On the Limitation of Fluid-based Approach for Internet Congestion Control

Do Young Eun
Department of Electrical and Computer Engineering
North Carolina State University
dyeun@ncsu.edu

Abstract—**Fluid models have been the main tools for Internet congestion control. By capturing how the average rate of each flow evolves, the fluid model proves to be useful as it predicts the equilibrium point to which system trajectory converges and also provides conditions under which the convergence is ensured, i.e., the system is stable. However, due to inherent randomness in the network caused by random packet arrivals or random packet marking, the actual system evolution is always of a stochastic nature. In this paper, we show that we can be better off using a stochastic approach toward the congestion control. We first prove that the equilibrium point of a fluid model can be quite different from the true average rate of the corresponding stochastic system. After we describe the notion of stability for two different approaches, we show that a stable fluid model can impose too much restriction on our choice of system parameters such as buffer size or link utilization. In particular, under fluid models, we show that there exists a fundamental tradeoff between the link utilization and buffer size requirement for large systems, while in a more realistic setting with stochastic models, there is no such tradeoff. This implies that the current congestion control design can be much more flexible, to the benefit of efficient usage of network resources.**

I. Introduction

Transmission Control Protocol (TCP) coupled with various Active Queue Management (AQM) schemes are key components of the current congestion control and responsible for more than 90% of total traffic carried by the Internet. Since the seminal work in [1], many variations and improvements on the congestion control have been made via the use of the so-called fluid modeling and through its stability analysis [2], [3], [4], [5]. By capturing the mean-level change in each flow's injection rate and ensuring its convergence to an equilibrium point, the fluid model approach, usually in the form of a differential or a difference equation, turns out to be extremely effective and versatile in the analysis and design of various congestion controlled systems depending on the type of underlying versions of TCP and AQM.

Another nice characteristic of this fluid-based approach is that the congestion control can be explained as a distributed algorithm toward solving a global optimization where the cost functions (or utility functions) of flows are chosen such that certain fairness criterion is achieved at the optimal allocation of the rates [1], [6], [7]. In this case, the fluid-based differential equation or difference equation for the rate of each flow can be interpreted as an iteration (or rate-updates) of its rates over a series of round-trip-times (RTTs), and by stability of the system we mean that these rate update equations converge to the equilibrium point, i.e., the optimal rate allocation.

In addition to the above, the criterion for system stability (convergence) produces many different design guidelines on several system parameters such as the buffer size, target link capacity per flow, marking probability, or its slope. Although the resulting stability criterion varies case by case depending on the choice of TCP/AQM and the modeling framework under consideration, they all share the same rules-of-thumb for system stability. For example, it is well known [3], [4], [8] that the marking slope at the equilibrium point should be kept small to make the system at least linearly stable. Similarly, the system is getting more unstable for larger RTT (delay).

On the other hand, there exists another body of work on the congestion control based on a stochastic description of the network [9], [10], [11]. This approach often provides the distribution of the performance measure of the system, e.g., window size evolution, possibly at the cost of more difficult and sometimes unwieldy analysis. The main tool in this area is the Markov process approach; given the current network status or event such as the window sizes (or rates) and the existence of any congestion signal, the TCP/AQM algorithm completely specifies (at least probabilistically) the network status in the next RTT based on the information available at the current RTT. In this case, the notion of stability of the system is not about the convergence of the mean rate to an equilibrium point (real number), but refers to a stationary behavior of the system (or "ergodicity") starting from any initial distribution.

In this paper, we provide several arguments in favor of the stochastic description of congestion control and address some possible limitations inherent in the fluid-based approach. First, as the simplest case with a single flow over a single bottleneck link, we show that the equilibrium (or the fixed point) of the fluid-based model can be much different from the true average of the stochastic counterpart, in which we have some randomness either through random packet marking or random packet arrivals. Depending on the marking function employed at the link, we prove that the fixed point can be strictly larger or smaller than the real expected value of the rate (or throughput). This implies that the equilibrium point of the fluid model may not be the actual steady-state point of the system and the corresponding (linear) stability criterion can also be problematic since the system would have been linearized around possibly an inaccurate equilibrium point.

In case of many flows accessing a shared link, however, it has been shown that a simple fluid-based model becomes accurate enough to capture the dynamics of the "averaged" behavior of the system [12], [5], [13], through which many stability criteria or design guidelines have been obtained. With fluid models, we point out that there exists a fundamental tradeoff between the link utilization and the buffer size requirement for system stability. In particular, if we fix the buffer size as the number of flows and the size of capacity both increase, the link will suffer reduced utilization (strictly less than one). In contrast, if we want to achieve full link utilization and system stability, then the buffer size should be scaled in proportion to the size of capacity and the number of flows. However, under the stochastic model as in reality, we show that there exists a way to scale the system to achieve full utilization, while keeping the buffer size requirement much smaller. In this regard, we believe that the current fluid-based approach toward congestion control may place too much unnecessary restriction on the choice of system parameters.

This paper is organized as follows. In Section II, we describe some basics of a fluid model and a Markovian model for congestion control and explain the different notions for stability in each case. In Section III, we prove that under some conditions the equilibrium points for two different models can be quite different. We also provide popular examples for the marking function for which our assumption is satisfied. In Section IV, we discuss the implication of the stability criterion for both approaches and illustrate that fluid-based approach can be unduly restrictive for system design. We then conclude in Section V.

II. Fluid Model vs. Markov chain for Congestion Control

Both fluid-models and Markov chains have been widely used in the analysis and design of network protocols and congestion controllers. The range of their applications is so vast in the literature that we by no means attempt to cover them all here. Instead, we here use a simple example to illustrate the main difference here.

A. Fluid Model in Discrete Time

Consider a simple congestion controller with a single flow adapting its rate according to an additive-increase-multiplicative-decrease (AIMD) algorithm. Let the round-trip-time (RTT) be equal to one time slot. Then, it is well-known that the *average behavior* of the flow's rate $x(t)$ at time slot t ($t = 1, 2, \ldots$) can be described by the following fluid model in a discrete time setting:

$$\begin{aligned} x(t+1) &= (x(t)+1)(1-p(x(t))) + \frac{x(t)}{2}p(x(t)) \\ &:= g(x(t)), \end{aligned} \tag{1}$$

where

$$g(x) := (x+1)(1-p(x)) + xp(x)/2. \tag{2}$$

Here, the function $p(x) \in [0,1]$ is the marking function denoting the probability that the flow receives congestion signal when the current rate of the flow is x. Then, as usual, (1) has an interpretation that the average flow's rate will increase by one if no congestion and reduced by half under congestion.

The stability of the fluid model in (1) requires that $\lim_{t\to\infty} x(t) = x^*$ where the equilibrium point x^* satisfies

$$x^* = g(x^*) \iff 1 = \left(1 + \frac{x^*}{2}\right) p(x^*). \tag{3}$$

The system is called globally stable if the convergence holds for any initial point $x(0)$ and locally stable if it holds only when the initial point $x(0)$ is sufficiently close to x^*. Similar to the techniques based on the Laplace transform for differential equations, we can also easily obtain a condition for local stability for the difference equation in (1). Note that

$$x(t+1) - x^* = g'(x^*)(x(t) - x^*) + o(x(t) - x^*).$$

The system is then linearly stable if and only if

$$|\ g'(x^*)\ | \ < \ 1, \tag{4}$$

where $g'(x)$ is the derivative of the function $g(x)$. In other words, the system is locally 'contractive' around x^*, and the linear stability ensures convergence to x^* when $x(0)$ is sufficiently close to x^*. Since $p(x)$ is non-negative and increasing (non-decreasing) in x, we can rewrite the condition in (4) in terms of p as follows:

$$\frac{p(x^*)}{2} + \frac{p'(x^*)}{p(x^*)} < 2, \tag{5}$$

where x^* is the unique solution of (3).

Examples: Let us consider the following examples for $p(x)$:

$$\begin{aligned} p_1(x) &= \left(\frac{x}{C}\right)^B && (6) \\ p_2(x) &= \exp\left(\frac{-2B(C-x)}{\sigma^2 x}\right). && (7) \end{aligned}$$

(6) can be interpreted as the probability of queue length being larger than B in an $M/M/1$ queue with arrival rate x and service rate C. Similarly, (7) corresponds to a case where packets over one unit time (RTT) arrive to a queue with capacity C according to a Gaussian process with mean x and variance $\sigma^2 x$. Then, $p_2(x)$ is the probability that an arriving packet finds the queue-length to be larger than B [3]. Both of these models for marking functions are known to capture the packet-level behavior as they explicitly take into account the random packet arrivals to the queue by modeling the arrival process either by a Poisson (for $p_1(x)$) or a Gaussian process (for $p_2(x)$). By substituting (6) and (7) into (3) and (5), we can calculate the equilibrium point x^* and the corresponding local stability conditions for each of these marking functions.

B. Markovian Models

We consider the same AIMD-like algorithm in Section II-A with one RTT set to one unit time slot. Let $X(t)$ be the *actual* rate (or the window size since RTT is set to be unity) of the flow at time slot t. Then, depending on the flow's current rate $X(t)$ at time slot t, the flow's actual rate at the next RTT will be always either incremented by one with probability $1 -$

$p(X(t))$ or reduced by half with probability $p(X(t))$. In other words, the actual rate of a flow $X(t)$ can be described as the following simple (state-dependent) Markov chain:

$$X(t+1) = \begin{cases} X(t)+1, & \text{with probability } 1-p(X(t)) \\ X(t)/2, & \text{with probability } p(X(t)) \end{cases} \quad (8)$$

Our interpretation is as follows. Suppose the current flow's rate (or simply the number of packets in the current RTT) is $X(t)$. Then, given that $X(t) = x$, the number of packets arriving to the queue of interest is x. Since each of these packet will be randomly jittered in its arrival time to the queue, it causes the queue-length to fluctuate even when $x < C$. This random arrival nature is already reflected in (6) and (7) depending on the choice of arrival model. However, since the rate (or the window size) of the flow at the next RTT can only be *either* $x+1$ with probability $1-p(x)$ *or* $x/2$ with probability $p(x)$ (never in between), it naturally leads to the Markovian description as in (8).

Another equivalent representation of (8) is possible via the so-called iterated random function [14]. Define

$$f(x,u) := (x+1)(1-1_{\{u\le p(x)\}}) + \frac{x}{2}1_{\{u\le p(x)\}}. \quad (9)$$

Let $U_1, U_2, \ldots,$ be $i.i.d.$ uniform random variables over $[0,1]$ and *independent of everything else*. Then, we see that (8) conveniently takes on the following form:

$$X(t+1) = f(X(t), U_t) \quad t = 1, 2, \ldots. \quad (10)$$

In view of (8) – (10), we can rewrite the fluid model in (1) as follows:

$$\begin{aligned} x(t+1) &= E\{X(t+1)|X(t) = x(t)\} \quad \text{or,} \\ x(t+1) &= E_{U_t}\{f(x(t), U_t)\} = \int_0^1 f(x(t), u)\, du, \end{aligned}$$

C. Stability of the Markov Chain

Any suitably designed network algorithm or protocol is expected to be "stable". The notion of the stability of communication networks is in general the most important concept and has been the first condition that needs to be satisfied in every occasion.

The stability of a Markov chain means "ergodicity" of the chain in that, starting from any initial distribution, the chain eventually converges (in the sense of total variation) to a stationary version, where the distribution of a stationary chain does not depend on time [15]. Thus, by enforcing the stationarity of the system, we expect that any performance measure, usually defined by an expectation (over the stationary distribution) of some function of the state $X(t)$, is well-defined and does not change in time. Let us consider again the chain in (8). It is easy to see that for any 'well-defined' non-decreasing function $p(x)$, the chain in (8) becomes aperiodic and irreducible. Further, observe that

$$\begin{aligned} &\limsup_{x\to\infty} E\{X(t+1) - X(t) \mid X(t) = x\} \\ &= \limsup_{x\to\infty} 1 - p(x) - \frac{xp(x)}{2} < 0, \end{aligned}$$

as long as $\lim_{x\to\infty} p(x) > 0$ (which is always the case). Thus from Pakes's Lemma [15], the chain is always ergodic and converges to a stationary version of $X(t)$ starting from any initial distribution. In other words, the "stability" of the Markov chain in (8) is always guaranteed.

III. Discrepancy in Equilibrium Points

The fluid model or the recursion in (1) serves two purposes. First, it is intended for capturing the change in the average rate over RTTs. Second, by ensuring the stability of the fluid model, we also expect that the original system works reasonably well with good performance. In this section, we show that under some conditions the equilibrium point of the fluid model can deviate a lot from the actual expected value obtained via the Markovian model.

In general, an exact solution (in a distributional sense) for the simple Markov chain given by (8), if exists, is quite complicated even for a very simple function $p(x)$. (See [9] for example.) Instead, we can think of an approximation using the fluid model in (1) to make the analysis feasible and tractable under more general settings. As suggested in [2], an approximation using a fluid model can be justified by capturing the average of system states (random variables), i.e., $x(t) \approx E\{X(t)\}$, and focusing on how this average changes in response to different system reactions triggered by the congestion signal. Also, observe that the Markov chain in (8) is always ergodic, and thus there always exists a stationary measure for the chain, i.e., a stationary distribution (in time t) for $X(t)$ satisfying (8) or (10). Let π be the stationary distribution of $X(t)$ and let $\hat{x}$ be the steady-state mean of $X(t)$, i.e.,

$$\hat{x} = E_\pi\{X(t)\}, \quad \forall\, t,$$

where the expectation is taken over π. Since the original Markov chain always enters the steady-state in which the distribution does not depend on t, we expect that $x^* = \hat{x}$ if the fluid model were to be "close" to the original Markov chain and to capture the "average behavior" of the system.

Proposition 1: Assume that $X(t)$ in (8) is not always a constant (non-degenerate). Suppose that $g(x)$ in (2) is either strictly convex or strictly concave. Then, we have

$$x^* \neq \hat{x}.$$

Proof: Suppose the function $g(x)$ is strictly convex. From the stationarity of $X(t)$ and (10), note that

$$\begin{aligned} \hat{x} = E_\pi\{X(t+1)\} &= E\{f(X(t), U_t\} \\ &= E_\pi\{E_{U_t}\{f(X(t), U_t) \mid X(t)\}\} \\ &= E_\pi\left\{\int_0^1 f(X(t), u)\, du\right\}, \quad (11) \end{aligned}$$

where $f(x,u)$ is from (9), and (11) follows since $X(t)$ and U_t are independent. Since $g(x) = \int_0^1 f(x,u)\, du$ from (9) and (2), the previous equation implies

$$\hat{x} = E_\pi\{g(X(t)\} > g\left(E_\pi\{X(t)\}\right) = g(\hat{x}), \quad (12)$$

where the inequality follows from the (strict) convexity of g and Jensen's inequality. Since x^* is the unique fixed point of g, i.e., $x^* = g(x^*)$ (see (3)) and from (12), we see that $\hat{x}$ cannot be the fixed point of g, i.e., $\hat{x} \neq x^*$. The proof for a concave g is identical except that the inequality in (12) is reversed. ■

If $p(x)$ is given by (6), it is immediate to see that the function $g(x)$ as in (2) is always strictly concave. Thus, from Proposition 1, we have $x^* \neq E_\pi\{X(t)\}$. Fig. 1 shows the function $g(x)$ with $p(x)$ given by (6). Note that the fixed point x^* increases as B increases and the function $g(x)$ is always concave in this case. For all three values of B, the fluid model is stable, i.e., it satisfies (4).

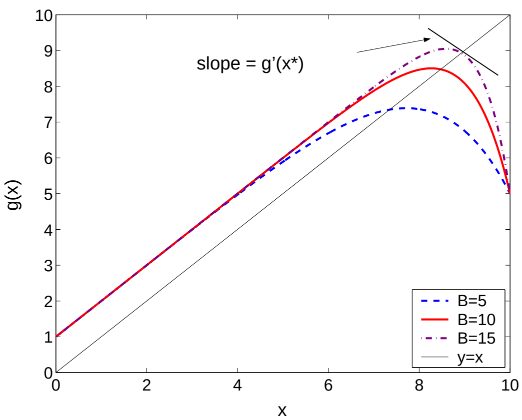

Fig. 1. $g(x)$ in (2) for a marking functions $p(x) = (x/C)^B$ with $C = 10$ and different buffer sizes B.

Tables I and II show the equilibrium points of the fluid model x^* and of the Markov chain $\hat{x} = E_\pi\{X(t)\}$ for two different functions $p(x)$, respectively. We fix $C = 10$ for all cases and set $\sigma^2 = 50$ only for $p_2(x)$ in Table II. All these values make the fluid model stable. As expected from Proposition 1, the two equilibrium points are quite different. In general, the higher the curvature of the function $g(x)$, the larger the difference between x^* and $\hat{x}$. For example, in case of $p(x) = (x/C)^B$ with $C = 10$ and $B = 15$ (for which the system is stable in the sense of (5)), the actual expected rate of the system can be about 17.7% smaller than what the fluid model predicts.

TABLE I
EQUILIBRIUM POINTS FOR $p(x) = p_1(x)$ IN (6)

	x^*	$\hat{x} = E_\pi\{X(t)\}$	$(x^* - \hat{x})/x^*$
B=5	7.34	6.30	14.1%
B=10	8.47	7.11	16%
B=15	8.93	7.35	17.7%

IV. Implication of the Stability of a System with Many Flows

In the previous section, we show that the equilibrium predicted by a fluid model can be quite different from the actual expected rate of the stochastic system. In this section we study the implication of the stability for these two different approaches and its impact on the choice of system parameters when there are many flows in the system.

TABLE II
EQUILIBRIUM POINTS FOR $p(x) = p_2(x)$ IN (7)

	x^*	$\hat{x} = E_\pi\{X(t)\}$	$(x^* - \hat{x})/x^*$
B=50	5.92	5.28	11%
B=100	7.23	6.29	13%
B=300	8.76	7.28	17%

A. *Tradeoff Between Link Utilization and Buffer Size in Fluid Models*

Consider now a single link with capacity NC shared by N flows, where the rate of each flow is governed by the same AIMD algorithm. Let $x_i(t)$ be the rate of flow i ($i \in \{1, 2, \ldots, N\}$) at time slot t in the fluid model. Assume that the congestion signal at the link is generated based only on the 'link utilization' of the link, i.e., $p(x)$ depends only on $\sum_{i=1}^N x_i(t)/N$. Note that the two examples in (6) and (7) satisfies this assumption. We then obtain a fluid model for flow i as follows:

$$\begin{aligned} x_i(t+1) &= (x_i(t)+1)\left(1 - p\left(\frac{\sum_{i=1}^N x_i(t)}{N}\right)\right) \\ &\quad + \frac{x_i(t)}{2} p\left(\frac{\sum_{i=1}^N x_i(t)}{N}\right) \end{aligned} \tag{13}$$

Define $y_N(t) := \sum_{i=1}^N x_i(t)/N$. Then, by summing (13) over i and dividing it by N yields

$$y_N(t+1) = (y_N(t)+1)(1 - p(y_N(t))) + \frac{y_N(t)p(y_N(t))}{2}. \tag{14}$$

Note that this is the same as the case of a single flow in (1). Hence, as far as the average rate (over flows) is concerned, the system with N flows preserves the same equilibrium point and stability property as the single flow case.

Consider again the two examples for $p(x)$ in (6) and (7). As shown in Fig. 1, it is not difficult to see that the stability condition for both cases becomes $B < \bar{B}$ for some finite $\bar{B}$ and we always have $y_N^* < C$, where y_N^* is the fixed point of (14), i.e., $g(y_N^*) = y_N^*$. (Note that $y_N^* = x^*$ since the fixed point is unique.) In other words, if the system in (13) with many flows were to be stable, then the buffer size B should be bounded (not too large), and the resulting link utilization becomes strictly less than one. If we arbitrarily increase B, then the equilibrium point y_N^* approaches to C (full utilization), but the fluid model becomes unstable, i.e., it never converges to y_N^*.

This type of observation between the link utilization and buffer size can also arise in other situations. Under the AQM scheme with a virtual queue [16], where we base the marking decision on the virtual queue whose capacity is strictly less than the real capacity, the real queue size remains bounded (very small), so the buffer requirement is minimal at the cost of reduced link utilization. Similarly, in [17], it is shown that starting from a queue-based marking where the marking happens at a fixed queue-length, the system dynamics becomes almost identical to a system with a rate-based marking scheme, for which we again obtain reduced utilization but bounded buffer-size requirement.

On the other hand, if we use queue-based AQMs (e.g., RED [18]) with suitably chosen marking functions with an appropriate buffer size, we can achieve both full utilization and stable system. Again, let there be N flows accessing a link with capacity NC and assume that the marking happens only based on the normalized queue-length. Specifically, let $q_N(t)$ be the actual queue-length at time slot t. Then, the normalized queue-length evolves as

$$\frac{q_N(t+1)}{N} = \left[\frac{q_N(t)}{N} + y_N(t) - C\right]^+,$$

where $y_N(t) = \frac{1}{N}\sum_{i=1}^N x_i(t)$. When the marking function $p(x)$ depends only on $q_N(t)/N$, it is shown that the system can be (locally) stable in the sense of the fluid model if $p(x)$ satisfies further conditions. In essence, the stability condition always requires that the slope of the marking probability at the equilibrium point be kept low. In case of queue-based AQMs with N flows and capacity NC, this means that the slope should be $O(1/N)$ [4], [13], indicating that the equilibrium queue-length and the buffer size should be at least $O(N)$. Note that this linear sizing for the buffer and the equilibrium queue-length ensures a 'stable' operation and full link utilization even when all the N flows drop their rates by half. As a result, for a system to be stable in a 'fluid' sense, we see that there exists a fundamental tradeoff between the link utilization and the buffer size requirement: (i) if the buffer size is bounded, i.e., $B = O(1)$, then the link utilization is also bounded away from one; (ii) if we need full link utilization, then the buffer size should be chosen as $B = O(N)$ where N is the number of flows accessing the link with capacity NC.

B. Stability of a Markov Chain with Many Flows

In this section we consider a stochastic version of the system via a Markov chain when there are N flows accessing a link with capacity NC. Let $X_i(t)$ be the random variable for the actual rate to the queue from flow i (or simply the window size or the number of packets) at time slot t. Suppose now that $X_i(t)$ is always bounded, i.e., $1 \le X_i(t) \le w_{max}$ where w_{max} is due to the receiver buffer space or bounded maximum rates at access networks. Let $X_i(t)$ for $i = 1, 2, \ldots, N$ be given. Then, based on the arguments as in Section II-B, we see that $X_i(t+1)$ will evolve as follows:

$$X_i(t+1) = \begin{cases} (X_i(t)+1) \wedge w_{max}, & \text{w.p. } 1 - p(Y_N(t)) \\ (X_i(t)/2) \vee 1, & \text{w.p. } p(Y_N(t)), \end{cases} \tag{15}$$

where $x \wedge y = \min\{x, y\}$ and $x \vee y = \max\{x, y\}$, and

$$Y_N(t) := \frac{1}{N}\sum_{i=1}^N X_i(t).$$

As before, we can interpret (15) as follows. Given $(X_i(t), \ldots, X_N(t)) = (x_1, \ldots, x_N)$, we see that the actual packet arrivals to the link will be a random process with parameter $\bar{x} = (x_1 + x_2 + \cdots + x_N)/N$. Depending on our choice for the arrival process, each flow will either decrease (or increase) with probability $p(\bar{x})$ (or $1 - p(\bar{x})$) independently from other flows. Thus, we obtain the a Markov chain as in (15) for $\vec{X}(t) = (X_1(t), \ldots, X_N(t))$, where $1 \le X_i(t) \le w_{max}$.

First, let N be fixed. Then, from our construction, $\vec{X}(t) = (X_1(t), \ldots, X_N(t))$ is an N-dimensional Markov chain with a finite number of states. Since the chain is also irreducible and aperiodic, it is always ergodic and converges in total variation to its stationary version [15]. This means that for any bounded function F, we have

$$\lim_{t\to\infty} E\{F(\vec{X}(t))\} = E_\pi\{F(\vec{X})\},$$

where $\vec{X} = (X_1, \ldots, X_N)$ denotes a stationary version of $\vec{X}(t)$ and π is the associated stationary distribution. Since $Y_N(t) \le w_{max}$, the above implies

$$\lim_{t\to\infty} E\{Y_N(t)\} = E_\pi\{Y_N\} = E_\pi\left\{\frac{1}{N}\sum_{i=1}^N X_i\right\} := \hat{y}_N.$$

In other words, the expected value of averaged rates (over N flows) converges to an equilibrium value. This result tells us that, given N, due to the stochastic nature of the system, the expected rate of the system *always* converges to the equilibrium point $\hat{y}_N$ regardless of initial distributions of the system. This is in stark contrast to the fluid model in which there exist some restrictive conditions on the system parameters for $y_N(t)$ to converges to the equilibrium point, which is given by the unique solution of $g(x) = x$.

This stochastic construction also allows much less restrictive parameter choice for a system with many flows. Similarly as above, since $p(\cdot)$ is bounded, we expect that $E\{p(Y_N(t)\}$ also converges to a steady-state probability $E_\pi\{p(Y_N)\}$ for any given N. To illustrate how this stochastic description affects the system performance and the choice of parameters, we take $p(x) = (x/C)^B$ as in (6) and assume that

$$0 < \alpha \le E_\pi\{p(Y_N)\} \le \beta < 1, \tag{16}$$

for all N, where $\alpha, \beta \in (0, 1)$ are constants. This assumption is reasonable since, in order for the system to be in a stationary regime, the steady-state marking probability $E_\pi\{p(Y_N)\}$ also needs to be bounded away from 0 and 1. Indeed, if this value is arbitrarily close to zero, then we expect that almost all the flows will increase their rates by one. This surely causes the distribution of the rates at the next RTT to be certainly different from the current one, thereby violating the stationarity condition.

Observe that Y_N is an average over N random variables X_i ($i = 1, 2, \ldots, N$). Thus, under a very general condition, we expect that Y_N will be very close to its mean due to the law of large numbers. Thus, for large N, we can write Y_N as

$$Y_N \approx \hat{y}_N + o(1),$$

where $\hat{y}_N = E_\pi\{Y_N\}$ and $\lim_{N\to\infty} o(1) = 0$. So, the condition in (16) with the choice of $p(x) = (x/C)^B$ becomes

$$\left(\frac{\hat{y}_N}{C} + o(1)\right)^B = \kappa(N), \tag{17}$$

where $\alpha \le \kappa(N) \le \beta$. Now, let $B = B(N)$, i.e., we want to scale the buffer size as the number of flows and the size of capacity grows. Then, we can rewrite (17) as

$$\hat{y}_N = C \cdot \left(\kappa(N)^{\frac{1}{B(N)}} - o(1)\right). \tag{18}$$

Since $\kappa(N) \ge \alpha > 0$ for all N, we see that $\hat{y}_N$ approaches to C as long as $\lim_{N\to\infty} B(N) = \infty$. In other words, we can have a 'stochastically stable' system with almost full utilization as long as the buffer size increases without bound as the system size (N) increases.

Recall that in the fluid-based approach there exists a fundamental tradeoff between the buffer size requirement $B(N)$ and the link utilization. In particular, we require that the buffer size increases linearly in N, i.e., $B(N) = O(N)$ for a stable system with full utilization, where the stability is in the sense of the convergence of $y_N(t)$ to y_N^* (see (14)). On the other hand, if we put fixed buffer size $B(N) = O(1)$ in a stable system, then the system utilization will be reduced and be strictly less than one. However, as we can see from (18), there is *no such tradeoff* between the utilization and the linearly increasing buffer size. Under the stochastic model, the system will achieve full utilization for $B(N) = O(N^p)$ with $0 < p < 1$ or even for $B(N) = \log N$, although the convergence can be very slow for such slowly increasing $B(N)$. This result is striking in that it offers a wide range of possibilities for system configuration without suffering "stability" issues. In fact, quite recently, it was shown [19] via real measurement that, under many TCP flows, the system utilization can be very high even when the buffer size is chosen as the bandwidth-delay product ($O(N)$) divided by $\sqrt{N}$, i.e., $B(N) = O(\sqrt{N})$. Further, more generally, we have also recently shown [20] that under a queue-based AQM, the system utilization approaches to 1 with low loss as N grows even when $B(N) = O(N^p)$ for any $p \in (0, 1)$, and verified this via ns-2 simulations.

V. Conclusion

The fluid-based modeling approach has proven to be extremely powerful and versatile for many problems in a network. For congestion control, the fluid model has been the main tool to derive any criterion for system stability around its equilibrium point. However, the actual behavior in a network is always stochastic, as there always exists inevitable randomness due to random packet arrivals as well as random marking/dropping at routers. In this paper, through simple analysis and numerical results, we maintain that there may exist some limitation on the fluid modeling approach from two distinct viewpoints. First, we have shown that the two equilibrium points can be quite different for a single flow system. In terms of the local stability of the fluid system, this implies that the system could have been linearized around an incorrect equilibrium point. Second, under many flows, we have also shown that the stability of the fluid model, which ensures the convergence of the average rate to its (possibly incorrect) equilibrium point, may impose excessive restriction of our choice of system parameters. In particular, under many flows with large capacity, we have shown that the system utilization still approaches to full utilization even when the buffer size is scaled far smaller than the bandwidth-delay product, while the buffer size on the order of bandwidth-delay product is necessary for the stability of the fluid model with full utilization. Our results in this paper thus generalize the new scaling laws for the buffer size that are recently reported in [19], [20] and also provide deep insight into the true meaning of a fluid model and its stability.

References

[1] F. P. Kelly, A. Maulloo, and D. Tan, "Rate control in communication networks: shadow prices, proportional fairness and stability," *Journal of the Operational Research Society*, pp. 237–252, 1998.

[2] V. Misra, W. Gong, and D. Towsley, "A Fluid-based Analysis of a Network of AQM Routers Supporting TCP Flows with an Application to RED," in *Proceedings of ACM SIGCOMM*, Stockholm, Sweden, September 2000.

[3] F. P. Kelly, "Models for a self-managed Internet," *Philosophical Transactions of the Royal Society A358*, pp. 2335–2348, 2000.

[4] S. H. Low, F. Paganini, J. Wang, and J. C. Doyle, "Linear stability of TCP/RED and a scalable control," *Computer Networks*, vol. 43, no. 5, pp. 633–647, Dec. 2003.

[5] S. Deb, S. Shakkottai, and R. Srikant, "Stability and Convergence of TCP-like Congestion Controllers in a Many-Flows Regime," in *Proceedings of IEEE INFOCOM*, San Francisco, CA, April 2003.

[6] S. H. Low and D. E. Lapsley, "Optimization Flow Control, I: Basic Algorithm and Convergence," *IEEE/ACM Transactions on Networking*, vol. 7, no. 6, pp. 861–875, Dec. 1999.

[7] H. Q. Ye, "Stability of Data Networks Under an Optimization-Based Bandwidth Allocation," *IEEE Transactions on Automatic Control*, vol. 48, no. 7, pp. 1238–1242, July 2003.

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

[9] V. Dumas, F. Guillemin, and P. Robert, "Limit results for Markovian models of TCP," in *Proceedings of IEEE GLOBECOM*, 2001.

[10] E. Altman, C. Barakat, and V. M. Ramos, "Analysis of AIMD protocols over paths with variable delay," in *Proceedings of IEEE INFOCOM*, Hong Kong, China, March 2004.

[11] F. Guillemin, P. Robert, and B. Zwart, "AIMD algorithms and exponential functionals," *The Annals of Applied Probability*, vol. 14, no. 1, pp. 90–117, 2004.

[12] P. Tinnakornsrisuphap and A. M. Makowski, "Limit Behavior of ECN/RED Gateways Under a Large Number of TCP Flows," in *Proceedings of IEEE INFOCOM*, San Francisco, CA, April 2003.

[13] S. Shakkottai and R. Srikant, "Mean FDE Models for Internet Congestion Control under a Many-Flows Regime," *IEEE Transactions on Information Theory*, vol. 50, no. 6, pp. 1050–1072, June 2004.

[14] P. Diaconis and D. Freedman, "Iterated Random Functions," *SIAM Review*, vol. 41, no. 1, pp. 45–76, 1999.

[15] P. Brémaud, *Markov Chains: Gibbs Fields, Monte Carlo Simulation, and Queues*, Springer-Verlag, 1999.

[16] S. S. Kunniyur and R. Srikant, "An adaptive virtual queue (AVQ) Algorithm for Active Queue Management," *IEEE/ACM Transactions on Networking*, vol. 12, no. 2, pp. 286–299, April 2004.

[17] S. Deb and R. Srikant, "Rate-Based versus Queue-Based Models of Congestion Control," in *Proceedings of ACM SIGMETRICS*, June 2004.

[18] S. Floyd and V. Jacobson, "Random Early Detection gateways for Congestion Avoidance," *IEEE/ACM Transactions on Networking*, vol. 1, no. 4, pp. 397–413, Aug. 1993.

[19] G. Appenzeller, I. Keslassy, and N. McKeown, "Sizing Router Buffers," in *Proceedings of ACM SIGCOMM*, Portland, OR, 2004.

[20] D. Y. Eun and X. Wang, "Stationary Behavior of TCP/AQM with Many Flows under Aggressive Packet Marking," in *IEEE International Conference on Communication*, Seoul, Korea, 2005.