Volume 3, Issue 9, September 2013 ISSN: 2277 128X

International Journal of Advanced Research in Computer Science and Software Engineering

Research Paper

Available online at: www.ijarcsse.com

MaxNet- A New Congestion Control Architecture

Balveer Singh
HOD CSE Deptt.
PKITM, Mathura (UP), India

Diwakar Saraswat
Astt.Prof. CSE Deptt.
PKITM, Mathura (UP), India

***Abstract*— In this paper, we will introduce a new model of performing congestion control on a source-flow controlled best-effort network. We will now explore the structural constraints of the Internet congestion signalling model and develop a new architecture for signalling congestion on an Internet-like network that we call MaxNet. MaxNet differs from Internet because source rates are controlled only by the most severely bottlenecked link on the end-to-end source to destination path. MaxNet redefines how congestion signalling is performed and results in unique fairness and scaling properties. MaxNet is in keeping with the very successful Internet philosophy of having a simple network infrastructure. Providing a minimalist service model and having the stateless waist in the protocol hourglass allows the Internet to scale with both the size of the network and heterogeneous applications and technologies.**

***Keywords*— Bottleneck, Congestion Control, Explicit Congestion Notification (ECN), Stability, Scalability.**

I. Introduction

The Internet communicates the congestion signal one bit at a time by packet dropping or marking ECN capable packets. The origins of this method of communicating congestion information are rooted in the fact that when the arriving traffic to a switch or router is greater than the outgoing capacity, the buffers will eventually overflow and a portion of the arriving packets will be lost. The rate of the individual packet loss events that the sources detect is interpreted as the congestion signal, which is a measure of the congestion of a network path. The addition of ECN abstracted this idea by substituting packet loss by packet marking. MaxNet abstracts the idea of congestion signal further and allows each link to communicate an arbitrary number of bits of congestion state [11].

MaxNet redefines how congestion signalling is performed and results in unique fairness and scaling properties. MaxNet is in keeping with the very successful Internet philosophy of having a simple network infrastructure. Providing a minimalist service model and having the ‘stateless waist’ in the protocol hourglass allows the Internet to scale with both the size of the network and heterogeneous applications and technologies. Together, they are two of the most important technical reasons behind the success of the Internet. MaxNet is an Internet like network because like the Internet it is a stateless network, where the only connection state is kept in the source and destination hosts, and sources transmit according to their utility functions. Connection state only exists in source and destination hosts and the network keeps no information about the relationship between the packets it transmits.

In this paper, we will show that MaxNet [5] is able to achieve Max-Min fairness. Although the scope of MaxNet goes beyond Max-Min fairness, we will discuss how MaxNet differs from existing solutions for achieving Max-Min fairness. Achieving Max-Min fairness on Internet like networks has till now only been shown for some special cases, where source utility functions are either globally manipulated, or in the limit of a particular family of utility functions. In this paper we will discuss how, unlike the Internet, MaxNet is able to achieve Max-Min fairness [8] for arbitrary well behaved homogenous sources, with no global coordination. Extensive research has been performed on achieving Max-Min fairness in ATM networks [7]. MaxNet is unique in that it is a fully distributed algorithm and does not require per-flow information storage or processing at the links to achieve Max-Min fairness.

II. Proposed MaxNet Architecture System

We will now describe the MaxNet congestion control architecture. The MaxNet architecture uses the maximum price of all the link prices on the end-to-end path as the feedback signal to the source. To achieve this, the packet format must include bits to communicate the complete congestion price.

TABLE I
MaxNet Packet Format

Packet Header	Congestion Price	Data Bits
H Bits	*S* Bits	*D* Bits

Each link replaces the current congestion price in the packet if the link’s congestion price is greater than the one in the packet. In this way, the maximum congestion signal on the path is communicated to the destination, which relays the information to the source in acknowledgement packets.

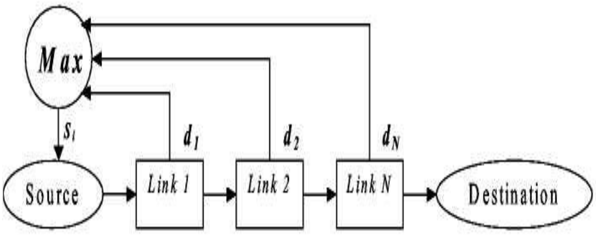

Fig.1 MaxNet control loop

Assuming each bottleneck link produces congestion notifications independently, we obtain:

$$M(t) = 1 - \prod_{i=1}^{N} (1 - m_i(t)). \tag{1.1}$$

We will now show the source rate allocation that MaxNet[5] achieves for general sources and later show that the rate allocation is Max-Min fair for a network of sources with homogenous demand functions. For this proof, the only requirement on the link congestion algorithm $G(y_l(t),d_l(t))$ is that the congestion price converges towards a steady state price at the point when the aggregate packet arrival rate, $y_l(t)$, is equal to the link target capacity C_l. The simplest form of such a process is the integrator.

$$D_l(t+1) = d_l(t) + \mu_1(y_l(t) - C_l) \tag{1.2}$$

A. MaxNet Rate Allocations

1) General Rate Allocation: Here, we describe the source rate allocation for heterogeneous sources on a MaxNet network. On a MaxNet network, each link marks the packet with its own price if it is greater than the price already marked in the packet; therefore each flow rate is controlled only by the maximum price of all prices encountered on the end-to-end path. Let L_l be the set of flows through the link l, T_f be the set of the links that flow f traverses and let N_l be the number of flows controlled by link l, i.e., flows whose maximum price (path price) is equal to the price at link l. Then the price control law for link l with price d_l, from (1.2) is

$$d_l(t+1) = d_l(t) + \mu_l \left(\sum_{x \in L_l} D_x(Max\{d_f : f \in T_x\}) - C_l \right) \tag{1.3}$$

In general, a source i, with demand function $D_i(S)$ will achieve rate r_i in steady state, where:

$$r_i = C_l \frac{D_i(Max(d_f \ f \in T_i))}{\sum_{x \in L_l} D_x(Max(d_f \ f \in T_x))} \quad i \in L_l$$

For the case where the source i is bottlenecked at the most severe bottleneck in the network l, with price d_0, it can be seen that the rate r_i is in proportion to the magnitude of the demand function relative to other demand functions.

$$r_i = C_l \frac{D_i(d_0)}{\sum_{x \in L_l} D_x(d_0)} \quad i \in L_l$$

Since sources receive bandwidth in proportion to their demand functions, MaxNet can achieve differentiated bandwidth allocation for a multi-service network. In the following section we show the unique properties of MaxNet when all of the source demand functions are the same.

2) Max-Min Fair Rate Allocation: We now show that the flow rates of MaxNet[5] in the steady state are Max-Min fair for sources with the same demand function. The assumption of homogeneity is a starting point for the analysis of MaxNet. This assumption of a single source algorithm is supported by the situation on the Internet where most traffic is generated by TCP, with TCP/Reno being widely deployed. Assume all sources share the same demand function $r_i = D(S_i)$, where r_i is the transmission rate of source i and S_i is the network congestion price seen by this source. The function $D(S_i)$ for $S_i \geq 0$ is assumed to be continuous, positive and decreasing. Let vector X contain all the source flow rates through the network arranged into subsets which contain flows of equal rates. Within the subset y of components of X, elements (flow rates) are denoted x_{ya}, x_{yb}, … The subsets are ordered as follows:

$$X = (x_{0a}, x_{0b}, \ldots), (x_{1a}, x_{1b}, \ldots), \ldots (x_{Na}, x_{Nb}, \ldots)$$
$$x_{0a} = x_{0b} = \ldots < x_{1a} = x_{1b} = \ldots < x_{Na} = x_{Nb} = \ldots \; . \tag{1.4}$$

Let X_z be the vector $(x_{za}, x_{zb}, \ldots)$. Also let x_z be the rate $x_z = x_{za} = x_{zb} = \ldots$. Let vector P contain the congestion prices that correspond to each source. Each element in P corresponds to an element in X such that $x_{yz}=D(p_{yz})$ or $p_{yz}=D^{-1}(x_{yz})$. Thusheading.

$$P = (D^{-1}(x_{0a}),\ D^{-1}(x_{0b}),...),$$
$$(D^{-1}(x_{1a}),\ D^{-1}(x_{1b}),...),\ ...$$
$$(D^{-1}(x_{Na}),\ D^{-1}(x_{Nb}),...)$$

Let p_z be the value being a decreasing function, the lowest rate in the network experiences the highest price, and the prices are ordered in reverse order to rates:

$$p_{0a}=p_{0b}\ ...\ > p_{1a}=p_{1b}=...\ > p_{Na}=p_{Nb}=...$$

$$p_0 > p_1 > p_N. \quad (1.5)$$

Notice that the flows of rate x_0 share their bottleneck links equally. If we apply the Max-Min [7]condition only to this set of minimum rate flows (the elements of the vector X_0), we see that they are Max-Min fair, since they are feasible and no rate can be increased without decreasing another rate within the flow represented by the vector X_0. We now extend this argument to include flows with price p_1.

The links in W_1 are the links with the second highest congestion prices p_1, $p_0 > p_1 > p_n$ $n \neq 0,1$. All flows through link(s) W_1 are either already controlled by nodes with price p_0, $p_0 > p_1$, or are controlled by the W_1 links, since they have the next highest congestion price in the network. Let a_{il} be the number of flows traversing link l that are controlled by a link with path price p_i. Then the price control law for links with price p_1 is:

$$p_1(t+1) = p_1(t) + \mu_1\left(D(p_1)\cdot N_l + a_{0l}\cdot D(p_0) - C_l\right) \quad l \in W_1$$

and in steady state:

$$p_1 = D^{-1}\left(\frac{C_l - a_{0l}\cdot D(p_0)}{N_l}\right) \quad l \in W_1. \quad (1.6)$$

Note that the allocation of rates to flows passing this link is Max-Min fair. The minimum flows through this link i.e., those with a congestion price p_0, are not controlled nor bottlenecked at this link, and we have already shown that their rates are maximized. All other flows receive an equal share of the remaining bandwidth by (1.6), which is also a Max-Min fair allocation. We now extend the set of flows to include all the flows in the network to show global Max-Min fairness by induction [8].

Let us assume that capacity allocation to flows with prices p_{k-1}, p_{k-2} … p_0 is Max-Min fair, and then we show that capacity allocation to flows with price p_k is also Max-Min fair. In general, links with price p_k will have a set of flows traversing them which are controlled by other links whose prices are higher, p_{k-1}, p_{k-2} … p_0, and a set of flows which are controlled by them with the price p_k. The reasoning which produced (1.6) can be generalising so that any link l with price p_k will set p_k according to:

$$p_k = D^{-1}\left(\frac{C_l - (a_{(k-1),l}\cdot D(p_{k-1}) - a_{(k-2),l}\cdot D(p_{k-2}) - ... - a_{0,l}\cdot D(p_0))}{N_l}\right) \quad l \in W_k \quad (1.7)$$

Notice that (1.7) allocates capacity so that flows with smaller rates than those controlled by p_k, $p_k < p_r$ $k > r$, are not bottlenecked at the W_k link(s). We have assumed these flows already have Max-Min allocation at other links[7]. The remaining capacity is allocated equally to the flows which are bottlenecked at the D_k link(s), thus satisfying the Max-Min condition for such flows. Therefore the Max-Min condition is satisfied by all flows traversing the W_k link(s). Because we have shown that the p_0 allocation is Max-Min fair, and that, if p_{k-1} is fair then p_k is fair, it follows by induction that all allocations are Max-Min fair and global fairness is achieved.

3) Weighted Max-Min Fairness: We will now show that MaxNet can achieve weighted Max-Min fairness [8]. Let the source demand function for source i be $r_i = a_i.D(S_i)$, where ai is the 'weight' for source i and D is the well behaved demand function shared by all sources in the network. We can consider the demand function $r_i = a_i.D(S_i)$, to be a_i virtual sources with demand function $D(S_i)$. We have proven Max-Min fairness when all demand functions are $r_i = D(S_i)$. This means that weighted Max-Min fairness is achieved for the actual sources, since they behave as a_i homogenous sources each.

B. MaxNet Stability and Robustness

In this section, we analyse the stability and robustness properties of MaxNet. Our analysis applies the stability and robustness analysis of SumNet to the case of MaxNet [5]. This section shows that MaxNet congestion control is arbitrarily scalable and maintains stability for arbitrary network topologies and arbitrary amounts of delays. Furthermore, whilst in SumNets are shown to need to estimate and communicate the number of bottlenecked links on the end-to-end path to achieve this property of arbitrary scaling, we show that MaxNet does not need to communicate or estimate the number of bottleneck links on the end-to-end path to be arbitrarily scalable.

In the analysis, we consider a network of L communication links shared by a set of S sources. The network modelled is shown in Figure 6.3. It is modelled in the Laplace domain. Each link, l, has capacity c_l, aggregate arrival rate y_l and price d_l. Each source, i, has source rate r_i and receives aggregate price q_i. The forward routing matrix describes the links loaded by each source:

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

Where $\tau^f_{i,l}$ is the delay from source i to link l. The backward routing matrix describes the aggregation method and feedback path of the congestion signals from links to sources. A SumNet is considered, all of the bottleneck links on the end-to-end path of a source have a backward feedback path to the source. With MaxNet, only one link, the link with the maximum congestion signal on the end-to-end path of the source, has a backward path to the source. The MaxNet [8] backward routing matrix changes if the maximum bottleneck links on a path changes; however, over each period when the order of bottleneck links remain unchanged, it is described by:

$$[\overline{R}_b(s)]_{l,i} = \begin{cases} e^{-\tau^b_{i,l}s} & \text{if source } i \text{ uses link } l \text{ and } d_l = Max(d_f\ f \in T_i) \\ 0 & \text{otherwise} \end{cases}$$

Note that the RTT is the sum of the forward and backward delays:

$$\tau_i = \tau^b_{i,l} + \tau^f_{i,l} \qquad (1.8)$$

The closed loop system can be described by:

$$y(s) = \overline{R}_f(s) \cdot r(s)$$
$$q(s) = \overline{R}_b(s)^T \cdot d(s)$$

In this analysis the dynamic properties around an equilibrium point, r_0, y_0, d_0 and q_0 are studied. Consider small perturbations around the equilibrium $r = r_0 + \delta r, y = y_0 + \delta y, d = d_0 + \delta d, q = q_0 + \delta q$. Assuming the set of bottlenecks is unchanged by this small perturbation, δq is only non-zero for bottleneck links. Therefore, for the local analysis, we can write the reduced model:

$$\delta\, y(s) = R_f(s) \cdot \delta\, r(s)$$
$$\delta\, q(s) = R_b(s)^T \cdot \delta\, d(s)$$

Where the matrices R_f and R_b and the vectors δr and δd are obtained by eliminating the rows corresponding to non-bottleneck links. In this small signal model, the backward and forward routing matrices are of full row rank, ruling out degenerate cases as detailed in

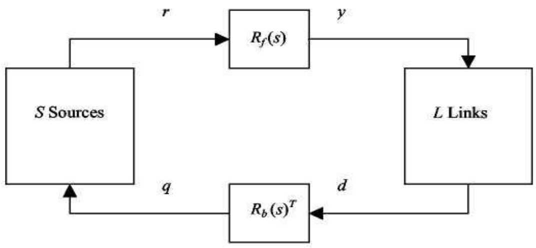

Fig.2 General flow control structure based on pricing signals

The SumNet closed loop system is stable and robust so long as each source i has a gain:

$$\kappa_i = \frac{\alpha_i r_{0i}}{M_i \tau_i} \qquad (1.9)$$

where $0 < \alpha i < 1$ is a gain parameter, r_{0i} is the steady state rate, τ_i the RTT and M_i is the number of bottleneck links in source i's path. This gain imposes some restrictions on the shape of the demand function. The relationship between this gain and the types of source algorithms possible. Each link must also have a gain scaled by the capacity:

$$d_l = \frac{1}{c_l s} y_l$$

We will prove that the same stability and robustness also holds for MaxNet, without needing to scale by M_i, with a simpler source gain:

$$\kappa_i = \frac{\alpha_i r_{0i}}{\tau_i}$$

Eliminating M_i has several advantages. Firstly, the additional signalling infrastructure required to determine M_i, as proposed for SumNet is removed. Without this signalling infrastructure, to remain stable, SumNets must assume an upper-bound on Mi and have a slow conservative control policy. With MaxNet M_i is always 1 and this avoids additional signalling infrastructure or a conservative control policy.

The overall multivariable feedback loop that describes the small signal model of the closed loop transfer function is:

$$L(s) = R_f(s) \cdot \mathrm{K} \cdot R_b^T(s) \cdot C \frac{I_L}{s}$$

where I_L is an identity matrix of size L and the gains are:

$$\mathrm{K} = \mathrm{diag}(\kappa_i), \qquad C = \mathrm{diag}\left(\frac{1}{c_l}\right)$$

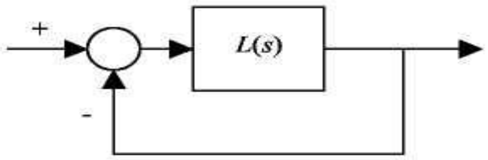

Fig.3 Overall feedback loop

Without loss of generality, let the price vector d be ordered, such that $d_0 \geq d_1 \geq K \geq d_L$. Associated with each source i, there is one link n_i which is the maximum bottleneck link that controls that source. Then let the sources be ordered such that $n_1 \leq n_2 \leq K \leq n_s$. Since every bottleneck link has at least one source it controls:

$$n_1 = 1 \qquad n_S = L \qquad n_i \leq n_{i+1} \leq n_i + 1 \tag{1.9.1}$$

Note, that we now use L as the number of bottleneck links only. We can now proceed to prove conditions (ii) and (iii).

Proposition 1. $F(0)$ has strictly positive eigenvalues.

Proposition 2. For all $\gamma \in (0,1]$, -1 is not an eigenvalue of $L(jw)$ $w \neq 0$.

MaxNet satisfies the three conditions of the proof; we have established that MaxNet is stable for arbitrary topologies and network delays, without having to communicate the number of bottlenecked links on the end-to-end path.

The scaling requirement on the feedback to the source has implications on the source flow control algorithm and possible utility functions. It is shown that it is still possible to achieve a wide variety of source behaviour given this feedback requirement. In particular, source algorithms like TCP/RENO or TCP/VEGAS have been modelled by source control laws which satisfy these requirements.

C. Simulation Results

To gain insight into the transient behaviour of MaxNet, a fluid-flow model of a MaxNet network was simulated. The network topology simulated is shown in Fig. 4 the network has four sources, S_0 to S_3, which transmit to four destination hosts D_1 to D_4 respectively. Each source has a source demand function described by (6.20) where r_{max}=15 and k=0.007. There are three bottleneck links, A,B,C of initial capacity 2 Mbps, 3 Mbps and 2 Mbps respectively. Bottlenecked links have no delay. Non-bottleneck links which interconnect the bottleneck links and sources have finite delay and infinite capacity. The number next to each non-bottleneck link represents the delay of the link in simulation time step units. The simulation was run over 8000 time step units, and the source rates are plotted in Fig 5 to 8. To generate a transient response, the capacity of link B is reduced from 3Mbps to 1Mbps at time step 4000.

Firstly, note that the steady state rates before and after the transient are Max-Min fair. This agrees with our steady state analysis. Despite the large change in capacity of link B, the convergence to new rates occurs without significant oscillatory behaviour. The simulated network is also stable under this transient, which gives some confirmation to our stability analysis.

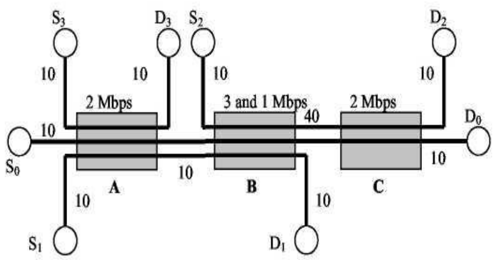

Fig.4 Simulated network topology

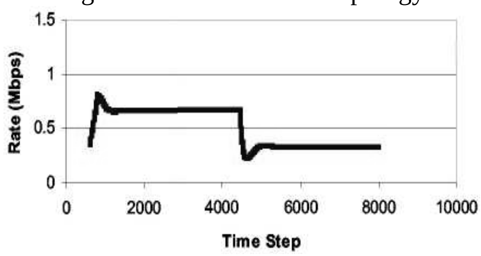

Fig.5 Source 0 rate

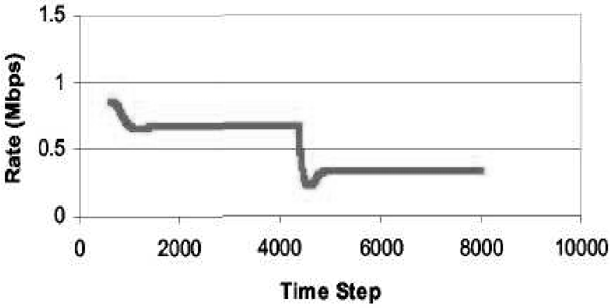

Fig.6 Source 1 rate

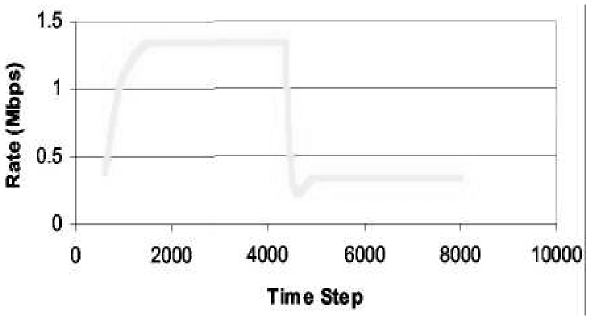

Fig.7 Source 2 rate

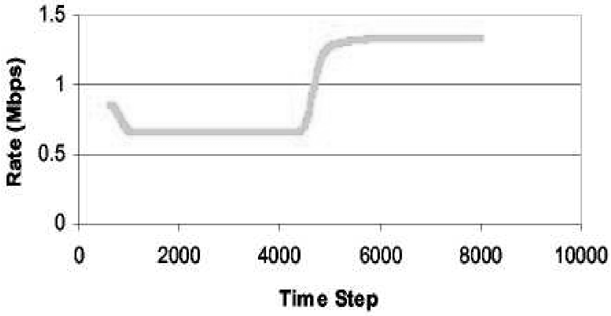

Fig.8 Source 3 rate

III. CONCLUSIONS

We introduced MaxNet, which is an Internet like congestion control architecture. We analysed its fairness properties and well behaved homogenous sources it results in Max-Min fairness. Like the Internet, MaxNet is a distributed architecture that requires no global information or per-flow state in the link. We proved that a MaxNet network is stable and robust, and its stability is invariant to the network size or topology. We also showed that MaxNet has these scalability properties without needing to communicate the number of bottleneck links on a transmissions path.

REFERENCES

[1] S. H. Low and D. E. Lapsley, "Optimization flow control, I: Basic algorithm and convergence," *IEEE/ACM Trans. Networking*, vol. 7, pp. 861–875, Dec. 1999.

[2]] F. Paganini, J. C. Doyle, and S. H. Low, "Scalable laws for stable network congestion control," *Proc. 40th IEEE Conf. on Decision and Control*, vol. 1, pp. 185–190, Dec. 2001.

[3] F. P. Kelly, "Charging and rate control for elastic traffic," *Eur. Trans.Telecommun.*, vol. 8, pp. 33–37, 1997.

[4] A. Charny, "*An algorithm for rate allocation in a packet-switching network with feedback*," MIT, Cambridge, MIT/TR-601, 19

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

[6] B. Wydrowski, "*Packet marking scheme efficiency: One-bit vs packedbit packet marking,"* Tech. Rep., CUBIN, Electrical and Electronic Engg. The University of Melbourne., www.ee.mu.oz.au/pgrad/bpw,March 2003.

[7] Y. Hou, H. Tzeng, and S. Panwar, "*A Generalized Max-Min Rate Allocation Policy and Its Distributed Implementation Using the ABR Flow Control Mechansim*" Proceedings of Infocom 98, San Francisco, 1998.

[8] L. Kalampoukas, A. Varma, and K. K. Ramakrishnan, "*An Efficient Rate Allocation Algorithm for ATM Networks Providing Maxmin fairness*" Proceedings of 6th IFIP Int'l Conf. High Performance Networking, 1995. pp. 143-154.

[9] L. Kalampoukas, "*Congestion Management in High Speed Networks*": University of California at Santa Cruz, April 1997.

[10] F. Kelly, A. Maulloo, and D. Tan, "*Rate control in communication networks: shadow prices, proportional fairness and stability*." Journal of the Operational Research Society, vol. 49, pp. 237-252, 1998.
[11] K. Laevens, P. Key, and D. McAuley, "*An ECN-based end-to-end congestion control framework: experiments and evaluation*" Microsoft Research MSR-TR2000 -104, October 2000.
[12] B. T. Doshi and H. Q. Nguyen, "*Congestion Control in ISDN Frame-Relay Networks*," AT&T Technical Journal, November/December 1988, pp. 35-46.
[13] S. Athuraliya, D. Laspsley, and S. H. Low, "*An enhanced random early marking algorithm for Internet flow control*" Proceedings of Infocom, Israel, 2000.
[14] D. M. Chiu and R. Jain, "*Analysis of the Increase and Decrease Algorithms for Congestion Avoidance in Computer Networks*" *Computer Networks*, vol. 17, pp. 1-14, 1989.