MaxNet: A Congestion Control Architecture for Scalable Networks

Bartek Wydrowski, *Member, IEEE*, Lachlan L. H. Andrew, *Member, IEEE*, and Moshe Zukerman, *Senior Member, IEEE*

***Abstract*—MaxNet is a distributed congestion control architecture. This paper analyzes the stability properties of MaxNet. We show that MaxNet is stable for networks with arbitrary delays, numbers of sources, capacities, and topologies. Unlike existing proposals, MaxNet does not need to estimate the number of bottleneck links on the end-to-end path to achieve these scaling properties.**

***Index Terms*—Congestion control, network flow control, stability, scalability.**

I. Introduction

The problem of network flow control is to control source rates so that link capacities are utilized. For Internet-like networks, where links and sources can only have local information, the challenge is to control the source rates in a fully distributed manner. There are many different possible source rate allocations that fulfill the requirement of utilising the capacity. It has been shown [1] that the source rate allocation achieved by TCP Reno maximizes a utility function. Most other flow control schemes also maximize (different) utility functions [1]. In [2], we showed that it is possible to achieve a different, fairer rate allocation by altering the way the network signals congestion information.

Models of Internet-like networks control the source rate by a scalar feedback congestion signal. This signal is generated by aggregating the congestion prices of links on the end-to-end connection path of the source. For networks which achieve maximum utility, including TCP Reno networks, the signal is aggregated by summing all of the link prices on the path. We refer to such networks as SumNets. In [2] we introduced MaxNet, where the aggregation function is *Max*, whereby only the maximum link price along the connection path controls the source rate. In [2], we showed that MaxNet results in the Max-Min fairness criteria for sources with homogenous demand functions.

This letter focuses on the stability properties of MaxNet. In [3], the gains required at sources and links were found such that SumNet remains stable for arbitrary capacities, delays and routing. Here we adopt the approach of [3] to prove that, with a suitable choice of gains, MaxNet is also stable for arbitrary capacities, delays and routing, but with fewer requirements on the source and link gains.

Manuscript received March 23, 2003. The associate editor coordinating the review of this letter and approving it for publication was Dr. J. Kim. This work has been supported by the Australian Research Council.

The authors are with the ARC Special Research Centre for Ultra-Broadband Information Networks, Department of Electrical and Electronic Engineering, The University of Melbourne, Melbourne, Vic. 3010, Australia (e-mail: b.wydrowski@ee.mu.oz.au; l.andrew@ee.mu.oz.au; m.zukerman@ee.mu.oz.au).

Digital Object Identifier 10.1109/LCOMM.2003.818888

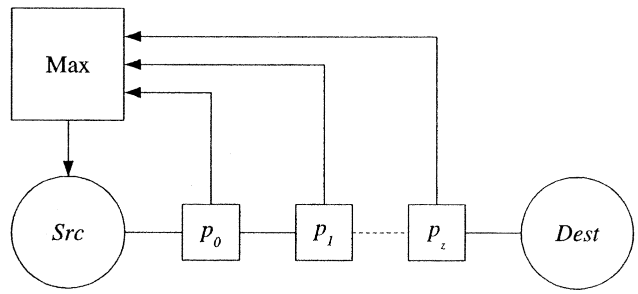

Fig. 1. MaxNet logical feedback loop.

In Section II, we start by briefly describing the MaxNet architecture presented in [2], and Section III recalls the steady state rate allocation properties. Section IV describes the control model, whose stability is analyzed in Section V.

II. MaxNet Architecture

In a MaxNet network, the congestion signal, q_i, communicated to source i is the maximum of all link prices on the end-to-end transmission path, as illustrated in Fig. 1. Let p_l be the price at link l and L_i be the set of links source i uses. Then

$$q_i = \max\{p_l : l \in L_i\}.$$

The controlling bottleneck link of the source is defined as the link, l, maximizing p_l. If multiple links achieve the maximum, then one is selected arbitrarily.

To determine q_i, the packet format must include sufficient bits to communicate the complete congestion price. Each link replaces the congestion price in the packet with its own congestion price if its own price is larger than the one in the packet. (This was shown in [4] to be more efficient than the single-bit signaling typically used in SumNets.) The congestion signal is relayed back to the source by the destination host in acknowledgment packets.

The behavior of source i is governed by an explicit demand function, $D_i(\cdot)$, such that its transmit rate is

$$x_i = D_i(q_i) \tag{1}$$

for a congestion signal q_i. The link Active Queue Management (AQM) algorithm is the well studied integrator process [3]:

$$p_l(t+1) = p_l(t) + (y_l(t) - c_l)\varphi_l \tag{2}$$

where $y_l(t) = \sum_{i:l \in L_i} x_i(t)$ is the aggregate arrival rate for link l at time t, φ_l is the control gain, and c_l is the target capacity of link l which is related to its physical capacity C_l by the target utilization $0 < \mu_l < 1$ such that $c_l = \mu_l C_l$.

III. Steady–State Rate Allocation

In steady state, when $p_l(t+1) = p_l(t)$, the rate allocation to source i, x_i will depend on the magnitude of its demand function relative to those of other sources sharing the controlling bottleneck link l of source i. If T_l is the set of sources traversing link l, then the rate allocation to i is

$$x_i = C_l \frac{D_i(q_i)}{\sum_{k \in T_l} D_k(q_k)}. \tag{3}$$

MaxNet can achieve a Max-Min fair rate allocation. A vector of rates, x, is defined as Max-Min fair if, for every feasible rate vector r with $r_i > x_i$ for some source i, there exists a source k such that $x_k \leq x_i$ and $x_k > r_k$. Put simply, a rate vector is Max-Min if it is feasible and no flow can be increased while maintaining feasibility without decreasing a smaller or equal flow.

The following proposition is proved in [2].

Proposition 1: Let $D(\cdot)$ be positive, continuous, and decreasing. The rate allocation for a MaxNet network of homogenous sources with $x_i = D(q_i)$ is Max-Min fair.

IV. Control System Model

This section derives the model of MaxNet which will be shown, in the following section, to be stable. The analysis applies the stability and robustness results derived in [3] for SumNet, and shows that MaxNet congestion control is also arbitrarily scalable and maintains stability for arbitrary network topologies and arbitrary amounts of delays.

We also show that MaxNet does not require knowledge of M_i, the number of bottleneck links (links with $p_l \neq 0$) on the end-to-end path of source i. SumNets require M_i to be estimated and communicated to the source in order to achieve stability under arbitrary network scaling [3]. Eliminating M_i removes the additional signaling infrastructure required to determine M_i, as proposed for SumNet in [3]. To remain stable without this signaling infrastructure, SumNets must assume an upper-bound on M_i and have a slow conservative control policy. With MaxNet, the number of controlling bottleneck links is always 1, which avoids these problems.

We will now describe the multi variable control system, shown in Fig. 2. Note that Fig, 2, for illustration only, shows a large-signal source, and small-signal links and network. The interconnection of sources with links is described in the Laplace domain by forward and backward routing matrices. The matrices specify the interconnection and the delay incurred in signal flow from source to link and vice versa. The forward routing matrix is

$$[\bar{R}_f(s)]_{l,i} = \begin{cases} e^{-\tau^f_{i,l} s}, & \text{if source } i \text{ uses link } l \\ 0, & \text{otherwise} \end{cases} \tag{4}$$

where $\tau^f_{i,l} s$ is the forward delay between source i and link l. Let n_i be the index of the controlling bottleneck link of source i. The backward routing matrix, which describes the flow of congestion information from each link back to sources, depends on n, and is given by

$$[\bar{R}_b(s;n)]_{l,i} = \begin{cases} e^{-\tau^b_{i,l} s}, & \text{if } n_i = l \\ 0, & \text{otherwise.} \end{cases} \tag{5}$$

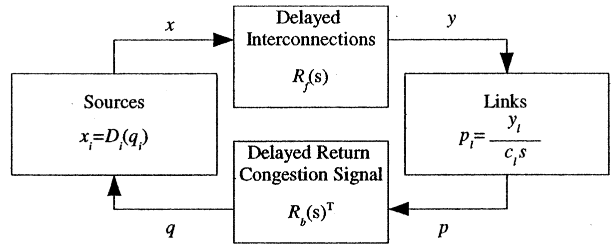

Fig. 2. Flow control structure.

Note that the round-trip time of source i's connection is $\tau_i = \tau^f_{i,l} + \tau^b_{i,l}$. Let L be the number of links in the network. Without loss of generality, order the link prices such that

$$p_1 \geq p_2 \geq \cdots \geq p_L. \tag{6}$$

The backward routing matrix remains static over a period where the variations in link prices do not change the ordering of link prices (6). The overall multi variable feedback loop in the configuration of Fig. 2 is

$$y(s) = \bar{R}_f(s)x(s) \tag{7}$$
$$q(s) = \bar{R}_b(s;n)^T p(s). \tag{8}$$

We can construct a small signal model as in [3]. Consider small perturbations around equilibrium, $x = x_0 + \delta x, y = \bar{y}_0 + \delta\bar{y}, p = \bar{p}_0 + \delta\bar{p}, q = q_0 + \delta q$, where $x_0, \bar{y}_0, \bar{p}_0, q_0$ are the steady state values and $\delta x, \delta\bar{y}, \delta\bar{p}, \delta q$ are the perturbations. When all link prices are distinct, the vector of controlling bottleneck links, n, is unchanged by a sufficiently small perturbation. In this case, $\delta\overline{p_l}$ is only nonzero for controlling bottleneck links and the small signal model (10) applies. A reduced small-signal model can then be written as

$$\delta y(s) = R_f(s)\delta x(s) \tag{9}$$
$$\delta q(s) = R_b(s)^T \delta p(s) \tag{10}$$

where the matrices R_f, R_b, and the vectors $\delta p(s), \delta y(s)$ are obtained by eliminating the rows corresponding to noncontrolling links.

The small-signal gain of perturbations in δq_i that determines the resulting perturbation in δx_i is

$$\kappa_i = D'_i(q_i). \tag{11}$$

To achieve stable control for networks of arbitrary dimensions, the gains that sources, and links introduce need to be prescribed as detailed in [3]. Because the control signal in MaxNet is not scaled up by M_i, the source gain

$$\kappa_i = \frac{\alpha_i x_{0i}}{\tau_i} \tag{12}$$

yields a stable system for any $0 < \alpha_i < 1$. The selectable parameter α_i controls the magnitude of the demand function to reflect the source's need for capacity. The term τ_i makes the stability invariant to delay. To make stability invariant to the number of sources, a gain x_{0i}/c_l is introduced in the closed-loop, with the x_{0i} component put into the source and the $1/c_l$ component in the link as $\varphi_l = 1/c_l$. In the Laplace domain, the integrator AQM of (2) with the required gain between the coupling of δp_l and δy_l is

$$\delta p_l = \frac{1}{c_l s}\delta y_l. \tag{13}$$

Note that (11) implicitly assumes a static demand function. As discussed in [3], the requirement (12) determines the shape of the static demand function. However, recent work in [5] provides dynamic source algorithms which allow arbitrary demand functions, whilst preserving the control gain required for stability invariance. They separate the high-frequency AC gain from the DC gain.

The open-loop transfer function that describes the small-signal system is

$$L(s) = \frac{1}{s}R_f(s)\mathcal{K}R_b(s)^T\mathcal{C} \tag{14}$$

where

$$\mathcal{K} = \text{diag}(\kappa_i), \quad \mathcal{C} = \text{diag}\left(\frac{1}{c_l}\right). \tag{15}$$

V. Stability

In this section, we show that the linearised closed loop MaxNet system (14) is stable for networks with arbitrary capacities, delays, numbers of sources, and links. The key is that the ordering (6) causes R_f to be block lower triangular, while R_b is block diagonal. This leads to Propositions 2 and 3 that bypass a critical symmetry requirement in the stability proof of [3], making the main result also hold for MaxNet.

Consider a standard unity feedback loop, with $L(s) = (\gamma/s)F(s)$. In [3], it is shown that the feedback system is stable for all $\gamma \in (0,1]$, given the following conditions:

1) $F(s)$ is analytic in $\text{Re}(s) > 0$, and there exists a β such that $\|F(s)\| \leq \beta$ in $\text{Re}(s) \geq 0$.
2) $F(0)$ has strictly positive eigenvalues.
3) For all $\gamma \in (0,1]$, -1 is not an eigenvalue of $L(j\omega)$ for $\omega \in \mathbb{R}$ and $\omega \neq 0$.

Here we prove the stability of MaxNet by showing that all of these conditions hold for MaxNet. Note that, as in [3], condition 1 is automatically satisfied. To prove that MaxNet satisfies the remaining conditions, we will first show that $F(s)$, and hence $L(s)$, is a lower triangular matrix.

For MaxNet, the mapping from sources to their controlling bottleneck links is many to one and onto. Without loss of generality, label the sources such that $1 \leq n_1 \leq n_2 \leq \cdots \leq n_S = L$, where n_i is again the index of the controlling bottleneck link of source i and S is the total number of sources. Each link l controls a nonempty set of sources, $m_l = \{i : n_i = l\} \neq \emptyset$.

For MaxNet, $R_b(s)$ is block diagonal since the only nonzero element of column i is in the n_ith row. Thus the lth diagonal block has size $1\times|m_l|$, where $|m_l|$ denotes the cardinality of m_l. Since $\mathcal{K}$ and $\mathcal{C}$ are both diagonal, $\mathcal{K}R_b(s)^T\mathcal{C}$ is block diagonal with the lth diagonal block having size $|m_l| \times 1$. Let B_{ll}^T denote the lth diagonal block of $\mathcal{K}R_b(s)^TC$.

By definition, a source does not use links with higher congestion prices than that of its controlling bottleneck link, i.e., $[R_f(s)]_{l,i} = 0$ if $n_i > l$. That is, $R_f(s)$ is block lower triangular, with the lth diagonal block, D_{ll}, having size $1 \times |m_l|$.

A block triangular matrix multiplied by a block diagonal matrix with conformable block structure is again block triangular. Thus $F(s) = R_f(s)(\mathcal{K}R_b(s)^T\mathcal{C})$ is block lower triangular. The lth diagonal block is $D_{ll}B_{ll}^T$, which is the scalar

$$[F(s)]_{l,l} = \frac{1}{c_l}\sum_{k\in m_l}\frac{\alpha_k x_{0k}}{\tau_k}e^{-\tau_k s}. \tag{16}$$

The stability of (14) now follows from the following two propositions.

Proposition 2: The eigenvalues of $F(0)$ are strictly positive.

Proof: The eigenvalues of a triangular matrix are the diagonal elements. By (16), the eigenvalues of $F(0)$ are

$$[F(0)]_{l,l} = \frac{1}{c_l}\sum_{k\in m_l}\frac{\alpha_k x_{0k}}{\tau_k}. \tag{17}$$

which are nonempty sums of strictly positive real numbers, and hence strictly positive. ∎

Proposition 3: For all $\gamma \in (0,1]$, -1 is not an eigenvalue of $L(j\omega)$.

Proof: Since $F(s)$ is lower triangular, so is $L(s)$, with eigenvalues

$$[L(s)]_{l,l} = \frac{1}{c_l}\sum_{k\in m_l}x_{0k}\lambda_k(s) \tag{18}$$

where

$$\lambda_i(s) = \frac{\alpha_i e^{-\tau_i s}}{\tau_i s}. \tag{19}$$

Note that $[L(s)]_{l,l}$ is a weighted sum of the $\lambda_k(s)$s, with weights $w_k = x_{0k}/c_l \geq 0$. Moreover, $\sum_k w_k \leq 1$, since the x_{0k}s are all steady state flow rates through link l, whose sum cannot exceed the capacity c_l, by (2). In [3], it was shown that if $0 < \alpha_i < 1$ then $\text{Re}[\lambda_i(j\omega)] > -1$ for $\omega \neq 0$. Thus their weighted sum must also have real part greater than -1, whence -1 is not an eigenvalue of $L(j\omega)$ for $\omega \neq 0$. ∎

VI. Conclusion

We have shown a set of control laws for MaxNet that provides local stability for arbitrary network routing, delays, and link capacities. Unlike SumNet, MaxNet can achieve this invariant stability without estimating the number of bottleneck links on the end-to-end path. MaxNet also produces a fairer steady state rate allocation than SumNet.

References

[1] S. H. Low, "A duality model of TCP and queue management algorithms," in *IEEE/ACM Trans. Networking*, vol. 11, Aug. 2003, pp. 525–536.

[2] B. Wydrowski and M. Zukerman, "MaxNet: A congestion control architecture for maxmin fairness," *IEEE Commun. Lett.*, vol. 6, pp. 512–514, Nov. 2002.

[3] F. Paganini, J. C. Doyle, and S. H. Low, "Scalable laws for stable network congestion control," *Proc. IEEE Conf. Decision Contr. (CDC)*, pp. 185–190, 2001.

[4] B. Wydrowski. (2003, Mar.) Packet marking scheme efficiency: One-bit vs packed-bit packet marking. CUBIN, Elect. Electron. Eng. Dept., The University of Melbourne, Melbourne, Australia. [Online]. Available: http://www.ee.mu.oz.au/pgrad/bpw

[5] F. Paganini, Z. Wang, S. Low, and J. Doyle, "A new TCP/AQM for stable operation in fast networks," *Proc. IEEE Infocom*, pp. 96–105, 2003.