

Stabilized MAX-MIN Flow Control Using PID and PII² Controllers[†]

Jeong-woo Cho and Song Chong

Department of Electrical Engineering and Computer Science
Korea Advanced Institute of Science and Technology
Daejeon 305-701, Republic of Korea.

Abstract—This paper describes an analytical framework for the *weighted max-min* flow control of elastic flows in packet networks using PID and PII² controller when flows experience heterogeneous round-trip delays. Our algorithms are *scalable* in that routers do not need to store any per-flow information of each flow and they use simple first come first serve (FCFS) discipline, *stable* in that the stability is proven rigorously when there are flows with heterogeneous round-trip delays. We first suggest two closed-loop system models that approximate our flow control algorithms in continuous-time domain where the purpose of the first algorithm is to achieve the target queue length and that of the second is to achieve the target utilization. The slow convergence of source rates traversing routers with empty buffers which is inherent in many flow control algorithms can be resolved by the second algorithm. Based on these models, we find the conditions for controller gains that stabilize closed-loop systems when round-trip delays are equal and extend this result to the case of heterogeneous round-trip delays with the help of Zero exclusion theorem. We simulate our algorithms with optimal gain sets for various configurations including a multiple bottleneck network to verify the usefulness and extensibility of our algorithms.

I. INTRODUCTION

Recently many efforts have been devoted to provide a framework for designing best-effort service networks that can offer low-loss, low-delay data services where flow control plays a major role in controlling congestion as well as allocating bandwidth among users by enforcing users to adjust their transmission rate in a certain way in response to congestion in their path. The potential advantages of such networks would be the ability to offer even real-time services without the need for complicated admission control, resource reservation or packet scheduling mechanisms.

Flow control is a distributed algorithm to fairly share network bandwidth among competing data sources while maximizing the overall throughput without incurring congestion. The most common understanding of fairness for a best-effort service network is *max-min fairness* as defined in [1]. The intuition behind the max-min bandwidth sharing is that any flow is entitled to as much as bandwidth use as is for any other flow with the assumption that all flows have equal priority. This intuition naturally leads to the idea of maximizing the

bandwidth use of flows with minimum bandwidth allocation, thus giving rise to the term *max-min flow control*. The appropriateness of max-min fairness as a bandwidth sharing objective has recently been questioned by Kelly [2] who has introduced the alternative notion of *proportional fairness*. The objective of proportional fairness may be interpreted as being to maximize the overall utility of rate allocations assuming each flow has a logarithmic utility function. The advantage of proportional fairness is that it can achieve greater overall throughput than max-min fairness by penalizing flows proportional to number of hops in the path in rate allocation [3]. However, this advantage is problematic regarding fairness in that the path of a flow is chosen by the routing protocol in the network, not by the user, so that it is obviously unfair to penalize flows because of their long path length. Thus, one can still argue that max-min fairness might be the better choice than proportional fairness from user's viewpoint since max-min fairness treats flows equally irrespective of their path length.

This paper concerns the design of *minimum plus weighted max-min flow control* [4], a generalization of max-min flow control, where each flow is associated with two parameters, its weight w_i and minimum rate requirement m_i , such that the minimum rate of each flow is guaranteed as requested during the entire holding time of the flow and the bandwidth unused after allocating the minimum rates is shared by all flows in the weighted max-min sense. An increase in the weight of a flow leads to an increase in the bandwidth share of the flow with the assumption that users pay more for a higher weight. Let us define flow i 's source rate, say a_i , to be $a_i \equiv w_i f_i + m_i$ where f_i is the max-min fair share of the bandwidth unused after allocating the minimum rates to all flows. Let us denote the set of all links, the set of all flows and the set of flows traversing through link l by L , N and $N(l)$, respectively. Then, the weighted max-min fairness can be defined as follows.

Definition 1: A rate vector $\langle a_1, \dots, a_{|N|} \rangle$ is said to be **feasible** if it satisfies $a_i \geq 0$, $\forall i \in N$ and $\sum_{i \in N(l)} a_i \leq \alpha_T^l \mu^l$, $\forall l \in L$.

Definition 2: A rate vector $\langle a_1, \dots, a_{|N|} \rangle$, where $a_i = w_i f_i + m_i$, is said to be **weighted max-min fair** if it is feasible, and for each $i \in N$ and feasible fair rate vector $\langle f_1, \dots, f_{|N|} \rangle$ for which $f_i < f_i$, there exists some i' with $f_i \geq f_{i'} > f_i$.

[†]This work was supported in part by Korea Science and Engineering Foundation under Grant R01-2001-000-00317-0 and in part by university IT research center program of the government of Korea.

Here μ^l denotes the capacity of link l and α_T^l is a constant defining target link utilization ($0 < \alpha_T^l \leq 1$). Note that admission control is necessary to ensure $\sum_{i \in N(l)} m_i < \alpha_T^l \mu^l$ for all $l \in L$ so that the minimum rate of each flow is guaranteed as requested during the entire holding time of the flow. Definition 2 can be restated more informally as follows: a rate vector $\langle a_1, \dots, a_{|N|} \rangle$ is said to be weighted max-min fair if it is feasible and for each user $i \in N$, its fair rate f_i cannot be increased while maintaining feasibility without decreasing the fair rate $f_{i'}$ for some user i' for which $f_{i'} < f_i$.

In this paper, our goal is to provide a control-theoretic framework based on deterministic fluid models that reveals not only the existence of such a distributed iterative algorithm but also an explicit stability condition of the algorithm in presence of flows with heterogeneous round-trip delays.

A. Our Contributions

We propose two control-theoretic max-min flow control models and algorithms. The first algorithm satisfies Definition 2 for $\alpha_T^l = 1$ such that in the steady-state, bandwidth at every bottleneck link is used to the full while the minimum plus weighted max-min fairness is maintained in bandwidth sharing. Moreover, the queue length at every bottleneck link converges to the target value, say q_T^l , thereby achieving constant queueing delay expressed by $\frac{q_T^l}{\mu^l}$. In contrast, the second algorithm satisfies Definition 2 for $0 < \alpha_T^l < 1$ such that in the steady-state, every bottleneck link achieves its target utilization ($\alpha_T^l \mu^l$) and hence virtually zero queueing delay while the minimum plus weighted max-min fairness is maintained. The motivation behind the second algorithm is making the queueing delay at each link to be virtually zero and improving transient performance by absorbing transient overshoots occurring before convergence at the expense of reduced link utilization. But the major advantage of the second algorithm is that the slow adaptation of source rates traversing routers with empty buffers is overcome with this algorithm. The sluggishness of PI controllers based on queue length is also pointed out in [5]. In summary, the former can offer zero-loss, constant-delay data services at full utilization of bottleneck links whereas the latter can offer zero-loss, zero-delay data services and faster rate adaptation at the expense of reduced bottleneck link utilization.

In the former, the difference between queue length and target queue length, i.e., $q^l(t) - q_T^l$, is used as a congestion measure at each link l and the max-min fair rate f_i is computed by a PID (proportional integral derivative) controller of this queue-length based congestion measure. In the latter, $\sum_{i \in N(l)} a_i (= w_i f_i + m_i) - \alpha_T^l \mu^l$ is used as a congestion measure at each link l and the max-min fair rate f_i is computed by a PII² (proportional integral double integral) controller of this aggregate-flow based congestion measure. We show that the closed-loop characteristics of the network under these two different algorithms are actually identical, yielding the identical stability condition. By appealing to the Nyquist stability criterion [6] and the Zero exclusion theorem in robust control theory [7], we derive the sufficient and necessary condition

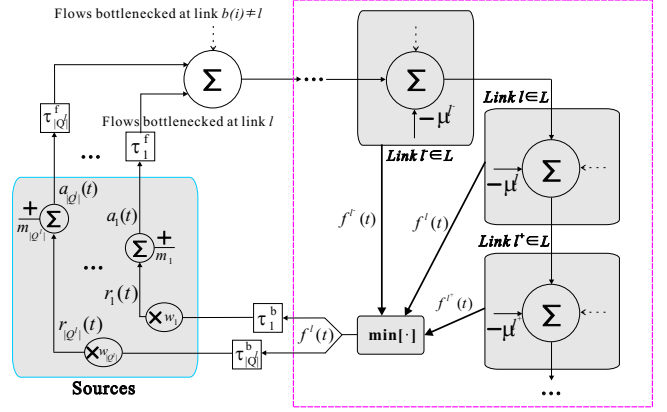


Fig. 1. The network architecture for weighted max-min fairness.

for the asymptotic stability of the network as an explicit and usable function of the upper bound $\bar{\tau}$ of all round-trip delays ($\bar{\tau} \geq \tau_i$ for all $i \in N$ where τ_i is the round-trip delay of flow i). Moreover, we find optimal controller gains for both PID and PII² controllers to maximize the asymptotic decay rate of the closed-loop dynamics, thereby achieving faster convergence. Finally, both PID and PII² controllers are highly scalable in that the computational complexity of the link algorithm is $O(1)$ with respect to number of flows passing through a link and no per-flow queueing implementation is necessary at any link. Some related works are introduced in [8].

II. NETWORK MODEL AND CONTROLLERS

In this section, we propose network models and controllers which achieve weighted max-min fairness. The network architecture with multiple sources and links is depicted in Fig. 1. Let us consider a bottleneck link $l \in L$. Then, the dynamics of the buffer of the link can be written by

$$\dot{q}^l(t) = \begin{cases} \sum_{i \in N(l)} a_i(t - \tau_i^{l,f}) - \mu^l, & q^l(t) > 0 \\ \left[\sum_{i \in N(l)} a_i(t - \tau_i^{l,f}) - \mu^l \right]^+, & q^l(t) = 0 \end{cases} \quad (1)$$

where $a_i(t)$ is the sending rate of source i , $\tau_i^{l,f}$ is the forward-path delay from source i to link l , μ^l is the link capacity of the link and the saturation function $[\cdot]^+ \equiv \max[\cdot, 0]$ represents that the $q^l(t)$ cannot be negative.

A source i sends packets according to fair rate value assigned by the network. To achieve weighted max-min fairness, let us assume that the source sends packets according to the minimum value among the fair rate values assigned by the links along the path of its flow. Thus we assume the following *source algorithm*.

$$a_i(t) = m_i + w_i \underbrace{\min_{l \in L(i)} [f^l(t - \tau_i^{l,b})]}_{f_i(t)}, \quad (2)$$

where $L(i)$ is the set of links through which flow i traverses, $f^l(t)$ is the rate value assigned by the link l on the path of flow i and $\tau_i^{l,b}$ is the backward-path delay from link l to source i . Because $\min[\cdot]$ operation is taken over a finite number of links,

there should exist at least one link l such that $f^l = \min[\cdot]$. Therefore, each flow i has at least one bottleneck $l \in L(i)$. There are several assumptions employed for the analysis of the network model.

- A.1. We assume that the sources are *persistent* until the closed-loop system reaches steady state. By persistent, we mean that the source always has enough data to transmit at the allocated rate.
- A.2. There are two delays, say, the forward-path delay $\tau_i^{l,f}$ and the backward-path delay $\tau_i^{l,b}$, which include propagation, queueing, and transmission and processing delays. We denote the sum of two delays by τ_i and assume that this is constant.

A. The PID Control Model

To control flows and to achieve weighted max-min fairness, we use a PID link controller at each link. In PID link controller model, there is a specified target queue length q_T^l to avoid underutilization of the link capacity. Because we have a nonzero target queue length q_T^l , PID model implies that $\alpha_T^l = 1$ in Definition 1. Each link calculates the common feedback rate value $f^l(t)$ for all flows traversing through the link according to PID control mechanism.

In general, a proportional term increases the convergence speed of transient responses and reduces errors caused by disturbances. An integral term is necessary to eliminate steady state error and it decreases the size of stability region. A derivative term adds some damping and extends the area of stability region and it also improves the performance of transient periods.

Let us denote the set of flows bottlenecked at link l and its cardinality by Q^l and $|Q^l|$. The *link algorithm with The PID controller* that uses the difference between $q^l(t)$ and q_T^l as its input is given by

$$f^l(t) = \left[-\frac{1}{|Q_w^l|} \left(g_P e_1^l(t) + g_I \int_0^t e_1^l(t) dt + g_D \dot{e}_1^l(t) \right) \right]^+ \quad (3)$$

where $e_1^l(t) \equiv q^l(t) - q_T^l$ is the error signal between control target and current output signal and, $g_P > 0$ and $g_I, g_D \geq 0$. Here, $|Q_w^l|$ denotes the sum of locally bottlenecked flows' weights, i.e., $|Q_w^l| \equiv \sum_{i \in Q^l} w_i$. For convenience in deriving our results, we use the definition $\rho_i \equiv w_i / |Q_w^l|$. Then it is satisfied that $\sum_{i \in Q^l} \rho_i = 1$ and $\rho_i > 0$ because we require $w_i > 0$.

Suppose that the closed-loop system has an equilibrium point at which the derivatives of the system variables are zero, i.e., $\lim_{t \rightarrow \infty} \dot{q}^l(t) = 0$, $\lim_{t \rightarrow \infty} q^l(t) = q_T^l$, $\lim_{t \rightarrow \infty} a_i(t) = a_{is}$ and $\lim_{t \rightarrow \infty} f^l(t) = f_s^l$. To be more formal, the set of flows bottlenecked at link l is given by

$$Q^l = \{i | i \in N(l) \text{ and } a_{is} = w_i f_s^l + m_i\} \quad (4)$$

and the set of all flows not bottlenecked at link l but traversing through link l , $N(l) - Q^l$, is given by

$$N(l) - Q^l = \{i | i \in N(l) \text{ and } a_{is} = w_i f_s^{b(i)} + m_i \text{ and } f_s^{b(i)} < f_s^l\}. \quad (5)$$

where $b(i) \in L(i)$ ($b_i \neq l$) is some bottleneck for flow $i \in N(l) - Q^l$. Then the equation (1) implies that the link capacity μ^l in PID link controller model is fully utilized as follows.

$$\sum_{i \in N(l)} a_{is} = \mu^l. \quad (6)$$

Using (6), and the definitions (4) and (5), we obtain

$$\sum_{i \in Q^l} (w_i f_s^l + m_i) + \sum_{i \in N(l) - Q^l} (w_i f_s^{b(i)} + m_i) = \mu^l$$

which establishes that the PID control model achieves the following *weighted max-min fairness property*

$$r_{is} \equiv w_i f_s^l = \frac{w_i}{|Q_w^l|} \left(\mu^l - \sum_{i \in N(l) - Q^l} w_i f_s^{b(i)} - \sum_{i \in N(l)} m_i \right). \quad (7)$$

B. The PII² Control Model

Instead of using $e_1^l(t)$, one can use $e_2^l(t) \equiv \sum_{i \in N(l)} a_i(t - \tau_i^{l,f}) - \alpha_T^l \mu^l$ as an input of link controllers where α_T^l is the target utilization of link l and should be a positive value smaller than 1. In this case, one can use PII² control model as follows because we now use rate error signal instead of queue error signal.

$$f^l(t) = \left[-\frac{1}{|Q_w^l|} \left(h_P e_2^l(t) + h_I \int_0^t e_2^l(t) dt + h_{I^2} \int_0^t \int_0^t e_2^l(t) dt dt \right) \right]^+ \quad (8)$$

where $h_P, h_{I^2} \geq 0$ and $h_I > 0$. It should be remarked that PID and PII² model are not identical because $\dot{e}_1^l(t) = \dot{q}^l(t) = \sum_{i \in N(l)} a_i(t - \tau_i^{l,f}) - \mu^l \neq e_2^l(t)$ for $q^l(t) > 0$. In this model, the purpose of control is to achieve the target utilization, α_T^l . In PII² model, note that $\dot{q}^l(t) = -(1 - \alpha_T^l) \mu^l < 0$ when $q^l(t) > 0$ and $e_2^l(t) = 0$. Therefore, this model controls flows so that the queue length at steady state becomes zero at the cost of some degree of underutilization. In PID model, note that $e_1^l(t)$ cannot be smaller than $-q_T^l$ because $q^l(t)$ cannot be negative. Thus, one axiomatic advantage of PII² model is that the control dynamics are not saturated at $q^l(t) = 0$ because the controller uses the rate error signal as its input instead of the queue error signal. Thus the main physical saturation nonlinearity of the PID control model can be overcome by this model.

For steady state analysis, following a similar way given in Section II-A except that $\mu^l \rightarrow \alpha_T^l \mu^l$ and $\lim_{t \rightarrow \infty} q^l(t) = 0$, r_{is} is given as follow.

$$r_{is} \equiv w_i f_s^l = \frac{w_i}{|Q_w^l|} \left(\alpha_T^l \mu^l - \sum_{i \in N(l) - Q^l} w_i f_s^{b(i)} - \sum_{i \in N(l)} m_i \right). \quad (9)$$

This shows that the PII² control model also achieves *weighted max-min fairness property*.

III. STABILITY ANALYSIS

Although we presented a multiple bottleneck network architecture in Section II, rigorous stability analysis of these kinds of models has been shown to be very difficult in [9] due to the dynamics coupling among links employing FCFS (first come first serve) discipline. In [9], though such dynamics

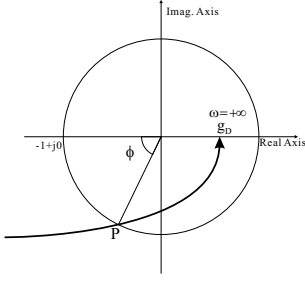


Fig. 2. Nyquist plot of $G_0(j\omega)$.

coupling exists in theory, the effect of coupling was shown to be negligible through simulations. Recently, Wydrowski et al. [10] also showed that the dynamics coupling is of a very weak form. Thus, in this section, we drop the superscript l and the analysis is focused on a single bottleneck model. We conjecture that our analytical results can be extended to multiple bottleneck models without significant modification.

We describe the stability conditions for controller gains for two network models when the saturation functions employed in (1), (3) and (8) are relaxed. The main contribution of our analysis is that we find the *equivalent* stability condition in continuous-time domain for the case flows experience heterogeneous round-trip delays, and the stability condition depends only on a given upper bound of round-trip delays. Due to space limitation, we concentrate on the PID control model and similar arguments for PII² control model are given in [8].

A. Single-Delay Case

To analyze the single-delay case of the PID control model, we simply set $|Q|$ to be 1, then there is only one flow which is bottlenecked at the link. We also regard this case as the situation where all round-trip delays of flows are equal and ρ_i are chosen to be $\rho_i = w_i/|Q_w|$, $\forall i \in Q$. This case allows us to drop the subscript of τ_1 , so that the round-trip delay of flow 1 be τ . By the single-delay assumption, ρ_i can be set as follows.

$$\rho_1 = \frac{w_1}{|Q_w|} = 1 \quad \text{and} \quad \rho_i = 0, \quad \forall i > 1. \quad (10)$$

By Eq. (3) and plugging Eq. (2) into Eq. (1), we can get the following equations.

$$\begin{aligned} \ddot{e}_1(t) &= w_1 \dot{f}(t - \tau), \\ \dot{f}(t) &= -\frac{1}{|Q_w|} [g_P \dot{e}_1(t) + g_I e_1(t) + g_D \ddot{e}_1(t)]. \end{aligned}$$

Then the corresponding Laplace transform of the open-loop system is given by

$$G(s) \doteq \underbrace{\left(g_D + \frac{g_P}{s} + \frac{g_I}{s^2} \right)}_{G_0(s)} \exp(-\tau s). \quad (11)$$

By $s = j\omega$, the following equations, to which we now apply Nyquist stability criterion [6], are obtained.

$$G(j\omega) = G_0(j\omega) \exp(-j\tau\omega), \quad G_0(j\omega) = g_D - j\frac{g_P}{\omega} - \frac{g_I}{\omega^2}. \quad (12)$$

Note that the Nyquist plot of $G_0(j\omega)$, which is depicted in Fig. 2 starts at $-j\infty - \infty^2$ in the third quadrant where $\omega = 0^+$ and ends at g_D where $\omega = +\infty$. Inferring from Fig. 2, we can see that the condition $|g_D| < 1$ is necessary because the Nyquist plot of $G(j\omega)$ will encircle or touch $-1 + j0$ unless the condition is satisfied.

Let us denote by P and $\bar{\omega}$ the point at which Nyquist plot of $G_0(j\omega)$ intersects with the unit circle and the value of ω at P , respectively. As shown in Fig. 2, ϕ is the angle between P and $-1 + j0$. More precisely,

$$\phi = \arccos(-\text{Re}[G_0(j\bar{\omega})]). \quad (13)$$

Since the Nyquist plot of $G(j\omega)$ is the Nyquist plot of $G_0(j\omega)$ rotated by $\tau\omega$ in the clockwise direction, it is required by Nyquist stability criterion that $\tau\omega < \phi$. Before proving the theorem for single-delay case, we need the following proposition. (Its proof is in [8].)

Proposition 1: If there exists a unique value $\bar{\omega}$ in $0 < \bar{\omega} < \infty$ and $\bar{\omega}\tau < \pi$ such that $|G(j\bar{\omega})| = 1$, $\text{Im}[G(j\bar{\omega})] < 0$, and $|G(j\omega)| > 1$ for all $\omega < \bar{\omega}$, then $\text{Im}[G(j\omega)] < 0$ is satisfied for all ω in $0 < \omega \leq \bar{\omega}$.

With the help of Proposition 1, the equivalent stability condition for the single-delay case now can be stated as follows. (Its proof is in [8].)

Theorem 1 (Single-Delay Case, PID Model): The closed-loop system of the PID model with a single delay $\tau \geq 0$ is asymptotically stable if and only if $|g_D| < 1$ and the delay is bounded by

$$0 \leq \tau < \frac{\arccos\left(\frac{g_I}{\bar{\omega}^2} - g_D\right)}{\bar{\omega}}. \quad (14)$$

B. Explicit Stability Conditions

Although we acquired the equivalent condition for the stability of our closed-loop system, the conditions are implicit and do not allow easy choice of controller gains, g_P , g_I and g_D . To obtain more explicit stability conditions, we proceed in the following way.

We assume that τ is fixed to a value and that $g_D \geq 0$, $g_P > 0$ and $g_I \geq 0$. We will find explicit conditions for controller gains. Now, there are three variables, i.e., g_P , g_I and g_D , concerned with the stability conditions. For mathematical tractability, we will ignore the case $\tau = 0$ and use the following definitions of variables.

$$\omega_1 \doteq \bar{\omega}\tau, \quad G_D \doteq g_D, \quad G_P \doteq g_P\tau, \quad G_I \doteq g_I\tau^2.$$

If we rewrite Eq. (14) and the condition for $\bar{\omega}$ in terms of new variables assuming $\tau > 0$, it follows that

$$0 < \omega_1 < \arccos\left(\frac{G_I}{\omega_1^2} - G_D\right) \quad (15)$$

$$\text{and} \quad \left(\frac{G_I}{\omega_1^2} - G_D\right)^2 + \left(\frac{G_P}{\omega_1}\right)^2 = 1. \quad (16)$$

Corollary 1 (Explicit Stability Region): The stability condition given in Theorem 1 is equivalent to the following equations.

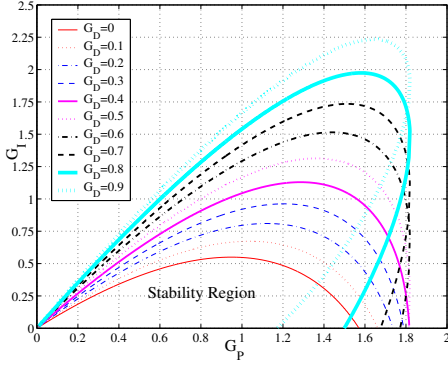


Fig. 3. Explicit stability region in terms of G_D , G_P and G_I .

$$0 \leq G_D < 1, \quad (17)$$

$$0 < G_P < \begin{cases} \arccos(-G_D)\sqrt{1-G_D^2} & \text{if } 0 \leq G_D < -\cos(\omega_0), \\ \omega_0 \sin(\omega_0) & \text{if } -\cos(\omega_0) \leq G_D < 1, \end{cases} \quad (18)$$

$$\begin{cases} 0 \leq G_I < \omega_{*1}^2(G_D + \cos(\omega_{*1})) & \text{if } \arccos(-G_D) \leq \omega_0, \\ \omega_{*2}^2(G_D + \cos(\omega_{*2})) < G_I < \omega_{*1}^2(G_D + \cos(\omega_{*1})) & \text{if } \omega_0 < \arccos(-G_D), \end{cases} \quad (19)$$

where $\omega_0 \approx 2.03$ is the value maximizing the function $\omega \sin(\omega)$ over the interval $0 < \omega < \pi$, ω_{*1} is the unique solution of $G_P = \omega \sin(\omega)$ over the interval $0 < \omega \leq \omega_0$, and ω_{*2} is the unique solution of $G_P = \omega \sin(\omega)$ over the interval $\omega_0 < \omega < \arccos(-G_D)$ which exists only when the condition $\omega_0 < \arccos(-G_D)$ is satisfied.

The proof of this corollary is in [8]. This corollary allows us to draw an exact stability region, provided that we are given a value of G_D . With the help of Corollary 1, an explicit stability region is depicted in Fig. 3 for various values of G_D . Notably, stability region corresponding to $G_D = 0$ is exactly the same to the stability region found in [11] where PI controller was used for flow control.

C. Heterogeneous-Delay Case

In this section, we prove a theorem that allows us to control flows with heterogeneous round-trip delays only with the knowledge of a given upper bound of round-trip delays. This point is important because a router may not store round-trip delay values of flows because doing so inevitably compels a router to store per-flow information.

With cancellation of Eq. (10), we now consider more general situation where all round-trip delays of flows can be different and the sum of ρ_i is less than or equal to 1. The reason for allowing $\sum_{i \in Q} \rho_i < 1$ will be clear soon. Similar to Section III-A, we can get the following open-loop transfer function.

$$G(s) \doteq \underbrace{\left(g_D + \frac{g_P}{s} + \frac{g_I}{s^2} \right)}_{G_0(s)} \sum_{i \in Q} \rho_i \exp(-\tau_i s). \quad (20)$$

Before proving the theorem, we need a proposition.

TABLE I
PARAMETERS USED FOR SIMULATION

$\bar{\mu}^l$	β	T_1	α_T^l	T_2	δ	λ	CPS	DPS	NCP	W
$1.1\mu_l$	0.02	30Δ	0.95	150Δ	0.9	0.98	40bytes	500bytes	30	300Δ

Proposition 2 (-I+j0 Exclusion Theorem): Given a fixed value ω , let us define the value set

$$V(\omega) = \left\{ Z = G(j\omega, \rho_i, \tau_i) \mid \sum_{i \in Q} \rho_i \leq 1, 0 \leq \tau_i \leq \tau \right\}.$$

The system is asymptotically stable if and only if the following two conditions are satisfied:

- There exists a $2|Q|$ -tuple vector $(\rho_1, \dots, \rho_{|Q|}, \tau_1, \dots, \tau_{|Q|})$ such that the system with the open-loop transfer function of $G(s, \rho_i, \tau_i)$ is asymptotically stable.
- For all $\omega \geq 0$, the value set $V(\omega)$ does not touch the point $-1 + j0$, i.e., $-1 + j0 \notin V(\omega)$.

Basically, this proposition is a direct application of the *Zero exclusion theorem* which is one of main results in robust control theory. Explanation and proof of the theorem can be found in [7]. Denoting by $\bar{\tau}$ an upper bound of τ_i , i.e., $\max_{i \in Q} \tau_i \leq \bar{\tau}$, we are ready to state our main result. (Its proof is in [8].)

Theorem 2 (Heterogeneous-Delay Case, PID Model):

The closed-loop system of the PID model with heterogeneous delays is asymptotically stable for all $0 \leq \tau_i \leq \bar{\tau}$ and for all ρ_i satisfying $\sum_{i \in Q} \rho_i \leq 1$ if and only if the closed-loop system of the single-delay case with delay $\bar{\tau}$ is asymptotically stable.

This theorem guarantees that a network is stabilized for all combinations of $0 \leq \tau_i \leq \bar{\tau}$ if routers know only one upper bound of round-trip delays, i.e., $\bar{\tau}$, by choosing a controller gain set $(G_D, G_P, G_I) = (g_D, g_P \bar{\tau}, g_I \bar{\tau}^2)$ contained in stability region depicted in Fig. 3. Observe that the closed-loop dynamics should be better when the $\bar{\tau}$ is more tightly chosen. A method for the estimation of $|Q_w|$ is explored in [8] because a router without per-flow information cannot know the exact sum of w_i . By appealing to Theorem 2, we can see that it is *completely safe* to overestimate $|Q_w|$, i.e., $|\hat{Q}_w| \geq |Q_w|$ where $|\hat{Q}_w|$ is the estimate of $|Q_w|$, because $\sum_{i \in Q} \rho_i = \sum_{i \in Q} w_i / |Q_w|$ is allowed to be smaller than 1.

IV. SIMULATION RESULTS

Due to space limitation, readers are encouraged to refer to [8] for finding optimal gain sets of PID, PII², PI and II² models. For determination of main parameters including β , $\bar{\tau}^l$ and q_T^l , and estimation of $|Q_w|$, readers are also encouraged to refer to [8].

Here we give simulation results in two scenarios to demonstrate the performance of our algorithms and to compare the performance of two models. The simulations are done using the ns-2 simulator [12]. The largest round-trip propagation delay on paths traversing through a link is set to 200ms. We assume that all packets are enqueued in the same buffer

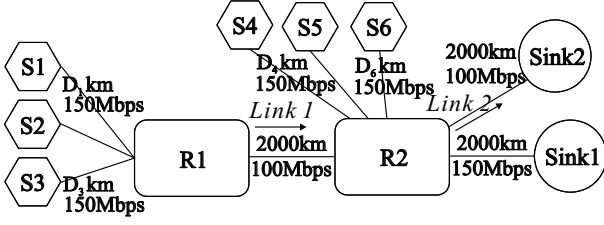


Fig. 4. Multiple bottleneck network used for Scenario 1

TABLE II

FLOW MODELS USED FOR SCENARIO 1 AND FAIR RATE. (THE UNITS OF m_i AND FAIR RATE ARE IN Mbps AND THE UNITS OF D_i ARE IN km. THE UNITS OF ARRIVAL AND DEPARTURE TIME ARE IN SECONDS.)

Src.	m_i	w_i	D_i	Arr.	Dept.	Sink
S1	5(4.75)	1	8000	$-\infty$	∞	Sink1
S2	15(14.25)	3	16000	$-\infty$	∞	Sink1
S3	0	2	4000	5	∞	Sink2
S4	20(19)	3	8000	$-\infty$	∞	Sink2
S5	0	3	12000	$-\infty$	20	Sink2
S6	0	2	4000	10	15	Sink2

Src.	Fair Rate in Time Interval				
	$-\infty \sim 5$	$5 \sim 10$	$10 \sim 15$	$15 \sim 20$	$20 \sim \infty$
S1	25(23.75)	20(19)	21(19.95)	20(19)	18.33(17.42)
S2	75(71.25)	60(57)	63(59.85)	60(57)	55(52.25)
S3	-	20(19)	16(15.2)	20(19)	26.67(25.3)
S4	60(57)	50(47.5)	44(41.8)	50(47.5)	73.33(69.7)
S5	40(38)	30(28.5)	24(22.8)	30(28.5)	-
S6	-	-	16(15.2)	-	-

and served simply with FCFS discipline and w_{max} is set to 3. Other parameters used for simulation are given in Table I where Δ is one data packet transmission time, i.e., $\Delta \equiv \text{DPS}/\mu^l$. Note that the target queue lengths are $q_T^l \approx 50\text{kbytes}$ with $\mu^l = 100\text{Mbps}$ and the queueing delay of a link with $q^l(t) = q_T^l$ is approximately 4ms for the PID model. Simulation results for PID, PII^2 , PI and II^2 models are respectively denoted by G_{PID}^3 , $G_{PII^2}^3$, G_{PID}^2 and $G_{PII^2}^2$.

A. Multiple Bottleneck Network

In the first scenario, we investigate various properties of our algorithm. In [9] and [10] authors showed through simulations and analysis that the local stability condition derived in the neighborhood works well for the FCFS discipline. By appealing to this result, we here consider a scenario where two bottlenecks exist. The network configuration is shown in Fig. 4 where the bottleneck links 1 and 2 have the link capacity of 100Mbps. The flow models used in this scenario and theoretical rates of flows satisfying the *weighted max-min fairness property* for the PID model are summarized in Table II. For the PII^2 model, changes are indicated in parentheses when needed. At $t = 0\text{s}$, the queue lengths at link 1 and 2 are already stabilized with 4 flows, S1, S2, S4 and S5. For four gain sets, values of $q^l(t)$ and $|\hat{Q}_w^l(t)|$ at link 1 and 2, and source transmission rates $a_i(t)$ are shown in Fig. 5. For G_{PID}^3 and G_{PID}^2 , the queue lengths are controlled to the target queue length except transient periods. For $G_{PII^2}^3$ and $G_{PII^2}^2$, the

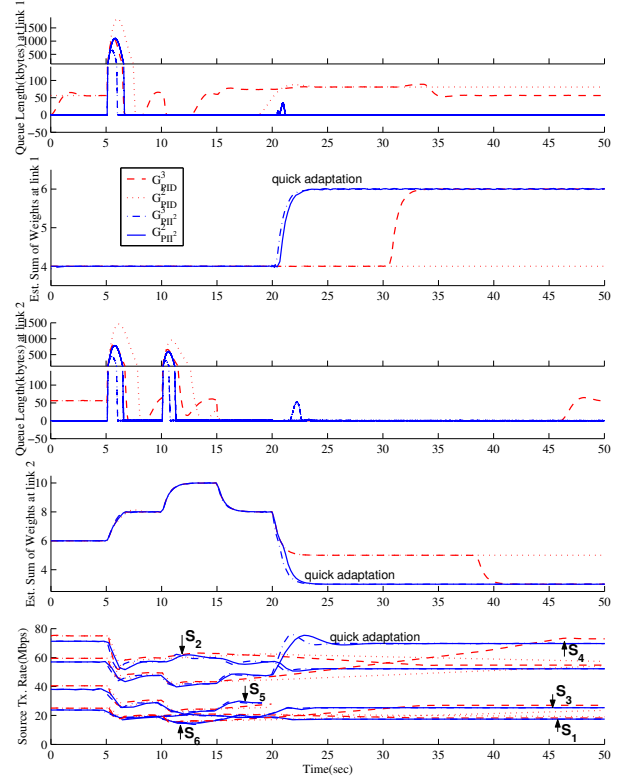


Fig. 5. Results of Scenario 1: From top to bottom - Queue length at link 1 ($q^1(t)$), Estimated sum of locally bottlenecked flows' weights at link 1 ($|\hat{Q}_w^1(t)|$), Queue length at link 2 ($q^2(t)$), Estimated sum of locally bottlenecked flows' weights at link 2 ($|\hat{Q}_w^2(t)|$) and Source Transmission Rates ($a_i(t)$).

queue length are nearly zero except transient periods, at the cost of 5% underutilization. It can be observed that the queue length overshoot is smaller and the rate adaptation is faster when a three-term controller (G_{PID}^3 or $G_{PII^2}^3$) is used instead of a two-term controller (G_{PID}^2 or $G_{PII^2}^2$). The overshoots of queue length at $t = 5\text{s}$ and $t = 10\text{s}$ are mainly due to the smallness of $|\hat{Q}_w^l|$ in this scenario. At $t = 50\text{s}$, the queue length with the gain set G_{PID}^2 is still being stabilized after the departure of S5 at $t = 20\text{s}$ because the error signal $e_1^2(t)$ is saturated to $-q_T^l$ due to the saturation nonlinearity at $q^2(t) = 0$. $G_{PII^2}^3$ and $G_{PII^2}^2$ achieve theoretical fair rates in advance of two gain sets used for the PID model. S3 is bottlenecked at link 2 from $t = 5\text{s}$ to $t = 20\text{s}$ and bottlenecked at link 1 as S5 stops its transmission at $t = 20\text{s}$. The actual source transmission rates approach to the theoretical fair rates given in Table II except transient period. We can see that our algorithms work well even if multiple bottlenecks exist.

B. Simple Network With Short-lived Flows

In the second scenario, we investigate the effect of short-lived flows to our algorithm. We use a simple network shown in Fig. 6 where 10 persistent sources with $D_i = i \times 1800\text{km}$, $i \in \{1, 2, \dots, 10\}$ and 15 on-off sources with $D_i = (i - 10) \times 1200\text{km}$, $i \in \{11, 12, \dots, 25\}$ exist. A on-off source is modelled by a two-state birth-death model where the dwell time periods in on and off state are exponentially distributed

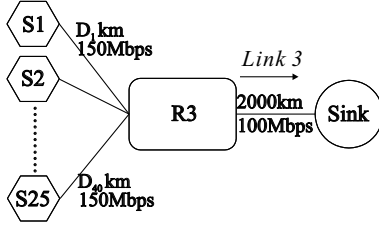


Fig. 6. Simple network used for Scenario 2.

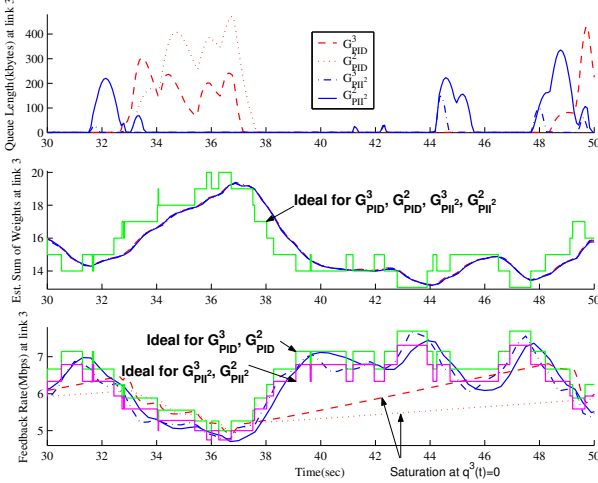


Fig. 7. Results of Scenario 2: From top to bottom - Queue length at link 3 ($q^3(t)$), Estimated sum of locally bottlenecked flows' weights at link 3 ($|\hat{Q}_w^3|$) and Feedback rate at link 3 ($u^3(t)$).

with the mean of 5s and 10s respectively. $w_i = 1$ and $m_i = 0$ are used. A portion of the simulation results is shown in Fig. 7. Although the results are oscillatory due to short-lived flows, the feedback rate and estimated sum of flows' weights at link 3 are kept in the neighborhood of 6.67Mbps (or 6.33Mbps) and 15 respectively. Note that the PII^2 model is good at tracking ideal fair rates and $G_{PII^2}^3$ has the smallest transient queue length. The utilization of link 3 is found to be 0.946, 0.913, 0.947 and 0.945 respectively with the gain set G_{PID}^3 , G_{PID}^2 , $G_{PII^2}^3$ and $G_{PII^2}^2$. The low utilization of G_{PID}^2 is caused by its slow rate adaptation when the queue length is zero. The utilization of $G_{PII^2}^3$ and $G_{PII^2}^2$ is very close to $\alpha_T^3 = 0.95$ due to their fast rate adaptation.

V. CONCLUSION

In this paper, we provided two network models which satisfy *weighted max-min fairness* and dispense with any kind of per-flow operation in routers. We found equivalent stability

conditions in two network models with heterogeneous round-trip delays. The theorem states that a stabilizing gain found with single-delay τ also stabilizes all the networks with heterogeneous delays less than or equal to τ and overestimation of the sum of flows' weights is completely safe. We also derived the equivalent condition for the asymptotic stability of the network as an explicit and usable function of the upper bound $\bar{\tau}$ of all round-trip delays. We showed in [8] that the gain sets maximizing the asymptotic decay rates do not cause the serious performance degradation even though they are obtained using only an upper bound of round-trip delays.

The PID model achieves not only full utilization but also the target queue length at its equilibrium point. The PII^2 model achieves zero queueing delays and absorbs transient overshoots in links sacrificing some degree of utilization, and α_T^l can be lowered to absorb the transient queues when many short-live flows exist. It also quickly achieves fair rates because the saturation nonlinearity at empty buffers is now eliminated. We believe that our analytical and experimental results will play an important role in encouraging the usage of more sophisticated flow control algorithms in packet networks.

REFERENCES

- [1] D. Bertsekas and R. Gallager, *Data Networks*. New Jersey: Prentice Hall, 1992.
- [2] F. Kelley, "Charging and rate control for elastic traffic," *Euro. Trans. Telecommun.*, vol. 8, pp. 33–37, 1997.
- [3] L. Massoulié and J. Roberts, "Bandwidth sharing: Objectives and algorithms," in *Proc. of IEEE INFOCOM '99*, 1999, pp. 1395–1403.
- [4] B. Vandalore, S. Fahmy, R. Jain, R. Goyal, and M. Goyal, "General weighted fairness and its support in explicit rate switch algorithms," *Comp. Commun.*, vol. 23, no. 2, pp. 149–161, Jan. 2000.
- [5] F. Ren, C. Lin, X. Ying, X. Shan, and F. Wang, "A robust active queue management algorithm based on sliding mode variable structure control," in *Proceedings of IEEE INFOCOM 2002*, 2002, pp. 13–20.
- [6] R. C. Dorf and R. H. Bishop, *Modern Control Systems*. Addison-Wesley, 1995.
- [7] J. Ackermann, *Robust Control: Systems with Uncertain Physical Parameters*. New York: Springer-Verlag, 1993.
- [8] J.-w. Cho and S. Chong, "Stabilized max-min flow control using PID and PII^2 controllers," *Draft Version*, 2003. [Online]. Available: <http://netsys.kaist.ac.kr/ggumdol/pidpii2.pdf>
- [9] L. Benmohamed and S. Meerkov, "Feedback control of congestion in packet switching networks: The case of multiple congested node," *Int. J. Commun. Syst.*, vol. 10, no. 5, pp. 227–246, 1997.
- [10] B. Wyrowski, L. L. H. Andrew, and M. Zukerman, "MaxNet: A congestion control architecture for scalable networks," *IEEE Communications Letters*, vol. 7, no. 10, pp. 511–513, Oct. 2003.
- [11] S. Chong, S. H. Lee, and S. H. Kang, "A simple, scalable, and stable explicit rate allocation algorithm for max-min flow control with minimum rate guarantee," *IEEE/ACM Trans. Networking*, vol. 9, no. 3, pp. 322–355, June 2001.
- [12] UCB/LBNL/VINT Network Simulator-ns (Version 2.1b9a). [Online]. Available: <http://www.isi.edu/nsnam/ns/>