Proceedings of **ITC-CSCC 2000**, Pusan, Korea

End-to-End Congestion Control of High-Speed Gigabit-Ethernet Networks based on Smith's Principle*

Seung-Hyub Lee and Kwang-Hyun Cho**

Division of Electronic Engineering
School of Electrical Engineering and Automation
University of Ulsan
San-29 Moogu-2 Dong, Namgu,
Ulsan 680-792, Korea
Tel: +82-52-259-2734, Fax: +82-52-259-1686
**E-mail: ckh@uou.ulsan.ac.kr

Abstract: Nowadays, the issue of congestion control in high-speed communication networks becomes critical in view of the bandwidth-delay products for efficient data flow. In particular, the fact that the congestion is often accompanied by the data flow from the high-speed link to low-speed link is important with respect to the stability of closed-loop congestion control. The Virtual-Connection Network (VCN) in Gigabit Ethernet networks is a packet-switching based network capable of implementing cell-based connection, link-by-link flow-controlled connection, and single- or multi-destination virtual connections. VCN described herein differ from the virtual channel in ATM literature in that VCN have link-by-link flow control and can be of multi-destination. VCNs support both connection-oriented and connectionless data link layer traffic. Therefore, the worst collision scenario in Ethernet CSMA/CD with virtual collision brings about end-to-end delay. Gigabit Ethernet networks based on CSMA/CD results in non-deterministic behavior because its media access rules are based on random probability. Hence, it is difficult to obtain any sound mathematical formulation for congestion control without employing random processes or fluid-flow models. In this paper, an analytical method for the design of a congestion control scheme is proposed based on Smith's principle to overcome instability accompanied with the increase of end-to-end delays as well as to avoid cell losses. To this end, mathematical analysis is provided such that the proposed control scheme guarantees the performance improvement with respect to bandwidth and latency for selected network links with different propagation delays. In addition, guaranteed bandwidth is to be implemented by allowing individual stations to burst several frames at a time without intervening round-trip idle time.

1. Introduction

The major networking function for high-speed Gigabit Ethernet networks [1] is the design of feedback congestion control scheme with capability of dynamically allocating network bandwidth to source [2]. In the communication networks area, congestion control is defined as the state that a network cannot supply a proper level services to the users due to the mismatch between the network resources and the amount of admitted traffic [2], [4]. Thus, congestion in the communication networks is becoming a major cause of network performance degradation. In most cases, algorithms dealing with congestion control exhibit oscillations due to propagation delay in network systems and can even be unstable. In [3], an analytic method for the design of a congestion controller, which ensured dynamic performance along with fairness in bandwidth allocation, has been proposed. However, this algorithm requires the adaptive and the robust control technique in order to ensure asymptotic stability under different network conditions. In packet-switched case, packets - blocks of data of varying length - are transmitted over networks. Especially, data field sizes of Gigabit Ethernet frame are varying from 46 bytes to 1500 bytes. Hence, the characteristic model in Gigabit Ethernet networks is not represented a fixed form. The Virtual-Connection Network (VCN) in viewing the cell-level multiplexing is most effective for networks, mainly LANs and MANs, where medium propagation delays are relatively small [5]. These networks allow fast, reliable and low-cost implementation of the cell level handshaking protocols required by the link-by-link flow control. The link-by-link flow control mechanism described herein are some desirable design goals: 1) the peak link bandwidth can be achieved, 2) the flow control overhead is relatively small, and 3) the flow control tolerant to transient link failure.

Gigabit Ethernet networks based CSMA/CD result in nondeterministic behavior because it media access rules are based on random probability [6]. Hence, it is difficult to obtain any sound mathematical formulation for congestion control without employing random processes or fluid-flow models. Moreover, the difference between round-trip delay of above-mentioned research paper [4] in case of ATM and proposed end-to-end delay in case of Gigabit Ethernet is referred to generally as *model uncertainty*. This paper focuses on the important stability issue of traffic control with end-to-end delay. With this approach, stability analysis for the controlled network becomes feasible with the knowledge end-to-end delay. In this paper, classical control theory and *Smith's principle* [9] are adopted as key

---

* This work was supported by the development program for the exemplary school in information and communications from the Ministry of Information and Communication (MIC).

tools for designing congestion control law in high-speed data networks. Also, the *Kharitonov's theorem* [8], [9] are used for discriminating stability for a given interval end-to-end delay of Gigabit Ethernet networks.

The objective of this paper is to find a control law for Gigabit Ethernet networks so that the network bandwidth is fully utilized without incurring network congestion. Especially, this paper focuses on the important stability issue of traffic control with end-to-end delay. The proposed control scheme has been applied to Gigabit Ethernet networks as well as high-speed communication networks. A major portion of this paper is devoted to an analytical method for the design of a congestion control in high-speed Gigabit Ethernet networks. The control algorithm is developed starting from an accurate mathematical model. Therefore, its stability is demonstrated.

The paper is structured as follows: Section 2 describes the Gigabit Ethernet queue model and the Gigabit Ethernet network model. In section 3, its analysis and design stability of the control law for the Gigabit Ethernet network model are addressed. And conclusions are given in Section 4.

2. The Gigabit Ethernet Network Model

2.1. The Gigabit Ethernet Queue Model

The Fig. 1 shows two connections sharing one outgoing link. Each output link has a common FIFO queue for all virtual connections sharing. Let $x_j(t)$ be the queue level associated with link l_j. By writing flow conservations, the level of occupancy $x_j(t)$, starting at $t=0$ with $x_j(0)=0$, is

$$x_j(t)=\int_0^t\sum_{i=1}^n u_{ij}(\tau-T_{ij})-\int_o^t d_j(\tau)d\tau\,. \qquad (1)$$

Where n is the number of connections sharing the queue, $u_{ij}(t)$ is the inflow rate due to the *i-th* connection, T_{ij} is the propagation dalay from the *i-th* source to *j*-th queue and $d_j(t)$ is the rate of packets leaving the *j-th* queue, that is the avaiable per-link bandwidth. It is assumed that the propagation dalay is dominant compared to other delay (processing, queuing, etc).

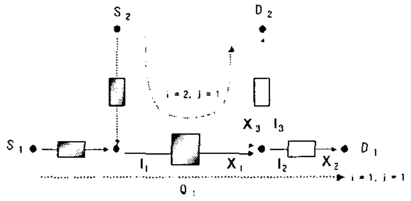

Figure 1. Scheme of connection (S_1, D_1) and (S_2, D_2) sharing link l_1 and queue Q_1.

2.2. The Gigabit Ethernet Network Model

We assume that a connection establishes a VCN with per-SPQ at the node output links. A particular link along the VCN will be the bottleneck queue for the flow, whereas other queues encountered by the flow will be empty. Thus, we take into consideration system model of n VCN sharing common on FIFO queue with the controller transfer function $G_{ge}(s)$.

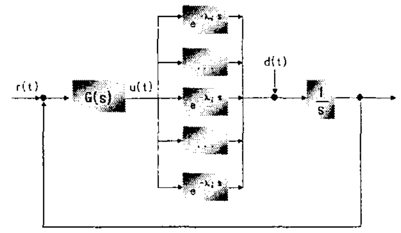

Figure 2. Gigabit Ethernet Networks model of n VCN sharing common FIFO queue.

The network employs a store-and-forward packet switching service where users are serviced without prior reservation. Mainly following the notation reported in [3], the network consists of N switching nodes and L communication links. For each node $i \in N$, let $O(i) \subset L$ denote the set of its outgoing links. Each link i is characterized by its transmission capacity $c_i = 1/t_i$ (packets/sec), where t_i is the transmission time of a packet, and a propagation delay of $1/t_{di}$ sec. The network is assumed by homogeneous. Finally, it is worth noting that, in high-speed networks, the bandwidth delay product $c_i t_{di}$ (in pipe cells) represents a large number of cells "in flight" on the transmission link. The network traffic is represented by source-destination pairs $(S,D) \in N \times N$. To each (S,D) connection is associated a Virtual-Connection Network (VCN) mapped on the path $p(S,D)$. The path contains one node for every network node and one directed link $e=(a,b)$ for every communication link from node a to node b. Therefore a VCN i is specified by the sequence of links $e^i_1\ e^i_2\ e^i_3 \ldots e^i_n$ that traverses as it goes throught the networks. The feedback scheme in [7] requires that each traffic source in enhanced flow control send one control frame (PAUSE frame) every N data frames. Each node i has a congestion controller which periodically computes PAUSE time for each outgoing link $j \in O(i)$. Some indication that some link through which a stream passes is congested passed back through switches. Each node encountered by the PAUSE frame along VCNs, stamps the computed value for input rate on the PAUSE frame only if this value results to be less than the rate already stored.

The dynamics of the bottleneck queue level in response to the source input rate is a *single-input-single-output dynamics*. Fig. 3 shows the Gigabit Ethernet networks model consisting of:

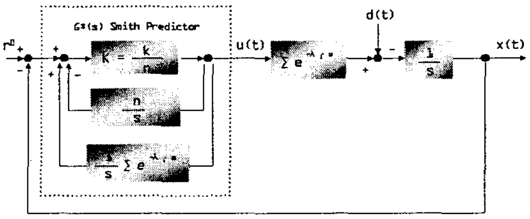

Figure 3. Queue model in Gigabit Ethernet networks.

1) The bottleneck FIFO queue modeled in the Laplace domain by the integrator $1/s$,
2) The available per-link bandwidth $d(t)$ modeled as a disturb,
3) The round trip delay T_i, $(i=1,2,\ldots,n)$ of each virtual connection sharing the queue,
4) The end-to-end delay $\lambda_i = \alpha_i T_i$, $(i=1,2,\ldots,n)$, where $\alpha_i \in [1/2, 1]$,
5) The controller transfer function $G_{ge}(s)$ in Gigabit Ethernet networks,
6) The set point $r(t)$.

3. The Control law for Network Model

In this section, it is to design a feedback control law $u(t)$ for the input rate of each VCN such that each network queue level $x(t)$ satisfies the following conditions.

(1) *Stability condition* : $x(t) \le r^0$ (2)

where r^0 is the queue capacity, which guarantees that this queue is always bounded, i.e. no packet loss.

(2) *Full link utilization* : $x(t) > 0$ for $t \ge T$ (3)

where T represents the tasnsient time after the starting of network operation.

(3) *Uncertainty of end-to-end delay factor* α_i : (4)

$\alpha_i \in [1/2, 1]$, where $i = 1,2,\ldots,n$

Due to the large delays inside the feedback loop, queue dynamics might exhibit oscillations, and even become unstable. In WAN, it is of important that round trip *delays* are mostly determined by propagation delay. We will analyze under the round-trip assumption: $S = timeout = W \times TRANSP$, where W is window size and TRANSP is transmission time of typical packet. And we assume that the boundary of RTT is [0.01, 10000]. The RTT=10,000 can coresspond to a cross-country connection ($\approx 24{,}000km$ round-trip) through a WAN with a typical bottleneck link capacity of 155Mb/sec [5].

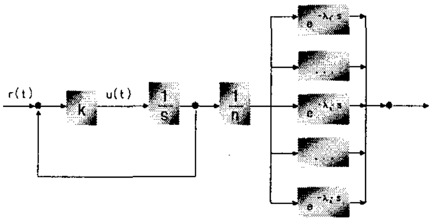

Figure 4. Model of the desired input-output dynamic.

Now we derive the controller that gives the desired input-output dynamics reported in Fig. 4. It is a first order system with a sum of delays in cascade. By letting the set point $r(t)$ to be the step function $r^0 \cdot 1(t)$. Thus, the output does not overshoot.

Proposition 1: The system shown in Fig. 4, where the reference signal function is the step function $r(t) = r^0 \cdot 1(t)$, satisfies the stability condition $x(t) \le r^0$ if and only if end-to-end delay time $\lambda_i \rangle 0$ and $k \rangle 0$.

Proof. The Laplace transform of the output $x(t)$ in response to the set point $r^0 \cdot 1(t)$ is:

$$X(s) = \frac{r^0}{n} \cdot \frac{1}{s} \cdot \frac{1}{\frac{s}{k}+1} \cdot \sum_{i=1}^{n} e^{-\lambda_i s} . \tag{5}$$

If we substitute the relation

$$X_{partial}(s) \cdot \sum_{i=1}^{n} e^{-\lambda_i s} \leftrightarrow \sum_{i=1}^{n} x_{partial}(t-\lambda_i)$$

into Eq. (5), we find that.

$$x(t) = \frac{r^0}{n} \cdot \sum_{i=1}^{n} \left[1 - e^{-k(t-\lambda_i)}\right] \cdot 1[t-\lambda_i] \;\le\; r^0 . \tag{6}$$

Therefore, in Eq. (6), when $\lambda_i \rangle 0$, the exponential expression decays, and we say the step response is stable. This completes the proof.

Proposition 2: The transfer functions $X(s)/R(s)$ of the system shown in Fig. 2 and Fig. 4 are made euqal by controller described by the transfer function

$$G_{ge}(s) = \frac{k/n}{1 + \frac{k/n}{s}\left[n - \sum_{i=1}^{n} e^{-\lambda_i s}\right]} . \tag{7}$$

Proof. By equating the transfer functions of the systems reported in Fig. 2 and Fig. 4

$$\frac{\frac{1}{s} \cdot G_{ge}(s) \sum_{i=1}^{n} e^{-\lambda_i s}}{1 + \frac{1}{s} \cdot G_{ge}(s) \sum_{i=1}^{n} e^{-\lambda_i s}} = \frac{\frac{k}{s}}{1 + \frac{k}{s}} \cdot \frac{1}{n} \sum_{i=1}^{n} e^{-\lambda_i s} . \tag{8}$$

With respect to $G_{ge}(s)$, we arrange in Eq. (8). Then, the controller (7) is derived. We design a controller based on Smith's principle. Therefore, Fig. 2 shows the controller of Gigabit Ethernet Networks.

By looking at Fig. 2, it is easy to write the rate control equation in time domain, that is

$$y(t) \equiv r(t) - x(t) - n \int_0^t u_i(\tau) \cdot d\tau + \sum_{i=1}^{n} \int_0^{t-\lambda_i} u_i(\tau) \cdot d\tau \tag{9}$$

In case of $r(t) = r^0 \cdot 1(t)$, we get

$$y(t) = r^0 - x(t) - n \int_0^t u_i(\tau) \cdot d\tau - \sum_{i=1}^{n} \int_{t-\lambda_i}^{0} u_i(\tau) \cdot d\tau$$

$$= r^0 - x(t) - \sum_{i=1}^{n} \int_{t-\lambda_i}^{t} u_i(\tau) \cdot d\tau$$

From Fig. 2, we can conclude that the inflow rate $u_i(t) = y(t) \cdot \left[\frac{k}{n}\right]$ due to the i-th connection is

$$u_i(t) = \frac{k}{n}\left[r^0 - x(t) - \sum_{i=1}^{n} \int_{t-\lambda_i}^{t} u_i(\tau) \cdot d\tau\right] . \tag{10}$$

This equation can be intutively as follows: the computed input rate is proportional, through the coefficient k/n, to available queue room (*buffer capacity*) $r^0 - x(t)$ decreased by the number of cells released by each connection during the last corresponding end-to-end delay λ_i, that is, the sum of "*in flight*" cells of all VCNs sharing the queue [4]. Hence, To be guarantee full utilization of VCN in Gigabit Ethernet networks should be that the each queue capacity for VCN is at least euqal to the number of "*in flight*" cells contained in a connection with end-to-end delay $\sum_{i=1}^{n} \frac{T_i}{n} + \tau$.

Proposition 3: The characteristic equation of the controller transfer function $G_{ge}(s)$ in Fig. 3 satisfies the stability condition if and only if $T_i \in [0.01,10000]$.

Proof. By (p,q) Pade approximant, we obtain an approimate values

$$\sum_{i=1}^{n} e^{-\lambda_i s} \approx \frac{1}{1+\lambda_1 s} + \frac{1}{1+\lambda_2 s} + \ldots + \frac{1}{1+\lambda_n s}. \qquad (11)$$

Eq. (11) means that it is sumed by each link end-to-end delay, where is $\lambda_i = \alpha_i T_i \in [0.005,10000]$ by means of $\alpha_i \in [1/2,1]$, $T_i \in [0.01,10000]$. For illustrative purpose, we consider in case of VCN $i = 1$

$$G_{ge}(s) = \frac{1}{\frac{n}{k} + \frac{1}{s}\left[n - \frac{1}{1+\lambda_1 s}\right]}. \qquad (12)$$

By applying the Kharitonov's theorem in Eq. (12), gives the corresponding interval polynomials as followed.

$$p(s,q,r) = \frac{q_2 s^2 + q_1 s}{r_2 s^2 + r_1 s^1 + r_0}, \qquad (13)$$

We consider parameter uncertainties $r_2 = n\lambda_1$, $r_1 = n(1 - k\lambda_1)$, $r_0 = -nk - 1$ and $q_2 = k\lambda_1 + k$, $q_1 = k$. Thus, the coefficients vary inindependent intervals $r_2 \in [0.005n,10000n]$, $r_1 \in [n(1-k0.005), n(1-k10000)]$, $r_0 = -nk - 1$ and $q_2 \in [1.005k,10001k]$, $q_0 = k$.

By taking all combinations of the $N_i(s)$ and $D_k(s)$, we obtain the sixteen Kharitonov plants in case of $n = 1$

$$P_{ik}(s) = \frac{N_i(s)}{D_k(s)}, \text{ for } i,k = 1,2,3,4. \qquad (14)$$

Then the Kharitonov polynomial coressponding to $N(s)$ and $D(s)$ are

$$\begin{aligned} &N_1(s) = ks + 10001ks^2, \quad N_2(s) = ks + 10001ks^2, \\ &N_3(s) = ks + 1.005ks^2, \quad N_4(s) = ks + 1.005ns^2. \end{aligned} \qquad (15)$$

$$\begin{aligned} &D_1(s) = -nk - 1 + n(1-k0.005)s + 10000ns^2, \\ &D_2(s) = -nk - 1 + n(1-k10000)s + 10000ns^2, \\ &D_3(s) = -nk - 1 + n(1-k0.005)s + 0.005ns^2, \\ &D_4(s) = -nk - 1 + n(1-k10000)s + 0.005ns^2. \end{aligned} \qquad (16)$$

Furthermore, the family of closed loop characteristic polynomial $\Delta(s,k) = D(s) + N(k)$. Therefore, the value of $k > 0.001$ is satisfied by the stability condition. We may now state the closed-loop polynomial expressed in Eq. (15) and (16) in serveral equivalent. Therefore, the set of all proportional gain k values is satisfied by the stability condition.. As previously stated in this paper, the accurate explanation in Eq. (7) requires the adaptive and the robust control technique in order to ensure asymptotic stability under different network conditions. Note that, due to be concentrated upon Gigabit Ethernet networks of simplification, it is not dealt with an above-mentioned problem. Thus, we show the stability for the end-to-end delay factor, that is α_i, only. This is satisfied by the proof.

4. Conclusions

The Smith's principle has been proposed to model the dynamics of high-speed Gigabit Ethernet networks. The properties of the proposed control law have been demonstrated based on the characteristic of linear system. Especially, the control scheme reveals better performance for the steady-state response. Henceforth, we have developed an analytical method for the transient and steady state behavior of the system response. However, there still remains a problem developing a method applicable to the uncertain network models incorporating end-to-end delay.

References

[1] Howard Frazier, The 802.3z Gigabit Ethernet Standard, IEEE Network, May-June 1998, Vol. 12, Issue 3, pp.6-7.

[2] Ray Jain, Congestion Control and Traffic Management in ATM Networks: Recent Advances and a Survey, Computer Networks and ISDN Systems, October 1996, vol.28, no.13, pp.1723-1738.

[3] L. Benmohamed, S. M. Meerkov, Feedback Control of Congestion in Packet Switching Networks: The Case of Single Congested Node, IEEE/ACM Trans. on Networking, 1993, vol.1, no.6, pp.693-708.

[4] Saverio Mascolo., Congestion control in high-speed communication networks using the Smith principle, Automatica 35, pp. 1921-1935, 1999.

[5] P. Varaiya and Jean Walrand, High-Performance Communication Networks, Morgan Kaufmann Publishers, 2000.

[6] Stephen Saunders, Data Communications Gigabit Ethernet Handbook, McGraw-Hill, 1998.

[7] Seung-Hyub Lee and Kwang-Hyun Cho, Congestion Control for Virtual-Connection Using a Smith Predictor in Hgh-Speed Gigabit Ethernet Networks, KORUS2000, Ulsan, Korea, June, 2000

[8] B. Ross Barmish, Christopher V. Hoolt, Frank J. Kraus and Roberto Tempo, Extreme Point Results for Robust Stabilization of Interval Plants with First Order Compensators, IEEE Transaction on Automatic Control, Vol. 32, No. 6, pp.707-709, June 1992.

[9] Gene F. Franklin, J. David Powell, Addas Emami-Naeini., *Feedback Control of Dynamic Systems*, Addison-Wesley, 1994.