Application Layer Based Flow Control in Wireless Networks: Optimality and Semi-Global Exponential Stability of the Equilibrium

Minghua Chen and Avideh Zakhor

***Abstract*— Flow control is an important problem in wireless networks from both theoretical and practical viewpoints. Specifically, transmitting data across networks without a properly designed control could potentially result in serious performance degradations in terms of both throughput and delay. The problem of flow control in wired networks has been widely and successfully addressed, making use of the assumption that packet loss is a synonym of congestion; this assumption is not directly applicable to wireless networks in which the bulk of the packet loss is due to wireless channel errors. In previous work, we have proposed a practical end-to-end solution to the flow control in wireless networks, by opening an appropriate multiple connections between sender and receiver. Specifically, we modeled the problem as a nonlinear singularly perturbed system, and showed uniqueness and local stability of its equilibrium. In addition, we have shown the promising performance of our previously proposed solution via simulations and actual experiments. In this paper, we take one step further to show that the equilibrium is the optimal solution of a convex optimization problem. This interpretation is general in the sense that it encompasses other existing formulations for wired networks in the literature. Finally, we further generalize the analysis by showing that the equilibrium is semi-globally exponentially stable.**

I. INTRODUCTION

Transmission Control Protocol (TCP) has recently been the focus of much theoretical and practical research. TCP has been modeled via continuous time differential equations that describe the evolution of the rates of a set of users exchanging information over a network. This is an instance of fluid flow model. The study of this model enables a better understanding of the intrinsic characteristics of the system. Investigating the properties of this scheme has nevertheless proven to be a rather challenging task; this is due to the presence of strong non linearities in the functions that come into play, as well as of the distributed nature of the scheme; moreover, the multiple couplings between its entities, i.e. senders, receivers and links, hampers the global understanding of its behavior.

The current fluid flow models for TCP have only been dealing with the case of wired networks, [1] [2]. Fundamental properties such as stability have been studied [3], and conditions for achieving robustness to disturbances [4] and to delays have been introduced [5].

Recently some researchers have turned their attention to flow control over wireless (see [6] and references of it). This new setting poses new, unexpected challenges, due to the presence of channel noise at the link level. We have recently proposed a new flow control scheme called E-MULTFRC [7] for video streaming over wireless networks. Specifically, we have intended a corresponding continuous-time model showing its properties, such as existence of a unique equilibrium and local exponential stability [8].

In this paper, we show semi-global exponential stability of the equilibrium. By semi-global exponential stability, we mean the equilibrium is exponentially stable as long as variables are constrained to a large enough compact set, which shall be trivially satisfied in practice. Our results also indicate the stability, optimality and scalability of proposed solution coexisting with TCP-like schemes. This is useful in incremental deployment of our solution in the current Internet where TCP-like schemes are dominant.

The paper is structured as follows: after a concise introduction to the TCP-like scheme for wired networks in Section II, we propose our solution abstracted from E-MULTFRC in Section III. A series of facts will elucidate the existence and uniqueness of the equilibrium, as well as its local exponential stability, also in Section III. Then in Section IV, we prove the optimality and semi-global exponential stability for the scheme. Explanations of the implications of these results will follow, and a discussion of its practical implementation in Section V will close the paper.

II. PROBLEM FORMULATION

In this section we first introduce the dynamical model of the well-known general flow control problem first introduced by Kelly et. al., [1]. Starting from the wired scenario, we motivate and extend to the more challenging wireless case.

A. Wired Networks

A communication network is described via its J *links* and its R *users*, i.e. sender-receiver pairs. Each $j \in J$ has a finite capacity $C_j < \infty$. The network interconnections are described via a routing matrix $A = (a_{jr}, j \in J, r \in R)$, where $a_{jr} = 1$ if $j \in r$. About a decade ago, researchers developed a fluid-flow, continuous-time model for TCP-like scheme [1]. From this perspective, flow control can be regarded as a dynamical system model, and as such, prone to be quantitatively analyzed. Assume that to each user a sending rate $x_r \geq 0$ and a utility function $U_r(x_r)$ are associated [1], [9]. $U_r(x_r)$ is assumed to be increasing, strictly concave and $\mathcal{C}^1$. Thus, the setting can also be interpreted as a concave maximization problem [9] [2], dependent on the aggregate utility functions for the rates and on costs on the links,

Minghua Chen and Avideh Zakhor are with Department of Electrical Engineering and Computer Sciences, University of California at Berkeley, Berkeley, CA 94720 {minghua, avz}@eecs.berkeley.edu

denoted by $P_j(\cdot)$:

$$\max \sum_{r\in R} U_r(x_r) - \sum_{j\in J} P_j\left(y_j(t)\right), \tag{1}$$

where $y_j(t) = \sum_{s:j\in s} x_s(t)$ is aggregate rates through link j, and the cost functions $P_j(\cdot)$ are defined as:

$$P_j(\upsilon) = \int_0^{\upsilon} p_j(\delta)\, d\delta. \tag{2}$$

The terms $p_j(\cdot)$'s are the prices at the links, and are assumed to be non-negative, continuous, and increasing functions; they represent a congestion measure and can be inferred from their structure, they have a local dependence on the aggregate rate passing through the link they are indexed by. As is common in the literature [2], in this paper we choose the price to correspond to "packet loss rate", defined as:

$$p_j(\lambda) = \frac{(\lambda - C_j)^+}{\lambda}. \tag{3}$$

Intuitively, the quantities $C_j < \infty$ represent the capacity of the link j, i.e. the maximum load it can accommodate without getting congested. User r will suffer a packet loss rate, which can be additively approximated as the quantity $\sum_{j\in r} p_j(y_j(t))$, under the assumptions of small $p_j(\cdot)$'s. The classical rate control scheme that we use has the following shape:

$$\dot{x}_r(t) = k_r\left(w_r^o - x_r(t)\sum_{j\in r} p_j(y_j(t))\right), r\in R \tag{4}$$

with k_r being a positive scale factor affecting the adaptation rate, and the constant w_r^o being interpreted as the number of connections that the user establishes within the network; the congestion signal depends on the sum of the prices, $p_j(y_j(t))$, along all the links that are crossed by the user. Interpreting the model in (4) as a dynamical system, it is easy to express its equilibrium implicitly. In [1] Lyapunov arguments are used to show that this equilibrium is unique and asymptotically stable. Kunniyur and Srikant [10] show the equilibrium is semi-globally exponential stable, i.e. it is exponentially stable as long as $x(t)$ is constrained to lie in a compact set. Moreover under certain condition, the schemes can be endowed with many interesting properties such as delay sensitivity and robustness [1].

B. Wireless Networks

Wireless channels are affected by errors, due to the fading and noisy nature of the physical channel. This physical channel loss affects packet loss within TCP-like schemes. We need to encompass this in the model by introducing a new price function $q_j(\cdot)$ for link j as follows:

$$q_j\left(\sum_{s:j\in s} x_s(t)\right) \triangleq p_j(y_j(t)) + \epsilon_j \geq p_j\left(\sum_{s:j\in s} x_s(t)\right). \tag{5}$$

This function takes into account both the congestion measure, i.e. $p_j(\cdot)$, and the channel error ϵ_j. The TCP-like model in (4) will now depend upon this new function $q_j(\cdot)$ for wireless scenarios. It is again straightforward to determine the equilibrium of this new dynamic system, which is different than the one derived from (4). Interpreting this fact through the underlying optimization problem, shown in (1), we can show that the new equilibrium is suboptimal, resulting in underutilization of the wireless bandwidth in practice. This fact has motivated us to investigate wireless flow control problem in the next Section.

III. A New Control Scheme for Wireless Networks

A. A Practical Flow Control Scheme with Indicator Function

In [11] we introduced a end-to-end practical solution for flow control in wireless networks. Specifically, we change w_r^o in (4 to $w_r(t)$, in order to arrive at the following dynamic system:

$$\begin{cases} \dot{x}_r(t) = k_r\left(w_r(t) - x_r(t)\sum_{j\in r} p_j(y_j(t))\right), r\in R \\ \dot{w}_r(t) = c_r\left(w_r^o - w_r(t) I\left(\sum_{j\in r} p_j(y_j(t))\right)\right), r\in R \end{cases} \tag{6}$$

where c_r is a constant, and $I(\sum_{j\in r} p_j(y_j(t)))$ is an indicator function implying the congestion status of route r:

$$I\left(\sum_{j\in r} p_j(y_j(t))\right) = I\left(\sum_{j\in r} \frac{(y_j(t) - C_j)^+}{y_j(t)}\right) \tag{7}$$

$$= \begin{cases} 1, & \text{if route } r \text{ is congested at time } t, \\ & \text{i.e. } \sum_{j\in r} \frac{(y_j(t)-C_j)^+}{y_j(t)} > 0; \\ 0, & \text{otherwise.} \end{cases}$$

We can interpret the dynamics of $w_r(t)$ as the adjustment of the number of connections that the user opens within the network. The control law of the aggregate rates for user r, i.e. $x_r(t)$, can be understood as the sum of rates of $w_r(t)$ individual, and parallel connections, each of which being controlled using the standard TCP-like algorithm. This interpretation is attractive in that the approach requires little, if any modifications to TCP-like protocol in practice.

The above scheme can be easily implemented in practice by adjusting the number of connections which an application opens in an actual network. Therefore, it is an *application layer*-based approach, and it is easy to deploy, since it does not require modification to the network's infrastructure, such as base stations and routers, or network protocols.

At another level, the system tries to achieve full utilization by adjusting $w_r(t)$ accordingly. In particular,

- If a route r is underutilized, then $I(\sum_{j\in r} p_j(y_j(t))) = 0$; this implies that $w_r(t)$ will increase in order to boost the user's rate $x_r(t)$, pursuing full utilization on any route r;
- If the route r is fully utilized, i.e. one of its links is congested, then $I(\sum_{j\in r} p_j(y_j(t))) = 1$, lowering $w_r(t)$, and hence $x_r(t)$, to prevent the system from further congestion.

The intuition behind our approach is as follows: when loss rate caused by channel error increases, individual connection's sending rate is lowered, making the users open more connections to increase the aggregate throughput. The $I(\sum_{j\in r} p_j(y_j(t)))$ is the one bit of information required from the end-to-end measurements. In practice, we can estimate $I_r(\cdot)$ using end-to-end round trip time measurements. In particular, we estimate the queuing delay by comparing current round trip time with the propagation delay, and set $I(\sum_{j\in r} p_j(y_j(t))) = 1$ if the queuing delay is cumulating, and $I(\sum_{j\in r} p_j(y_j(t))) = 0$ otherwise.

B. Continuous Approximations of the System and the Two Time Scales Assumption

The discontinuities of the indicator functions $I_r[\sum_{j\in r} p_j(y_j(t))]$ and the non-smoothness of $p_j(y_j(t))$ hinder the analysis the scheme, especially the study of the equilibria and their stability. To carry out the analysis, we approximate these discontinuous functions using continuous ones. Specifically, $\forall j \in J, r \in R$, we introduce the following approximations:

$$p_j(y_j(t)) \approx \frac{1}{\beta} \ln\left(1 + e^{\beta\frac{y_j(t)-C_j}{y_j(t)}}\right) \triangleq g_j(y_j(t)), \quad (8)$$

$$I\left(\sum_{j\in r} p_j(y_j(t))\right) \approx \frac{e^{\beta\sum_{j\in r} g_j(y_j(t))} - 1}{e^{\beta\sum_{j\in r} g_j(y_j(t))} + 1} \triangleq f(\sum_{j\in r} g_j(y_j(t)), \quad (9)$$

where β is a positive constant. It should be clear that $f(\sum_{j\in r} g_j(y_j(t)) \rightarrow I(\sum_{j\in r} g_j(y_j(t))$ and $g_j(y_j(t)) \rightarrow p_j(y_j(t))$ as $\beta \rightarrow \infty$.

Thus, the corresponding approximated continuous version of original system in (6) is as follows: $\forall r \in R$,

$$\begin{cases} \dot{x}_r(t) = k_r\left[w_r(t) - x_r(t)\sum_{j\in r}(\epsilon_j + g_j(y_j(t)))\right]; \\ \dot{w}_r(t) = c_r\left[w_r^o - w_r(t) f(\sum_{j\in r} g_j(y_j(t))\right]. \end{cases} \quad (10)$$

Since the approximated system in (10) is continuous, we can then analyze its equilibrium and stability for arbitrary values of β. As $\beta \rightarrow \infty$, the system in (10) approaches the original system in (6); therefore by analyzing the properties of the system (10), and letting $\beta \rightarrow \infty$, we expect to understand the behavior of the system in (6).

Even though the approximated system in (10) is continuous, it is difficult to analyze in general. Specifically, it is a nonlinear, coupled, multivariable system, and the two equations are not exactly symmetrical even though they might appear to be so. Hence we make an important assumption to facilitate the analysis, which in practice turns out to easy to satisfy.

In [12] we argue that in actual TCP schemes, the rate of change of $w(t)$ representing the number of connections that a user opens, is dimensionally slower than that of $x(t)$, representing the source sending rate. Therefore, inspired by a wealth of control literature on singular perturbation systems [13], we carefully make the following key assumption: *the dynamics corresponding to $x(t)$ and $w(t)$ evolve in two different time scales; the former in a faster timescale, while the latter in a slower one.*

The two time scale assumption applied to the approximated system in (10) highlights two kinds of dynamics: a fast one, which is described by a *boundary-layer* system, and a slow one, which is characterized by the *reduced-order* system. The fast interconnection is modeled as:

$$\begin{cases} \dot{x}_r(t) = k_r\left[w_r(t) - x_r(t)\sum_{j\in r}(\epsilon_j + g_j(y_j(t)))\right], \\ w_r(t) = \text{constant}, \end{cases} \quad (11)$$

$\forall r \in R$. For the slower timescale we instead have the following dynamics:

$$\begin{cases} x_r(t) = \frac{w_r(t)}{\sum_{j\in r}[\epsilon_j + g_j(y_j(t))]}, \\ \dot{w}_r(t) = c\left(w_r^o - w_r(t) f(\sum_{j\in r} g_j(y_j(t))\right) \end{cases} \quad (12)$$

$\forall r \in R$. Under the two time scale framework, the behavior of the system can be described as follows. On the fast timescale, $w_r(t)$ is assumed to be constant, and the entire system evolves according to Kelly's model for wired network as expressed in (4), other than the fact that w_r^o is replaced by $w_r(t)$ and the price function $p_j(y_j(t))$ is replaced by $\sum_{j\in r}(\epsilon_j + g_j(y(t))$. The behavior of the boundary layer system can be easily inferred from the known results about the model in (4). Specifically, it has a semi-globally exponentially stable unique equilibrium, which is a function of $w_r(t)$ [10]. A first conclusion is that, on the fast timescale, $x_r(t)$ converges to the equilibrium manifold defined as follows:

$$x_r(t) = \frac{w_r(t)}{\sum_{j\in r}[\epsilon_j + g_j(y_j(t))]}, \quad r \in R. \quad (13)$$

On the slow timescale, on the other hand, the rates x_r have already converged to the equilibrium manifold in (13), and the system can be reformulated as the reduced order model described in (12). Its behavior determines the way the approximated system evolves in the long run. Together with the boundary layer system, this reduced order system fully characterizes the behavior of the complete model. Motivated by the above considerations, we shall mainly focus on investigating properties of the reduced system in (12) to infer those of the full model, as discussed in the following Section.

C. Existence, Uniqueness and Local Exponential Stability of Equilibria

The following theorem originally developed in [8] shows the system in (10) has a unique equilibrium

Theorem 1: For arbitrary $\beta > 0$, the approximated system (10) has one unique equilibrium.

Proof: Please refer to [8]. ■

Given the existence of a unique equilibrium for the system in (10), our previous work shows that it is in fact locally exponentially stable by examining the stability of its linearization around the equilibrium [8]. This is contained in the following theorem:

Theorem 2: The unique equilibrium (x^*, w^*) of system in (10) is locally exponentially stable, for arbitrary $\beta > 0$.
Proof: Refer to [8]. ■

IV. Optimality and Semi-Global Exponential Stability of The Equilibrium

In this section we take further efforts to show that the equilibrium of system in (10) is optimal in the sense of solving a concave utility optimization problem. Furthermore, it is fact semi-globally exponentially stable, i.e. it is exponentially stable as long as $w(t)$ and $x(t)$ are constrained in a large enough compact set, which is defined in the proof of Theorem 4. This is trivial to satisfy in practice, as argued in the proof of Lemma 1 in Appendix A.

The following theorem states that the unique equilibrium solves a concave utility optimization problem similar to (1).

Theorem 3: For any arbitrary $\beta > 0$, the unique equilibrium of the approximate system in (10), denoted by (x^*, w^*), solves the following concave optimization problem

$$\max_{x \geq 0} \sum_{r \in R} U_r(x_r) - \sum_{j \in J} \int_0^{y_j} g_j(z)\, dz, \tag{14}$$

with $U_r(\cdot), r \in R$ being the concave function:

$$U_r(x_r) = \int_0^{x_r} h_r^{-1}\left(\frac{w_r^o}{\nu}\right) d\nu, \quad r \in R,$$

and where $h_r^{-1}(\cdot), r \in R$ is the inverse of the monotonically increasing function h_r:

$$h_r(z) \triangleq \left(\sum_{j \in r} \epsilon_j + z\right) f(z) = \left(\sum_{j \in r} \epsilon_j + z\right) \frac{e^{\beta z} - 1}{e^{\beta z} + 1}.$$

Proof: First it is easy to see the net utility function in (14) is concave. Then the claim follows by setting to zero the derivative of the net utility function with respective to x. ■

The first observation about Theorem (3) is as follows: the unique equilibrium for the system in (10) in the wireless scenario solves a concave optimization problem which is of the same form as the general one in Eq. (1). It is similar to the one for the wired networks, proposed by Kelly et. al.'s [1], but with different utility functions $U_r(x_r)$ for each user. More precisely, while the $U_r(x_r)$ in the wired network case is only a function of x_r, in wireless scenario it is also a function of $\sum_{j \in r} \epsilon_j$, i.e. the wireless packet loss rate associated with route r. If we let $\beta \to \infty$ and $\epsilon_j = 0, \forall j \in J$, the problem is reduced to the wired network scenario and we obtain $h_r(z) = z$. In this case, the equilibrium (x^*, w^*) is exactly the same as (x^o, w^o), implying the optimization problem in the wired network is merely a special case of that in Theorem 3.

We state three lemmas that are needed for proving the global exponential stability. The first lemma explores an interesting structure of the vector field of the boundary layer system and the reduced system.

Lemma 1: There exists a compact set, denoted by Ω_1 for $w(t)$ in the reduced system in (12) with arbitrary $\beta > 0$, such that any compact set containing it is a positively invariant one. The same observation is also true for $x(t)$ in the boundary layer system (11), and the corresponding compact set is defined as $\Omega_2(w)$, a function of w.
Proof: Refer to Appendix A. ■

The second lemma investigates the global asymptotical stability of the equilibrium in the reduced system. The particular non-linear shape of the vector field for $\dot{w}_r(t)$ shown in (12) makes the search for a suitable Lyapunov function, or a function on which to apply the La Salle principle, a challenging task. We have found that none of the techniques applied in [1] and [12] work in this case. We therefore believe that our solution may provide insight to searching Lyapunov functions, or functions for the La Salle principle, in similar cases.

Lemma 2: The unique equilibrium of reduced layer system in (12), with arbitrary $\beta > 0$, is a globally asymptotically stable one.
Proof: Refer to Appendix B. ■

The third lemma states that for continuous systems, local exponential stability and global asymptotical stability is equivalent to semi-global exponential stability. This lemma is quite general and as such, its use is not restricted to the use in particular problem discussed in this paper.

Lemma 3: Consider a system $\dot{\xi} = \varphi(\xi, t, \xi_0)$ satisfying the following assumptions:

- it has a unique equilibrium at 0 that is locally exponentially stable and globally asymptotically stable;
- $\varphi(\xi, t, \xi_0)$ is continuous.

Then the equilibrium of the system is semi-globally exponentially stable.
Proof: Refer to Appendix C. ■

These three lemmas enable us to assert the semi-global exponential stability of the equilibrium of the system in (10), as follows:

Theorem 4: The unique equilibrium of singularly perturbed system in (10) with arbitrary $\beta > 0$ is semi-globally exponentially stable.
Proof: Refer to Appendix D. ■

Remarks: Note that the value of constant c_r can be chosen arbitrarily in system (10), as long as the two timescale decomposition holds true. Practically, this implies that each user can adjust $w_r(t)$ according to a different rate. A global setting among all the users, as assumed in [11], is not necessary. Furthermore, allowing some of the c_r to be equal to zero represents a scenario according to which the proposed scheme coexists with TCP-like schemes. In this situation, all the results including lemmas and theorems still hold, except for a modification to Theorem **??**. More precisely, the utility for users applying TCP-like schemes should be $U_r(x_r) = w_r^o \log x_r$, rather than the one defined in the theorem.

Our results imply that in a network where our proposed schemes coexist with TCP-like approaches, all users' rates will again converge, semi-globally exponentially, to a unique optimal equilibrium. These observations on stability and scalability facilitates incremental deployment of our scheme in the current Internet where TCP-like schemes are dominant, and directly addresses a major concern from the networking point of view.

V. Discussions

An actual implementation of the proposed scheme in (6) requires discretization of continuous quantities such as time and space. For instance, controlling $w_r(t)$ is implemented by adjusting the number of connections, which has to be an integer number; controlling $x_r(t)$ is implemented by adjusting the number of packets of finite size to be sent out in a time interval. Therefore, it is very unlikely for the system to operate at the points of discontinuity or non-smoothness in the model. From this point of view, the analysis based on the approximated system in (10), with a very large $\beta > 0$, is sufficient to predict and interpret the performance of the actual implementation of the algorithm. Our efforts in [11] have demonstrated the feasibility of proposed scheme and estimation of the indicator function (7) in practice.

Theorem 4 only asserts the exponential stability of the equilibrium when $(w(t), x(t))$ is constrained to a large enough compact set. However, it is sufficient to guarantee and predict proposed scheme's convergence performance in practice, since the number of connections and sending rates are always constrained in practice, and the "large enough" requirement is trivial to satisfy.

Acknowledgments

The authors would like to express their thanks to Alessandro Abate, Claire J. Tomlin, and Shankar Sastry for their helpful discussion and help in preparing this paper.

Appendix

A. Proof of Lemma 1

Proof: To construct the compact set Ω_1 for $w_r(t)$ in the reduced system, first note

a) if route r is not congested, then $g_j(y_j(t)) \leq \frac{1}{\beta}\ln 2$; so $x_r(t) \geq \frac{w_r(t)}{\sum_{j\in r}(\epsilon_j + \frac{1}{\beta}\ln 2)}$. As $w_r(t)$ increases, $x_r(t)$ will eventually hit $\min_{j\in r} C_j$ and route r becomes congested (cross traffic only helps to cause congestion). Hence, if $w_r(t)$ is sufficiently large, route r will be congested.

b) if route r is congested, at least at one link j along the route, the aggregate arriving rate $y_j(t)$ must exceed its capacity C_j, therefore $f(\sum_{j\in r} g_j(y_j(t)) \geq 2\frac{e^{\ln 2}}{1+e^{\ln 2}} - 1 = \frac{1}{3}$. Hence $w_r^o - w_r(t) f(\sum_{j\in r} g_j(y_j(t)) < w_r^o - \frac{1}{3}w_r(t)$ as long as route r is congested.

Therefore, there exists large enough w_r^{max} such that route r is congested regardless of cross traffic, i.e. even only user r is active and others are inactive, and $w_r^o - w_r^{max} f(\sum_{j\in r} g_j(y_j(t)) < w_r^o - \frac{1}{3}w_r^{max} \leq 0$. Eventually, Ω_1 can be defined as

$$\Omega_1 = [0, w_1^{max}] \times [0, w_2^{max}] \cdots [0, w_N^{max}].$$

It is easy to check the flow of the vector field satisfies:

- if $w_r(t) \leq 0$, $\dot{w}_r(t) > 0$ according to (12);
- if $w_r(t) \geq w_r^{max}$, then by w_r^{max}'s definition, $\dot{w}_r(t) \leq 0$.

Therefore, on the boundary of any compact set containing Ω_1, vector field defined in (12) points inward. Hence by definition, the compact set is positive invariant.

Similarly, a compact set $\Omega_2(w)$ can be defined for $x(t)$ in boundary layer system in (11), as follows:

$$\Omega_2(w) = [0, x_1^{max}(w_1)] \times [0, x_2^{max}(w_2)] \cdots [0, x_N^{max}(w_N)],$$

where $x_r^{max}(w_r) = \frac{w_r}{\sum_{j\in r}(\epsilon_j + g_j(x_r))}$ satisfies that given w_r and regardless of cross traffics along route r. If $x_r(t) \geq x_r^{max}(w_r)$, then $\dot{x}_r(t) \leq 0$.

From networking point of view, containing Ω_1 implies the number of connections of each user should be allowed to take large enough values to fully utilize his bottleneck, even if only he is active. Similarly, containing $\Omega_2(w)$ means sending rate of each user can be sufficiently large to achieve its equilibrium. These requirements are trivial to meet in practice. ■

B. Proof of Lemma 2

Proof: First by Lemma 1 in [8], for reduced system,

$$\dot{w} = diag(x) D\, \dot{x}, \tag{15}$$

where

$$D = \left[diag(\sum_{j\in r} \frac{\epsilon_j + g_j(y_j)}{x_r}) + A^T diag(g_j'(y_j))A\right]$$

is positive definite. From Appendix B in [8], we also know $E \triangleq diag(x_r^2/w_r) - D^{-1}$ is positive semi-definite.

Define $z_r = \frac{w_r}{x_r}, r \in R$, we have $\sum_{j\in r} g_j(y_j) = z_r - \sum_{j\in r}\epsilon_j$ and

$$\begin{aligned} \dot{z} &= diag(\frac{1}{x_r})\dot{w} - diag(\frac{w_r}{x_r^2}) D^{-1} diag(\frac{1}{x_r})\dot{w} \\ &= diag(z_r) diag(\frac{1}{x_r}) E\, diag(\frac{1}{x_r})\dot{w} \end{aligned} \tag{16}$$

Given initial condition $w(0)$, let Σ be the minimal connected compact set containing $w(0)$ and the invariant set Ω_1 defined in Lemma 1. Define *monotonically increasing* functions $\phi_r(\lambda), r \in R$ as follows:

$$\phi_r(\lambda) = \int_{\sum_{j\in r}\epsilon_j}^{f^{-1}(\lambda)+\sum_{j\in r}\epsilon_j} \frac{f(y - \sum_{j\in r}\epsilon_j)}{y} dy \tag{17}$$

Define a continuous function over Σ as follows:

$$\begin{aligned} V(z, w) &= -\sum_{r\in R} c_r \Bigg(\ln(z_r) + \int_{\sum_{j\in r}\epsilon_j}^{z_r} \frac{w_r}{y} f(y - \sum_{j\in r}\epsilon_j) \\ &\quad - \int_0^{w_r} \phi_r(\frac{1}{\lambda}) d\lambda \Bigg) \end{aligned} \tag{18}$$

Its Lee derivative is

$$\begin{aligned}\dot{V}(z,w) &= -\sum_{r\in R} c_r\left(\frac{1}{z_r}-\frac{w_r}{z_r}f(z_r-\sum_{j\in r}\epsilon_j)\right)\dot{z}_r - \\ &\quad \sum_{r\in R} c_r\left(\phi_r(1/w_r)-\phi_r(f(z_r-\sum_{r\in R}\epsilon_j))\right)\dot{w}_r \\ &= -\dot{w}^T\, diag(\frac{1}{x_r})E\, diag(\frac{1}{x_r})\dot{w} - \sum_{r\in R} \\ &\quad c_r\left(\phi_r(\frac{1}{w_r})-\phi_r(f(z_r-\sum_{r\in R}\epsilon_j))\right)\dot{w}_r \qquad (19) \\ &\leq 0,\end{aligned}$$

in which equality is taken if and only if $\dot{w} = 0$, i.e. at the equilibrium. In above derivation, we use the fact $\phi_r(\lambda), r \in R$ are monotonically increasing functions and hence the second part in (19) is non-positive.

Hence, we can apply La Salle principle [13] to conclude the trajectory starting from $w(0)$ converges to the equilibrium asymptotically. As $w(0)$ is arbitrarily, we conclude the equilibrium is globally asymptotically stable. ∎

C. Proof of Lemma 3

Proof: By definition of local exponential stability, there exists a $r > 0$ s.t.

$$\|\xi(t,\xi_0)\| \leq K\|\xi_0\|e^{-\gamma t}, \quad \forall \|\xi_0\| \leq r,$$

where K, r and $\gamma > 0$ are constant. Without loss of generality, we assume $K > 1$ and $r < 1$.

Define $T_r(\xi_0) = \inf\{t \geq 0 : \|\xi(t,\xi_0)\| \leq r\}$. By definition of global asymptotical stability, $T_r(\xi_0) < \infty$.

Since $\xi(t,\xi_0)$ is continuous with respect to t, let $M_r(\xi_0) = \max\{K, \xi(t,\xi_0) : 0 \leq t \leq T_r\} < \infty$.

Define $L = M_r(\xi_0)e^{\gamma T_r(\xi_0)}/r$, we claim

$$\|\xi(t,\xi_0)\| \leq L\|\xi_0\|e^{-\gamma t}. \qquad (20)$$

To see this, we study two cases:

- if $\|\xi_0\| \leq r$, then $T_r(\xi_0) = 0$, by setting, we already have $\|\xi(t,\xi_0)\| \leq K\|\xi_0\|e^{-\gamma t} \leq L\|\xi_0\|e^{-\gamma t}$;
- if $\|\xi_0\| > r$, the following is true for $t \in [0, T_r(\xi_0)]$:

$$\begin{aligned}L\|\xi_0\|e^{-\gamma t} &= M_r(\xi_0)\frac{\|\xi_0\|}{r}e^{\gamma T_r(\xi_0)}e^{-\gamma t} \\ &\geq M_r(\xi_0) \geq \|\xi(t,\xi_0)\|.\end{aligned}$$

  For $t \geq T_r(\xi_0)$, let $t' = t - T_r$, we have $\|\xi(t'=0)\| \leq r$. Applying the first case's result concludes our claim.

Therefore, by Definition 5.10 in [13], (20) implies semi-global exponential stability of the equilibrium. The semi-global term is due to L's dependency on initial condition ξ_0. ∎

D. Proof of Theorem 4

Proof: We have shown that the reduced system has a unique equilibrium (Theorem 3) that is globally asymptotically stable (Lemma 2) and locally exponentially stable (Theorem 2). Hence by Lemma 3, we conclude the unique equilibrium of reduced system in (12) is semi-globally exponentially stable.

Similar arguments are also true for boundary layer system. It has a unique equilibrium that is globally asymptotically stable and locally exponentially stable [1]. Hence, again by Lemma 3, we conclude the unique equilibrium of boundary layer system in (12) is semi-globally exponentially stable.

If we constrain $w(t)$ to a compact set containing Ω_1, denoted by Σ_1 and let $w_r^{max} = \max(w_r : w_r \in \Sigma_1)$, then any compact set containing $\Omega_2(w^{max})$ is positive invariant, following the arguments in Appendix A. Constrain $x(t)$ to a compact set containing $\Omega_2(w^{max})$, denoted by Σ_2.

On $\Sigma_1 \times \Sigma_2$, the equilibrium of the reduced system in (12) is exponentially stable, and the one of the boundary layer system in (11) is exponential stable uniformly in w (verification similar to in Appendix II.F in [10]).

Together with the fact $\Sigma_1 \times \Sigma_2$ is a positive invariant, by Theorem 11.4 in [14], we conclude that the singularly perturbed system in (10) has a unique equilibrium that is exponentially stable in $\Sigma_1 \times \Sigma_2$. Hence, the equilibrium of singularly perturbed system in (10) is semi-globally exponentially stable. ∎

REFERENCES

[1] F. P. Kelly, A. Maulloo, and D. Tan, "Rate control for communication networks: shadow prices, proportional fairness, and stability," *Journal of the Operationl Research Society*, vol. 49, pp. 237–252, 1998.

[2] F. P. Kelly, "Fairness and stability of end-to-end congestion control," *European Journal of Control*, vol. 9, pp. 159–176, 2003.

[3] F. Paganini, Z. Wang, J. Doyle, and S. Low, "A new TCP/AQM for stable operation in fast networks," in *Proc. IEEE INFOCOM*, San Francisco, CA, Mar. 2003.

[4] R. Johari and D. Tan, "End-to-end congestion control for the internet: delays and stability," *IEEE/ACM Trans. Networking*, vol. 9, no. 6, pp. 818–832, Dec. 2001.

[5] G. Vinnicombe, "On the stability of end-to-end congestion control for the internet," University of Cambridge, Cambridge, UK, Tech. Rep. CUED/F-INFENG/TR.398, 2001.

[6] M. Chen and A. Zakhor, "Rate control for streaming video over wireless," *IEEE Wireless Commun. Mag.*, vol. 12, no. 4, Aug. 2005.

[7] ——, "Flow control over wireless network and application layer implementation," in *Proc. IEEE INFOCOM*, Barcelona, Apr. 2006, to appear.

[8] M. Chen, A. Abate, and S. Sastry, "Analysis of an implementable application layer scheme for flow control over wireless networks," in *Proceeding of 17th International Symposium on Mathematical Theory of Networks and Systems*, Kyoto, Japan, 2006.

[9] S. H. Low and D. E. Lapsley, "Optimization flow control, i: Basic algorithm and convergence," *IEEE/ACM Trans. Networking*, vol. 7, no. 6, pp. 861–875, Dec. 1999.

[10] S. Kunniyur and R. Srikant, "A time scale decomposition approach to adaptive ecn marking," *IEEE Trans. Automat. Contr.*, vol. 47, no. 6, pp. 882–894, June 2002.

[11] M. Chen and A. Zakhor, "Multiple TFRC connections based rate control for wireless networks," *IEEE Trans. Multimedia*, accepted.

[12] M. Chen, A. Abate, and S. Sastry, "New congestion control schemes over wireless networks: Stability analysis," in *Proceedings of the* 16^{th} *IFAC World Congress*, Prague, 2005.

[13] S. Sastry, *Nonlinear Systems, Analysis, Stability and Control*. New York, NY: Springer Verlag, 1999.

[14] H. Khalil, *Nonlinear Systems (3rd edition)*. Prentice Hall, 2001.