain ∞ and such that its complementary is closed and compact, (topology of the Alexandroff compactification, also know as one-point compactification, see Engelking (1977)). For each K, we define the following random measure on $[0,\infty)\times\bar{\mathbb{Z}}_+^{N-1}$:

$$\nu^K((0,t)\times\Gamma) = \int_0^t \mathbb{1}_{\{(X_2^K(Ks),\ldots,X_N^K(Ks))\in\Gamma\}}\,\mathrm{d}s,\ \forall\Gamma\subset\bar{\mathbb{Z}}_+^{N-1},\ \text{and } \forall t\geq 0.$$

We denote $\mathcal{L}_0(\bar{\mathbb{Z}}_+^{N-1})$ the set of measures on $[0,\infty)\times\bar{\mathbb{Z}}_+^{N-1}$ such that, for all measures ν in $\mathcal{L}_0(\bar{\mathbb{Z}}_+^{N-1})$ and all $t\geq 0$, we have $\nu((0,t)\times\bar{\mathbb{Z}}_+^{N-1}) = t$. Since $\bar{\mathbb{Z}}_+$ is compact for the chosen topology, we have that $\mathcal{L}_0(\bar{\mathbb{Z}}_+^{N-1})$ is compact and we deduce that $\{\nu^K,\ K\in\mathbb{Z}_+\}$ is relatively compact.

In order to prove the relative compactness of $\{(Y_1^K,\nu^K),\ K\in\mathbb{Z}_+\}$, we then just have to prove the relative compactness of $\{Y_1^K,\ K\in\mathbb{Z}_+\}$. We define the following process

$$\begin{aligned} M_1^K(t) &= Y_1(t) \\ &\qquad - \frac{1}{K}\int_0^{Kt}\lambda_1(s)\,\mathrm{d}s + \frac{1}{K}\int_0^{Kt}\phi_1\left(\frac{X_1^K(s)}{K}, X_2^K(s),\ldots,X_N^K(s)\right)\mathrm{d}s \end{aligned} \tag{5}$$

The martingale characterization of jump processes (see Rogers and Williams (2000 (1987))) shows that M_1^K is a local martingale and its increasing process is given by

$$\langle M_1^K \rangle = \frac{1}{K^2}\int_0^{Kt} \lambda_1(s)\,\mathrm{d}s + \frac{1}{K^2}\int_0^{Kt} \phi_1\left(\frac{X_1^K(s)}{K}, X_2^K(s), \dots, X_N^K(s)\right)\mathrm{d}s$$

Using Doob's inequality[2], it follows that M_1^K converges in probability to 0 on any compact set when $K \to \infty$, i.e. for any $T \geq 0$ and any $\varepsilon > 0$,

$$\lim_{K\to\infty} \mathbb{P}\left(\sup_{0\leq s\leq T} |M_1^K(s)| > \varepsilon\right) = 0. \tag{6}$$

We then define w_h the modulus of continuity for any function h defined on $[0,t]$:

$$w_h(\delta) = \sup_{s_1,s_2\leq t;\ |s_1-s_2|<\delta} |h(s_1) - h(s_2)|.$$

Using Equations (5) and (6), since the process $Y_1^K(\cdot)$ decomposes into an absolutely continuous part and a martingale part converging to 0 we have that for any $\varepsilon > 0$ and $\eta > 0$, there exists $\delta > 0$ and A such that for $K > A$, we have

$$\mathbb{P}\left(w_{Y_1^K(.)}(\delta) > \eta\right) \leq \varepsilon.$$

The conditions of (Billingsley, 1999, 7.2 p. 81) are then fulfilled and the set $\{Y_1^K,\ K \in \mathbb{Z}_+\}$ is relatively compact. Moreover, any limiting point is a continuous process.

Step 2. We now suppose that (Y_1^K) converges in distribution to a limit Z_1. We have to characterize any limiting point of the sequence (ν^K) and then deduce the existence and uniqueness of the limit of (ν^K). In the following, we consider a convergent subsequence (Y^{K_l}, ν^{K_l}) and its limit process (Z_1, ν).

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be the probability space on which they are defined. We call $\{\mathcal{F}_t\}$ the natural filtration of (Z_1,ν). We then define γ such that

$$\forall A \in \mathcal{F},\ \forall B \in \mathcal{B}([0,\infty)), \forall C \in \mathcal{B}(\bar{\mathbb{Z}}_+^{N-1})\ \gamma(A\times B\times C) = \mathbb{E}(\mathbb{1}_A \nu(B\times C)).$$

According to (Ethier and Kurtz, 1986, Appendix 8), γ can be extended to a measure on $\mathcal{F}\otimes\mathcal{B}([0,\infty))\otimes\mathcal{B}(\bar{\mathbb{Z}}_+^{N-1})$. Furthermore, there exists ϑ such

[2]For any martingale M, using Cauchy Schwartz and Doob's inequality Darling and Norris (2008), we get that:

$$\begin{aligned}\mathbb{E}\left(\left|\sup_{0\leq s\leq t} M_s\right|\right)^2 &\leq \mathbb{E}\left(\sup_{0\leq s\leq t}|M_s|\right)^2\\ &\leq \mathbb{E}\left(\sup_{0\leq s\leq t} M_s^2\right),\\ &\leq 4\mathbb{E}\left(M_t^2\right).\end{aligned}$$

that for all t, $\vartheta(t,\cdot)$ is a random probability measure on $\bar{\mathbb{Z}}_+^{N-1}$, for any $B \in \mathcal{B}(\bar{\mathbb{Z}}_+^{N-1})$, $(\vartheta(t,B), t \geq 0)$ is $\{\mathcal{F}_t\}$-adapted and for any $A \in \mathcal{F} \otimes \mathcal{B}([0,\infty))$,

$$\gamma(A \times B) = \mathbb{E}\left(\int_0^{+\infty} \mathbb{1}_A(s)\vartheta(s,B)\,\mathrm{d}s\right). \tag{7}$$

Let us now define

$$M_B(t) = \nu([0,t] \times B) - \int_0^t \vartheta(s,B)\,\mathrm{d}s.$$

M_B is $\{\mathcal{F}_t\}$-adapted and continuous. We consider $t \geq s$, and $D \in \mathcal{F}_s$. We define $\mathbb{1}_C(\omega,\theta) = \mathbb{1}_D(\omega)\mathbb{1}_{[s,t)}(\theta)$ and we have

$$\begin{aligned}
\mathbb{E}\left(\mathbb{1}_D \nu([s,t) \times B)\right) &= \gamma(D \times [s,t) \times B), \\
&= \gamma(C \times B), \\
&= \mathbb{E}\left(\int_0^\infty \mathbb{1}_C(\theta)\vartheta(\theta,B)\,\mathrm{d}\theta\right), \text{ (according to (7))} \\
&= \mathbb{E}\left(\mathbb{1}_D \int_s^t \vartheta(\theta,B)\,\mathrm{d}\theta\right).
\end{aligned}$$

Since the previous equality is true for all $D \in \mathcal{F}_t$, it follows that

$$\mathbb{E}\left(\nu([s,t) \times B) \mid \mathcal{F}_s\right) = \mathbb{E}\left(\int_s^t \vartheta(\theta,B)\,\mathrm{d}\theta \,\middle|\, \mathcal{F}_s\right).$$

and immediately, we have

$$\mathbb{E}\left(M_B(t) \mid \mathcal{F}_s\right) = M_B(s).$$

Then, M_B is a continuous $\{\mathcal{F}_t\}$-martingale. It has finite sample paths from which we deduce that it is almost surely identically null. Then, the following equation holds for all t, almost surely,

$$\forall B \subset \bar{\mathbb{Z}}_+^{N-1},\ \nu([0,t) \times B) = \int_0^t \vartheta(s,B)\,\mathrm{d}s. \tag{8}$$

We have to characterize the random measures $\vartheta(t,.)$ associated to ν. For any uniformly continuous bounded function g on $\bar{\mathbb{Z}}_+^{N-1}$ and any $K \in \mathbb{Z}_+$, we define,

$$\begin{aligned}
M_g^K(t) = \frac{1}{K}&\Big(g\big(X_2^K(Kt),\dots,X_N^K(Kt)\big) - g(0)\Big) \\
&- \sum_{i=2}^N \lambda_i \int_0^t \Big(g\left(X_2^K(Kt),\dots,X_i^K(Kt)+e_i,\dots,X_N^K(Kt)\right)
\end{aligned}$$

$$
\begin{aligned}
&- g\left(X_2^K(Kt),\dots,X_N^K(Kt)\right)\Big)\,\mathrm{d}s\\
-\sum_{i=2}^{N}\mu_i\int_0^t\Big(&g\left(X_2^K(Kt),\dots,X_i^K(Kt)-e_i,\dots,X_N^K(Kt)\right)\\
&-g\left(X_2^K(Kt),\dots,X_N^K(Kt)\right)\Big)\\
&\phi_i\left(Y_1^K(s),X_2^K(Kt),\dots,X_N^K(Kt)\right)\mathrm{d}s.
\end{aligned}
$$

As M_1^K is a martingale, M_g^K is a martingale. Using the same argument as for M_K^1, (i.e. controlling the quadratic variation), we obtain that $M_g^{K_l}$ converges in distribution to 0. $|K_l|^{-1}(g(X_2^{K_l}(K_lt),\dots,X_N^{K_l}(K_lt))-g(0))$ also converges to 0 because g is bounded. As a consequence, the following term

$$
\begin{aligned}
&\sum_{i=2}^{N}\lambda_i\int_0^t\Big(g\left(X_2^{K_l}(Kt),\dots,X_i^{K_l}(K_lt)+e_i,\dots,X_N^{K_l}(K_lt)\right)\\
&\quad- g\left(X_2^{K_l}(K_lt),\dots,X_N^{K_l}(K_lt)\right)\Big)\,\mathrm{d}s\\
&-\sum_{i=2}^{N}\mu_i\int_0^t\Big(g\left(X_2^{K_l}(K_lt),\dots,X_i^{K_l}(K_lt)-e_i,\dots,X_N^{K_l}(K_lt)\right)\\
&\quad- g\left(X_2^{K_l}(K_lt),\dots,X_N^{K_l}(K_lt)\right)\Big)\phi_i\left(Y_1^{K_l}(s),X_2^{K_l}(K_lt),\dots,X_N^{K_l}(K_lt)\right)\mathrm{d}s
\end{aligned}
\tag{9}
$$

also converges in distribution to 0. But, by the continuous mapping theorem, using the continuity of g and ϕ, the processes $g(X_2^{K_l}(Kt),\dots,X_i^{K_l}(K_lt)+e_i,\dots,X_N^{K_l}(K_lt))$ as well as $\phi_i(Y_1^{K_l}(s),X_2^{K_l}(K_lt),\dots,X_N^{K_l}(K_lt))$ converge also in distribution. Now using (8), the term in (9) converges in distribution to

$$
\begin{aligned}
\int_0^t\sum_{i=2}^{N}\Big(&\lambda_i\sum_{y\in\mathbb{Z}_+^{N-1}}g(y+e_i)-g(y)\\
&+\mu_i\sum_{y\in\mathbb{Z}_+^{N-1}}\left(g(y-e_i)-g(y)\right)\phi_i\left(Z_1(s),y\right)\Big)\vartheta(s,y)\,\mathrm{d}s.
\end{aligned}
$$

Consequently, this is null almost surely for all t and we have then, for Lebesgue-almost every t,

$$
\sum_{i=2}^{N}\Big(\lambda_i\sum_{y\in\mathbb{Z}_+^{N-1}}g(y+e_i)-g(y)
$$

$$+ \mu_i \sum_{y\in\mathbb{Z}_+^{N-1}} (g(y-e_i)-g(y))\phi_i(Z_1(t),y)\Bigg)\vartheta(t,y)=0.$$

We deduce immediately that

$$\int_{\bar{\mathbb{Z}}_+^{N-1}} \Omega^{Z_1(t)}(g)(y)\vartheta(t,\mathrm{d}y)=0$$

where $\Omega^{Z_1(t)}$ is the infinitesimal generator of $(U^{Z_1(t)}(s))$. This proves exactly that $\vartheta(t,\cdot)$ is invariant for $U^{Z_1(t)}$. By uniqueness of the invariant distribution of $(U^{Z_1(t)}(s))$, this implies that, given Z_1, $\vartheta(t,\cdot)$ is a deterministic measure for all t. We can deduce that, if (Y^{K_l}) is a converging subsequence, then (v^{K_l}) is also converging and its limit is a random measure in $\mathcal{L}_0(\mathbb{Z}_+^{N-1})$. This implies in particular that (v^{K_l}) is tight in $\mathcal{L}_0(\mathbb{Z}_+^{N-1})$. We can now proceed of the last part of this step.

We consider $\varepsilon>0$, $\eta>0$ and $t\geq 0$. Because the sequence (ν^{K_l}) is tight in $\mathcal{L}_0(\mathbb{Z}_+^{N-1})$, there exists $\kappa>0$ and a compact $\Gamma\subset\mathbb{Z}_+^{N-1}$ such that:

$$\mathbb{P}\left(\sup_{l\geq\kappa}\nu^{K_l}([0,t)\times\Gamma^c)\geq\varepsilon\right)\leq\eta/2.$$

Because Z_1 is almost surely continuous and f is Lipschitz-continuous, we have

$$\mathbb{P}\left(\sup_{l\geq\kappa,y\in\Gamma,s\geq t}\left|f(Y_1^{K_l}(t),y)-f(Z_1(t),y)\right|\geq\varepsilon\right)\leq\eta/2.$$

Since f is bounded, we can deduce:

$$\mathbb{P}\Bigg(\sup_{k\geq\kappa}\Bigg|\int_{[0,t]\times\bar{\mathbb{Z}}_+^{N-1}} f(Y_1^{K_l}(s),y)\nu^{K_l}(\mathrm{d}s\times\mathrm{d}y) - \int_{[0,t]\times\bar{\mathbb{Z}}_+^{N-1}} f(Z_1(s),y)\nu(\mathrm{d}s\times\mathrm{d}y)\Bigg|\geq 2\varepsilon\|f\|\Bigg)\leq\eta.$$

According to (8), there exists a family $(\vartheta(t,\cdot))$ of random measures on $\bar{\mathbb{Z}}_+^{N-1}$ such that

$$\sup_{0\leq s\leq t}\left|\int_0^s f(Y_1^{K_l}(\theta),X_i^{K_l}(K_l\theta)) - \sum_{y\in\bar{\mathbb{Z}}_+^{N-1}} f\left(Z_1(\theta),y\right)\vartheta(\theta,y)\,\mathrm{d}\theta\right|$$

converges in probability to 0 when K_l tends to infinity.

Since f is bounded, we can apply the dominated convergence theorem and we have that

$$\lim_{K_l\to\infty} \mathbb{E}\left(\sup_{0\le s\le t}\left|\int_0^s f\left(Y_1^{K_l}(\theta), X_i^{K_l}(K_l\theta)\right)\right.\right.$$
$$\left.\left.- \sum_{y\in\bar{\mathbb{Z}}_+^{N-1}} f\left(Z_1(\theta), y\right)\vartheta(\theta,y)\,\mathrm{d}\theta\right|\right) = 0.$$

We further have that

$$\lim_{K\to\infty} \mathbb{E}\left(\sup_{0\le s\le t}\left|\int_0^s f\left(Y_1^{K}(\theta), X_i^{K}(K\theta)\right)\right.\right.$$
$$\left.\left.-\mathbb{E}\left(f\left(Z_1(\theta), U_i^{Z_1(\theta)}(\theta)\right) \,\Big|\, Z_1(\theta)\right)\mathrm{d}\theta\right|\right) = 0.$$

Step 2 is now complete.

Step 3. Using the martingale decomposition of X_1^K,

$$\begin{aligned} Y_1^K(t) &= \frac{X_1^K(Kt)}{K} \\ &= x_1 + M_K(t) + \frac{1}{K}\int_0^{Kt} \lambda_1^K(s)\,\mathrm{d}s \\ &\quad - \frac{1}{K}\int_0^{Kt} \phi_1\left(\frac{X_1^K(s)}{K}, X_2^K(s), \dots, X_N^K(s)\right)\mathrm{d}s. \end{aligned}$$

As already remarked in Step 1, since ϕ is bounded it follows that

$$\mathbb{E}\left(M_K(t)^2\right) \le \frac{At}{K}$$

which implies using Doob's inequality that there exists a constant A' such that for K big enough:

$$\mathbb{E}\left(\left|\sup_{0\le s\le t} M_K(s)\right|\right) \le A'\sqrt{\frac{t}{K}} \le \varepsilon.$$

Using the convergence of the arrival process (see assumption (A_3)) together with the convergence of the martingale M_K, we obtain the uniform integrability of Y_1^K. (The tightness of Y_1^K has already been obtained in Step 1). Now consider a converging subsequence $Y_1^{K_l}$ towards Z_1. Using the results of Step 2, the convergence of the arrival process together and the convergence of the martingale M_K, we obtain that Z_1 must satisfy:

$$Z_1(t) = x_1 + 0 + a_1(t) - \int_0^t \bar{\phi}_1(Z_1(s))\,\mathrm{d}s.$$

Hence the limit is unique and deterministic.

This in turn shows the convergence of Y_1^K in distribution and completely characterizes the measure ϑ introduced in Step 2 as a deterministic measure. We can now prove the convergence in L^1. Let ε be given. Define the error estimate:

$$n_K(t) = \sup_{0\le s\le t} |Y_1^K(s) - u_1(s)|.$$

Define now the noise amplitude as:

$$\bar{M}_K(t) = \sup_{0\le s\le t} |M_K(s)|\,.$$

Using the convergence of the intensity of the arrival process,

$$\begin{aligned} &n_K(t) \le \bar{M}_K(t) + \varepsilon \\ &\quad + \sup_{s\le t}\left|\frac{1}{K}\int_0^{Ks} \phi_1\left(\frac{X_1^K(z)}{K}, X_2^K(z), \ldots, X_N^K(z)\right)\mathrm{d}z - \int_0^s \bar{\phi}_1(u(z))\,\mathrm{d}z\right|. \end{aligned}$$

Using Step 2, ϕ_1 being Lipschitz and bounded, for K large enough:

$$\begin{aligned} \mathbb{E}\left(\sup_{s\le t}\left|\frac{1}{K}\int_0^{Ks} \phi_1\left(\frac{X_1^K(z)}{K}, X_2^K(z), \ldots, X_N^K(z)\right)\mathrm{d}z - \int_0^s \bar{\phi}_1(u(z))\,\mathrm{d}z\right|\right) \\ \le \varepsilon, \end{aligned}$$

which concludes the proof for the L^1 convergence of Y_1^K.

4. Qualitative behavior of the limiting processes. We now describe the different qualitative behaviors that may occur depending on the traffic conditions. Assume that class i, $i = 1, \ldots, S$, have entered a traffic surge. Under the scaling considered in Theorem 3.1, we observe three qualitative types of behaviors for the network responses, which are completely characterized using the stationary distributions of the family of processes $U_i^x, i = S+1, \ldots, N$. Defining

$$\forall x \in \mathbb{R}_+^S,\ \forall i \in \{1, \ldots, S\},\ \bar{\delta}_i(x) = \lambda_i - \mu_i \bar{\phi}_i(x), \tag{10}$$

let $\mathcal{A}$ the set of positive solutions of the equation

$$\bar{\delta}(x) = 0.$$

Given the classical results on asymptotic stability of non-linear autonomous systems, we can partially classify the possible situations using in particular the Hartman-Grobman theorem (see for instance Hartman (1960)). For a C^1 flow δ, we write $D\delta(x) < 0$ if the linearization of δ has only eigenvalues with strictly negative real parts and no eigenvalue on the unit complex circle. Assume in the following $\bar{\delta}$ is C^1. The possible behaviors are:

1. The network continues to see class i, $i \leq S$, saturated (at a macroscopic time and space scales), even after any large (macroscopic) amount of time. A sufficient condition is that there exists $x \in \mathcal{A}$ such that $D\bar{\delta}(x) < 0$ and $x > 0$, with initial conditions sufficiently close to x. Then the differential equation is asymptotically stable with stable point $x > 0$. In this case, limits in time and K cannot commute since there is always a part of the bandwidth of the network used for surging classes, while taking first the limit in time and then the limit in K always converge to the system with allocation $\phi_i(0, \ldots, 0, x_{S+1}, \ldots, x_N)$ for stable classes.
2. The differential equation (2) governing the dynamics of u is unstable, which means that the traffic surge cannot be absorbed and keeps building up. It might lead to the instability of other classes in the network.
3. The traffic surge will be absorbed at macroscopic time, i.e. the differential equation is asymptotically stable with stable point 0. Necessary conditions for this situation are that:
   (a) $0 \in \mathcal{A}$,
   (b) $D\bar{\delta}(0) < 0$, and
   (c) the initial condition is close enough to 0, or $\mathcal{A} = \{0\}$.

   In this case, note that the stationary measure of $(Y_i^K(t))_{i>S}$ converges when $K \to \infty$ to the stationary measure of the original system with allocation $\phi_i(0, \ldots, 0, x_{S+1}, \ldots, x_N))$, which boils down to the fact that the limit in time and in K commute for classes $S + 1$ to N.

4.1. *Robust bandwidth sharing networks.* Bandwidth sharing networks constitute a natural extension of a multi-class processor-sharing queue, and have become a standard stochastic model for the flow level dynamics of Internet congestion control (they were introduced by Massoulié and Roberts (2002)).

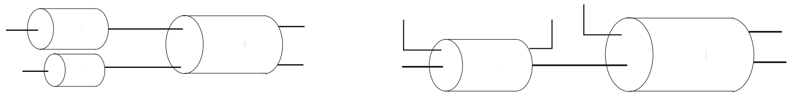

FIG 1. *Tree network and linear network.*

Consider for example the tree network represented on the left of Figure 1, with two traffic routes, each passing through a dedicated link, followed by a common link. If each dedicated link has a capacity $c_i \leq 1$, $i = 1$, 2, and the common link has capacity 1, the flow on each route gets a capacity $\phi_i(x)$ that lies in the polyhedron $\mathcal{C}$:

$$\sum_{i=1}^{2} \phi_i(x) \leq 1, \tag{11}$$

$$\phi_i(x) \leq c_i, \quad i = 1, 2. \tag{12}$$

Another example of interest is the linear network represented on the right of Figure 1 with 3 routes sharing two links.

In general, like for the specific foregoing examples, the capacity constraints determine the space over which a network controller can choose a desired allocation function. It has been argued by Kelly, Maulloo and Tan (1998) that a good approximation of current congestion control algorithms such as TCP (the Internet's predominant protocol for controlling congestion) can be obtained by using the weighted proportional fair allocation, which solves an optimization problem for each vector x of instantaneous numbers of flows. Specifically, the weighted proportional fair allocation $\eta(x)$ for state vector x maximizes

$$\sum_{i=1}^{N} w_i x_i \log(\eta_i), \eta \in \mathcal{C},$$

where the weights w_i are class-dependent control parameters.

REMARK 4.1. By definition of this optimization program, if $\phi(\cdot) = \eta(\cdot)$ is the standard (unweighted) proportional fair allocation with $w_i \equiv 1$, then the allocation $\phi^r(x) = \phi(r.x)$ corresponds to the weighted proportional fair allocation with weights $w_i \equiv r_i$.

This framework has been generalized to so-called weighted α-fair allocations, which provide flexibility to model different levels of fairness in the network. Another important alternative is the balanced fair allocation (Bonald et al., 2006), which allows a closed form expression for the stationary distribution of the numbers of flows in progress. In addition, the balanced fair allocation gives a good approximation of the proportional fair allocation while being easily evaluated, which is attractive for performance evaluation.

Remind that all α-fair bandwidth-sharing are stable for $\alpha > 0$ (in the sense that the process X is positive recurrent, Bonald and Massoulié (2001); de Veciana, Konstantopoulos and Lee (2001)) if

$$\rho \in \mathcal{S} = \{\eta, A\eta \leq C\}.$$

From now on, we refer to the interior of the set $\mathcal{S}$ as the "usual conditions of stability". We also use the term frontier of the stability set to refer to the closure of $\mathcal{S}$ minus its interior. We now refine the concept of stability by saying that the network is robust stable if classes undergoing a surge (penalized adequately) eventually drain while the other classes stay stochastically stable. More formally, let $\mathcal{I} \subset [1, N]$ be the set of indexes of surging classes and $\mathbb{P}([1, N])$ the power set of $[1, N]$.

Definition 4.1. The network is robust stable if for all $i = 1, \ldots, N$:

$$\limsup_{t\to\infty} \sup_{\mathcal{I}\in\mathbb{P}([1,N])} \limsup_{|x_j|\to\infty, j\in\mathcal{I}} \frac{E^x[X_i^{|x|}(|x|t)]}{|x|} = 0.$$

We hence define the robust stability region as the set of parameters such that the network is robust stable, i.e.:

$$\mathcal{S}^r = \Bigg\{\rho \in \mathbb{R}_+^N : \ \forall i = 1, \ldots, N, \quad \limsup_{t\to\infty} \sup_{\mathcal{I}\in\mathbb{P}([1,N])} \limsup_{|x_j|\to\infty, j\in\mathcal{I}} \frac{E^x[X_i^{|x|}(|x|t)]}{|x|} = 0.\Bigg\} \tag{13}$$

Remark that when all classes are undergoing a surge of traffic, we retrieve the usual notion of fluid limit, for which the convergence to 0 implies the network (usual notion of) stability. Hence it holds in general that:

$$\mathcal{S}^r \subset \mathcal{S} = \Bigg\{\rho \in \mathbb{R}_+^N : \ \forall i = 1, \ldots, N, \quad \limsup_{t\to\infty} \limsup_{x_i\to\infty, \forall i=1,\ldots,N} \frac{E^x[X_i^{|x|}(|x|t)]}{|x|} = 0\Bigg\}.$$

In the sequel, we show that:

- we can apply Theorem 1 to bandwidth sharing networks operating under α-fair policies under the usual conditions of stability,
- for work-conserving allocations, the robust stability set coincides with the usual stability set (except possibly on its frontier),
- for monotonic networks, the robust stability set coincides with the set of parameters under which a 'priority allocation' is stable and can hence be *strictly included* in $\mathcal{S}$;
- the situation is much more complex for non-monotone allocations and a full characterization of the robust stability set is still an open problem. The complexity of the dynamics can be explained through the fact that a surging fluid class can "bounce" at 0. We give a simple example of such a phenomenon.

4.2. *Existence of the fluid limit.*

LEMMA 4.1. *For bandwidth sharing networks under α-fair allocations with $\alpha > 0$, if $\rho \in \mathcal{S}$, then the processes U^z are positive recurrent for any z, and Theorem 3.1 applies.*

PROOF. All the mentioned allocations are stable if $\rho \in \mathcal{S}$. U^z represents in that context a network with z permanent customers of classes $i = 1, \dots, S$ and no arrivals and departures for these classes. The usual proof of stability for networks with α-fair allocation relies on the following Lyapunov function (de Veciana, Konstantopoulos and Lee (2001)):

$$F(x) = \sum_{i=1}^{N} x_i^2 r_i \lambda_i^{\alpha-1},$$

using the simple fact that for any vector s in the interior of $\mathcal{S}$, there exists $\varepsilon > 0$, such that $\nabla_\eta U(x,\eta) \cdot (s - \phi(x)) \le -\varepsilon$, where ∇ denotes the usual gradient. We adapt this Lyapunov argument to our context. For $x^S = (x_{S+1}, \dots, x_N)$ such that $|x^S|$ is big enough we have that

$$\begin{aligned}
\Delta F(x^S) &= \sum_{i=S+1}^{N} \lambda_i (F(x^S + e_i) - F(x^S)) \\
&\quad + \phi_i(z, x^S)(F(x^S - e_i) - F(x^S)), \\
&= \sum_{i=S+1}^{N} x_i^S r_i \lambda_i^{\alpha-1} (\lambda_i - \phi_i(z, x^S)) + o(|x^S|), \\
&= \sum_{i=S+1}^{N} x_i^S r_i \lambda_i^{\alpha-1} (\lambda_i - \phi_i(z, x^S))
\end{aligned}$$

$$
\begin{aligned}
&+ \sum_{i=1}^{S} z_i r_i \lambda_i^{\alpha-1}(\lambda_i - \phi_i(z, x^S)) + o(|x^S|), \\
= \quad & \nabla_\eta U(z, x^S).(\lambda - \phi(z, x^S)) + o(|x^S|) \leq -\varepsilon + o(|x^S|),
\end{aligned}
$$

where we used that $\sum_{i=1}^{S} z_i r_i \lambda_i^{\alpha-1}(\lambda_i - \phi_i(z, x^S))$ is bounded. □

4.3. *Work-conserving allocations.* Consider a work conserving allocation such that

$$\forall x \neq 0, \ \sum_{i=1}^{N} \phi_i(x) = 1.$$

It is immediate from a simple Lyapunov argument that every work-conserving allocation has the same stability region, namely $\mathcal{S} = \{\sum_{i=1}^{N} \rho_i < 1\}$. If this stability condition is satisfied, then the priority mechanism considered is asymptotically equivalent to giving full priority to class i, $i \geq S+1$. In other words, the fluid limit obtained for class i, $i \leq S$, is in that case the same as the fluid limit of an allocation that gives a full priority to class i, $i \geq S+1$, which we prove in the following Proposition.

Proposition 4.1. *For a work conserving network, $\mathcal{S}^r = \mathcal{S}$ (except possibly on the frontier of the stability set) and for all $i = 1, \ldots, S$:*

$$Y_i^K(t) \xrightarrow{L^1} u_i(t) = \left(u_i(0) + \lambda_i - \mu_i \left(1 - \sum_{j=S+1}^{N} \rho_j \right) t \right)^+ .$$

Proof. Note that, by definition, $\mathcal{S}^r \subset \mathcal{S}$. We need, therefore, to show that $\mathcal{S}^r \supset \mathcal{S}$ to conclude the proof.

Assume $\sum_{i=1}^{N} \rho_i < 1$ in which case the network is stable. Fix $z_1 \in \mathbb{R}$. Using the conservation of the rates at equilibrium for the process U^{z_1} (which boils down in the Markovian context to saying that at equilibrium the drift of $y \to y_i$ should be 0), we can write that:

$$\sum_{i=S+1}^{N} \sum_{y \in \mathbb{Z}_+{}^{N-S}} \phi_i(z_1, y) \pi^{z_1}(y) = \sum_{i=S+1}^{N} \rho_i.$$

We now calculate $\bar{\phi}_i$ for $z_1 > 0$:

$$\bar{\phi}_i(x) = \sum_{y \in \mathbb{Z}_+{}^{N-S}} \phi_i(z_1, y) \pi^{z_1}(y),$$

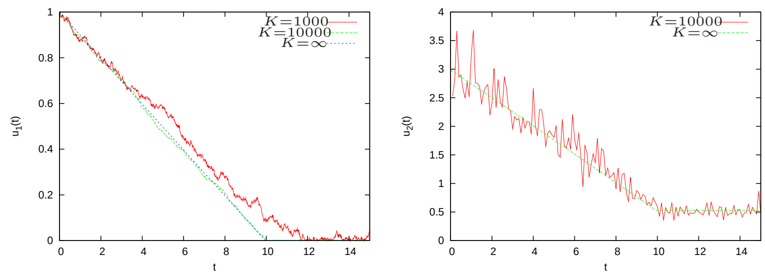

FIG 2. *DPS with three classes: scaling of class-1 (left) and of class-2 (right).*

$$\bar{\phi}_i(x) = \sum_{y \in \mathbb{Z}_+{}^{N-S}} (1 - \sum_{j \geq S+1} \phi_j(z_1, y))\pi^{z_1}(y),$$
$$\bar{\phi}_i(x) = 1 - \sum_{j \geq S+1} \rho_j.$$

Hence, the capacity seen asymptotically by class 1 is $(1-\sum_{j\geq S+1}\rho_j)$, which concludes the proof. □

Example: One link with the DPS allocation. The simplest instance of a network consists of one link shared by several competing classes of traffic. If the initial policy is supposed to be the classical processor sharing policy: $\phi_i(x) = \frac{x_i}{|x|}$ then the prioritized version of the model becomes the so-called discriminatory processor sharing (DPS): $\phi_i(r \cdot x) = \frac{r_i x_i}{\sum_j r_j x_j}$.

Consider a single link of capacity 1 shared by three classes. The bandwidth is allocated according to DPS with weight r_i for class i, $i = 1, 2, 3$. Proposition 4.1 says that $u_1(t)$ is a straight line with slope $\lambda_1 - \mu_1(1 - (\rho_2 + \rho_3))$. This behavior is illustrated in Figure 2, for which $\lambda_1 = 0.5$, $\mu_1 = 1$, $\rho_2 = 0.3$, $\rho_3 = 0.1$. The slope calculated using the proposition is thus 0.1, which is verified in the figure.

In Figure 2, we plot the empirical mean of class 2 at a macroscopic scale, (i.e. $\frac{1}{s}\int_t^{t+s} f(Y^K(h))\,\mathrm{d}h$) for a temporal window of $s = 0.1$.

4.4. *Monotone allocations.* Define the allocation ψ giving full priority to classes $S+1$ to N, given by

$$\psi(x) = \phi(0, \ldots, 0, x_{S+1}, \ldots, x_N), \text{ if } x_i > 0, \text{for some } i > S.$$
$$\psi(x) = \phi(x), \text{ otherwise.}$$

Denote $\mathcal{S}(\psi)$ the stability region of the network with allocation ψ.

PROPOSITION 4.2. *Consider a monotonic allocation (i.e. such that ϕ_i is decreasing in x_j, $j \neq i$).*

If $\bar{\delta}(0) < 0$, where $\bar{\delta}(\cdot)$ is defined by (10), then the network is robust stable and surging classes do not influence asymptotically stable classes. Conversely if $\bar{\delta}(0) > 0$, the network is not robust stable.

Moreover $\bar{\delta}(0) < 0$ if the network with allocation ψ is stable i.e.:

$$\mathcal{S}^r(\phi) = \mathcal{S}(\psi).$$

PROOF. Using stochastic comparisons (see Borst, Jonckheere and Leskelä (2008) for more details on stochastic comparisons of multidimensional birth-and-death processes), we obtain that for all $i = S+1, \dots, N$,

$$U_i^0 \leq_{\mathrm{st}} U_i^z, \, \forall z,$$

which implies that $\forall z, \; \bar{\phi}_i(z) \geq \bar{\phi}_i(0)$, $i = 1, \dots, S$. This in turn implies that, for $i = 1, \dots, S$,

$$\frac{d}{dt} u_i(t) < \bar{\delta}(0) < 0, \quad \forall t \geq 0, \text{ such that } u_i(t) > 0.$$

This implies that u_i will reach 0 in finite time.

The reverse statement follows along the same lines. □

Example: A tree network. Let us consider the tree network shown in Figure 1 with $c_1 = 0.4$ and $c_2 = 0.8$. We shall assume the following bandwidth allocation: Define $\mathcal{S}_1 = \{(x_1, x_2) : (r_1 x_1 + r_2 x_2) c_1 < r_1 x_1\}$. For $x_1 > 0$ and $x_2 > 0$,

$$\phi_1(x_1, x_2) = \begin{cases} c_1, & \text{if } (x_1, x_2) \in \mathcal{S}_1, \\ \max\left(\frac{r_1 x_1}{r_1 x_1 + r_2 x_2}, 1 - c_2\right), & \text{if } (x_1, x_2) \in \mathcal{S}_1^c, \end{cases} \tag{14}$$

and $\phi_2 = 1 - \phi_1$.

For this network, the allocation becomes a strict priority allocation for class 2 when $r_1 = 0$, in which case class 1 gets capacity c_1 if there are no class 2 flows, and $1 - c_2$ otherwise. Thus, for a fixed value of ρ_2, class 1 is stable if $\rho_1 < (1 - \frac{\rho_2}{c_2}) c_1 + \frac{\rho_2}{c_2}(1 - c_2)$. The stability regions for $r_1 = 0$ and $r_1 > 0$ are shown in Figure 3. This is an example where $\mathcal{S}^r$ is strictly included in $\mathcal{S}$.

The dynamics of $u_1(t)$ for two different values of ρ_1 – one in each region – is plotted in Figure 4, for which $\rho_2 = 0.5$.

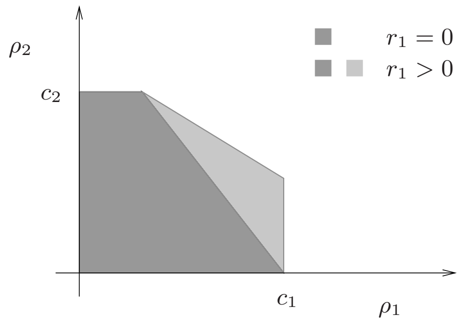

FIG 3. *Partitioning of the stability region for the tree network.*

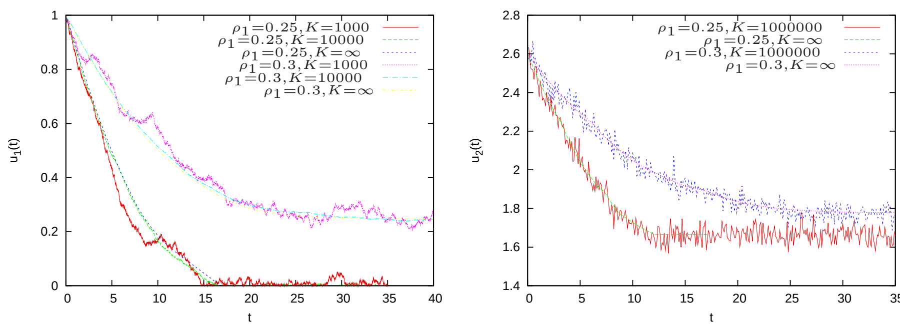

FIG 4. *Tree network: scaling of class* 1 *(left) and of class* 2 *(right).*

For class 2, when the priority allocation is stable the dynamics of the average number of customers converges to the one of the priority allocation, that is $\rho_2/(c_2 - \rho_2)$, as is illustrated in Figure 4.

In Figure 5, we show how class 1 is actually favored by asymptotically using the bandwidth of class 2, compared to the case where class 2 is given a strict priority. Though the stability conditions are identical for the two allocations, their performances at a fluid scale are *very different*.

4.5. *Non-monotone networks with multiple unstable classes bouncing at* 0. For a general multi-class network, characterizing the robust stability region is still an open question and cannot be answered using priority allocations as for monotonic networks. We now present a scenario with multiple unstable classes emphasizing the complexity of the dynamical systems obtained at the limit when the network is not monotonic. Consider the linear network of two links and three classes shown in Figure 6. Each unstable class is penalized in inverse proportion to the scale of its initial surge. Since the scale of the initial congestion is the same for the three classes, and the rate allocation

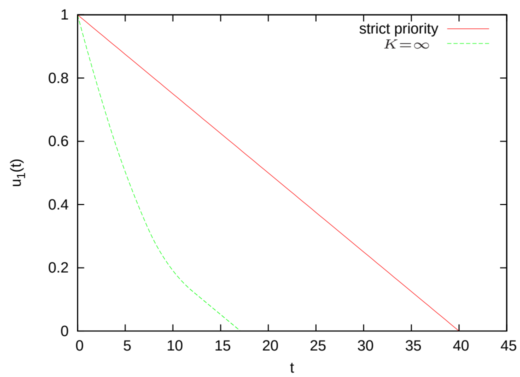

FIG 5. *Tree network: comparisons of trajectories of class* 1 *for a proportional fair allocation and a priority (to class* 2*) allocation.*

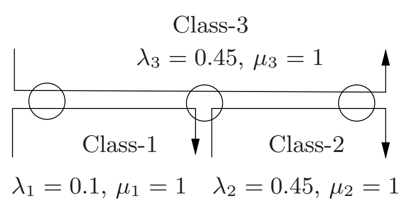

FIG 6. *A linear network with two links.*

to each class is a homogeneous function of degree zero, our proposed scaling results in the usual fluid limit.

Even though this classical fluid limit has been widely investigated, the following phenomenon has not been studied in detail: a class which drains out (that is, becomes stable) could become unstable later on. As an example, consider the following network parameters with the arrival rates and the service rates as shown in Figure 6:

$$c_1 = c_2 = 1,\ r_1 = r_2 = r_3 = 1,$$
$$X_1(0) = 10 \cdot K, X_2(0) = K, X_3(0) = K,$$

that is, the three classes have a macroscopic state at the beginning. The trajectories of the number of flows as a function of the scaled time for $K = 10000$ is shown in Figure 7. Observe that class 2 becomes stable around the 2 time unit mark, then becomes unstable around the 15 time unit mark, and becomes stable again around the 40 time unit mark. This behavior can be explained as follows.

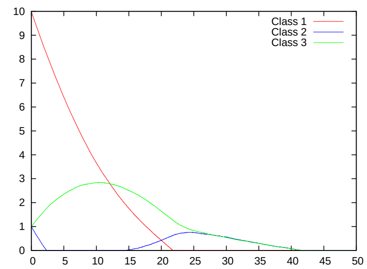

FIG 7. *The trajectories of the scaled number flows as a function of the scaled time.* $K = 10000$.

Class 1 has a larger initial surge than that of class 3 which has the same initial surge as class 2. Thus, class 1 gets a higher share of bandwidth on link 1 which creates a bottleneck for class 3 flows on that link. This leaves sufficient capacity for class 2 to drain out while the number of class-3 flows grows.

Once the number of class-1 flows has reduced sufficiently (around the same number as that of class-3 flows), class 3 starts to get sufficient capacity to go towards stability. However, the arrival rate of class-3 flows is higher than that of class-1 flows which means that the rate of decrease of the number of class-3 flows is smaller than that of the class-1 flows. Eventually, the proportion of flows of class 1 is smaller compared to that of class 3. This allows class- 3 flows to get a larger share of bandwidth link 1. More importantly, since link 2 is not a bottleneck for class-3 flows, this also allows class 3 to get a larger share on link 2. The arrival and service rates of class 2 and class 3 being the same, the imbalance in the rate allocation means that class 2 is now unstable and its number of flows starts to grow until it reaches the same number as that of class-3 flows at which time they both share the link capacity equally. Since the network is stable, all the three classes drain out eventually.

5. Integration of streaming and elastic traffic. Consider now a system where two intrinsically different types of traffic – "streaming" and "elastic" traffic – coexist and share a link of capacity 1. Such models have been considered by Núñez-Queija, van den Berg and Mandjes (1999); Delcoigne, Proutière and Régnié (2004); Bonald and Proutière (2004).

It is natural to equip streaming traffic with a fixed required rate, say, c per flow. On the other hand, the rate allocated to the elastic traffic is variable. Giving priority to streaming traffic (class 2) the allocation of service may be chosen as:

$$\phi_1(x) = \min\left(\frac{r_1 x_1}{r_1 x_1 + c x_2}, 1 - c x_2\right),$$
$$\phi_2(x) = c x_2,$$

where the parameter r_1 quantifies the level of priority. The allocated capacity to class 1 cannot exceed its fair share, which is assumed to be $r_1 x_1/(r_1 x_1 + c x_2)$. Since each streaming session is guaranteed a rate c, the total allocation to class 1 cannot exceed $1 - c x_2$ either.

The allocated capacity cannot exceed the total capacity, and the state-space must be restricted to states x_2 such that $\phi_1(x)+\phi_2(x) \leq 1$. Then, if the number of current streaming flows x_2 is such that $\phi_1(x+e_2)+\phi_2(x+e_2) > 1$, arriving streaming flows must be blocked from the network. The capacity that remains is a provision for future incoming streaming calls. The above allocation is asymptotically work-conserving, and it follows that the system is stable under the usual condition: $\rho_1 + \rho_2 < 1$.

In Kumar and Massoulié (2007), fluid and diffusion approximations are derived for a similar model when both streaming and elastic flows are evolving at a fast time-scale. In contrast, in this section we will be interested in the blocking probability of streaming flows when the elastic flows undergo a traffic surge which means that only the elastic flows will be scaled to a deterministic limit whereas the streaming flows will retain their stochastic nature.

In networks in which streaming and elastic traffic do not interact, the probability of blocking can be computed using an Erlang Fixed-Point approximation Kelly (1986). However, in the context of the present example, the interaction of these two types of traffic makes it more difficult to apply these fixed-point approximations, mainly due to the fact that the state space of the elastic flows is unbounded. In the regime when r_1 is small, we propose a rule-of-thumb, based on the Theorem 3.1, that can guarantee a blocking probability smaller than a desired value.

Let p_m denote the desired maximal blocking probability of class-2 flows. We shall set the priority level of class 1 (by varying r_1) such that the probability of blocking of class-2 flows is always less than p_m. Performing the scaling previously defined, remark that the state-space of class 2 depends, for a fixed macroscopic state z of class 1, on both z and c. Denote

$$\mathcal{S}_z = \left\{ x_2 : \frac{z}{z + c x_2} + c x_2 \leq 1 \right\},$$

the state-space of class 2 given that $u_1(t) = z$. It can be seen that an arrival of class 2 is blocked if and only if there are already $\lfloor \frac{1-z}{c} \rfloor$ calls of class 2 present. Define $\rho_2 = \frac{\lambda_2}{\mu_2 c}$. The process U_2^z is a birth-death process with birth rate λ_2 and death rate $\mu_2 c x_2$, and whose stationary distribution conditioned on z is given by

$$\pi_2^z(x_2) = \frac{1}{\sum_{j \in \mathcal{S}_z} \rho_2^j / j!} \frac{\rho_2^{x_2}}{x_2!}.$$

For $u_1(t) = z$, the blocking probability of class 2 is then

$$\pi_2^z(x_2)|_{x_2 = \lfloor \frac{1-z}{c} \rfloor} =: g(z).$$

Let $\bar{z} := \sup\{z : g(z) \leq p_m, 0 < z < 1\}$. Since $g(z)$ is a non-decreasing function of z, in order to guarantee a maximal blocking of p_m, $u_1(t)$ has to be smaller than $\bar{z}$ for all t. This leads us to the following necessary and sufficient condition which guarantees the desired quality of service:

$$\bar{u}_1 := \sup_{0 \leq t < \infty} u_1(t) < \bar{z}. \tag{15}$$

We can compute $\bar{u}_1$ as follows. From Theorem 3.1, the dynamics of $u_1(t)$ depends on the average capacity allocated to class 1, which an be computed as:

$$\bar{\phi}_1(z) = \sum_{x_2 \in \mathcal{S}_z} \min\left(\frac{z}{z + c x_2}, 1 - c x_2\right) \frac{\rho^{x_2}}{x_2!} C(z),$$

where $C(z) = (\sum_{x_2 \in \mathcal{S}_z} \frac{\rho^{x_2}}{x_2!})^{-1}$. In the case that c is very small ($c \ll 1$), we might consider as a reasonable approximation a Poisson distribution for class 2, whatever the state of class 1. In that case, $\bar{\phi}_1$ takes a slightly simpler form. After simple calculations:

$$\bar{\phi}_1(cz) = H(z) = \frac{z \int_0^{\rho_2} u^{z-1} \exp(u) \, \mathrm{d}u}{\rho_2^z \exp(\rho_2)}.$$

Using the monotonicity of ϕ_1 in its first variable, we can conclude that $u_1(t)$ converges monotonically to its limit point, and that

$$\bar{u}_1 = \begin{cases} u_1(0), & \text{if } \lambda_1 < \bar{\phi}_1(u_1(0)); \\ \bar{\phi}_1^{-1}(\lambda_1), & \text{otherwise,} \end{cases}$$

The inequality (15) can be ensured by scaling the process $u_1(t)$ by a factor $\bar{z}/\bar{u}_1$, which, in turn, can be achieved by scaling the priority level (or, equivalently, r_1) by this very same factor. This additional scaling results in a larger share of the bandwidth for class-1 flows in case $\bar{z} > \bar{u}_1$. Conversely, if $\bar{z} < \bar{u}_1$, the priority level of class-1 flows is appropriately decreased so that the blocking probability constraint of class-2 flows is not violated.

REMARK 5.1. In a network of links shared by several classes of streaming flows and one class of elastic flow, we could use fixed-point approximations to compute the blocking probability for the different classes of streaming flows as a function of z_1. Assuming that this probability is increasing in z_1, we could then compute the maximum value that z_1 can attain without the streaming classes violating their individual blocking probability.

6. Conclusions.

We analyzed the flow-level performance of multi-class communication networks when one or more classes undergoes a traffic surge. We showed that, under an appropriate scaling of space and time and amount of penalization, the dynamics of the temporarily unstable class can be described by a deterministic differential equation in which the time derivative at a given point depends on the conditional stationary distribution of the other classes calculated at that point. For work-conserving allocations, the differential equation is the same as the one of the network in which other classes have strict priority over the temporarily unstable class, that is, the scaled process evolves linearly and is either absorbed at zero or grows indefinitely depending on whether the network is stable or not.

For non-work conserving allocations, the trajectory is much more complex to describe as it depends on the mean residual bandwidth left over by the other classes which in turn depends on the current state of the first class. The limit point of the fluid trajectory can hence be non-zero and finite. We characterized the robust stability region of monotone allocations. We illustrated this behavior through several examples of network topologies and bandwidth allocations that are commonly used to model communication networks.

The time-space-transitions scaling that we considered raises several open questions which would give a better understanding of the network dynamics. In particular, finding necessary and sufficient conditions for the limit point of non work-conserving allocations to be zero would constitute a very interesting result. Also, error bounds estimates would be necessary to obtain a reliable performance evaluation tool.

Acknowledgment. M.J. and B.P. would like to warmly thank R. Núñez Queija for many fruitful discussions on the subject.

REFERENCES

BILLINGSLEY, P. (1999). *Convergence of Probability Measures (second edition). Wiley Series in Probability and Statistics*. Wiley-Interscience. MR1700749

BONALD, T. and MASSOULIÉ, L. (2001). Impact of fairness on Internet performance. In *Proceedings of ACM SIGMETRICS/Performance* 82–91.

Bonald, T. and Proutière, A. (2003). Insensitive bandwidth sharing in data networks. *Queueing Syst. Theory Appl.* **44** 69–100. MR1989867

Bonald, T. and Proutière, A. (2004). On performance bounds for the integration of elastic and adaptive streaming flows. In *SIGMETRICS* 235-245.

Bonald, T., Massoulié, L., Proutière, A. and Virtamo, J. (2006). A queueing analysis of max-min fairness, proportional fairness and balanced fairness. *Queueing Syst. Theory Appl.* **53** 65–84. MR2230014

Borst, S. C., Jonckheere, M. and Leskelä, L. (2008). Stability of parallel queueing systems with coupled service rates. *Discrete Event Dyn. Syst.* **18** 447–472. MR2443653

C. Kipnis, C. L. (1991). *Scaling Limits of Interacting Particle Systems*. Springer. MR1707314

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. MR1325041

Darling, R. W. R. and Norris, J. R. (2008). Differential equation approximations for Markov chains. *Probability Surveys* **5** 37. MR2395153

de Veciana, G., Konstantopoulos, T. and Lee, T.-J. (2001). Stability and performance analysis of networks supporting elastic services. *IEEE/ACM Trans. Netw.* **9** 2–14.

Delcoigne, F., Proutière, A. and Régnié, G. (2004). Modeling integration of streaming and data traffic. *Perform. Eval.* **55** 185–209.

Deshpande, M., Amit, A., Chang, M., Venkatasubramanian, N. and Mehrotra, S. (2007). Flashback: a peer-to-peer web server for flash crowds. *International Conference on Distributed Computing Systems* 15.

Engelking, R. (1977). *General Topology. Monografie Matematyczne*. PWN. MR0500780

Ethier, S. N. and Kurtz, T. G. (1986). *Markov Processes: Characterization and Convergence*. Wiley. MR0838085

Feuillet, M. (2012). On the flow-level stability of data networks without congestion control: the case of linear networks and upstream trees. *Queueing Systems, Theory and Applications* **70** 105–143. MR2886479

Gromoll, H. C. and Williams, R. J. (2009). Fluid limits for networks with bandwidth sharing and general document size distributions. *Annals of Applied Probability* **19** 243. MR2498678

Hartman, P. (1960). A lemma in the theory of structural stability of differential equations. *Proc. A.M.S.* **11(4)** 610–620. MR0121542

Jonckheere, M., Núñez-Queija, R. and Prabhu, B. (2010). Performance analysis of traffic surges in multi-class communication networks. In *ITC 22*.

Kandula, S., Katabi, D., Jacob, M. and Berger, A. . W. (2005). Botz-4-Sale: surviving organized DDoS attacks that mimic flash crowds. In *2nd Symposium on Networked Systems Design and Implementation (NSDI)*.

Kelly, F. (1979). *Reversibility and Stochastic Networks*. Wiley. MR0554920

Kelly, F. (1986). Blocking probabilities in large circuit-switched networks. *Adv. Appl. Probab.* **18** 473–505. MR0840104

Kelly, F. P., Maulloo, A. K. and Tan, D. K. H. (1998). Rate control for communication networks: shadow prices, proportional fairness and stability. *The Journal of the Operational Research Society* **49** 237–252.

Kumar, S. and Massoulié, L. (2007). Integrating streaming and file-transfer Internet traffic: fluid and diffusion approximations. *Queueing Syst.* **55** 195-205. MR2324634

Kurtz, T. G. (1992). Averaging for martingale problems and stochastic approximation. In *Applied Stochastic Analysis, US-French Workshop. Lecture Notes in Control and Information Sciences* **177** 186–209. Springer Verlag. MR1169928

Massoulié, L. and Roberts, J. (2002). Bandwidth sharing: objectives and algorithms. *IEEE/ACM Trans. Netw.* **10** 320–328.

Meyn, S. (2008). *Control Techniques for Complex Networks*. Cambridge University Press. MR2372453

Núñez-Queija, R., van den Berg, J. and Mandjes, M. (1999). Performance evaluation of strategies for integration of elastic and stream traffic. In *ITC 16* 1039-1050.

Peng, T., Lecki, C. and Ramamohanarao, K. (2007). Survey of network-based defense mechanisms countering the DoS and DDoS problems. *ACM Comput. Surv.* **39**.

Robert, P. (2003). *Stochastic Networks and Queues*. *Stochastic Modelling and Applied Probability Series*. Springer-Verlag, New York. xvii+398 pp. MR1996883

Rogers, L. C. G. and Williams, D. (2000 (1987)). *Diffusions, Markov Processes & Martingales Vol. 2: Itô Calculus*. Cambridge University Press. MR1780932

Segel, L. A. and Slemrod, M. (1989). The quasi-steady-state assumption: a case study in perturbation. *SIAM Review* **31** 446–477. MR1012300

Stavrou, A., Rubenstein, D. and Sahu, S. (2004). A lightweight, robust P2P system to handle flash crowds. *Selected Areas in Communications, IEEE Journal on* **22** 6–17.

Paris-Rocquencourt
Domaine de Voluceau
Rocquencourt, B.P. 105
78153 le Chesnay Cedex, France
E-mail: mathieu.feuillet@inria.fr

Departamento de Matemática
Facultad de Ciencias Exactas y Naturales
Universidad de Buenos Aires
Pabellón 1, Ciudad Universitaria
1428 Buenos Aires, Argentina
E-mail: mjonckhe@dm.uba.ar

CNRS, LAAS
7 avenue du Colonel Roche
F-31400 Toulouse, France
and
Univ de Toulouse, LAAS
F-31400 Toulouse, France
E-mail: balakrishna.prabhu@laas.fr