Diffusion-Type Autonomous Decentralized Flow Control for Open Networks

Chisa Takano† and Masaki Aida‡
† Traffic Engineering Division, NTT Advanced Technology Corporation
2-4-15, Naka-cho, Musashino-shi 180-0006, Japan
chisa@m.ieice.org
‡ NTT Information Sharing Platform Laboratories, NTT Corporation
Musashino-shi 180-8585, Japan
aida.masaki@lab.ntt.co.jp

Abstract

In our previous studies, we have proposed diffusion-type flow control for high-speed networks. The reason for proposing this type of control is that decision-making at each node leads to high performance of the whole network although each individual node handles its local traffic flow only on the basis of the information it is aware of from its immediately adjacent nodes. Our control method exhibits the desired stability and adaptability of network performance even when the network has an heterogeneous configuration and the capacity of links is dynamically changed. However, in order to make a comparison with other flow control methods under the same conditions, our evaluation used a closed network model, in which the number of packets was unchanged. This paper investigates the performance of our flow control method using an open network model, in order to show that our flow control method is still effective in more realistic networks. We identify the key issues associated with our flow control in an open network model, and demonstrate a solution.

1 Introduction

Delay in communication networks consists of processing delay and propagation delay. Processing delay is the delay experienced prior to packet transmission and in waiting for transmission at nodes. This is reduced if the processing speed of the nodes is increased. Propagation delay is the delay experienced by a packet during propagation over a link and is determined by the link length and the speed of light. Unlike processing delay, the propagation delay is fixed even if the processing speed of the nodes is increased. So, if the processing speed of the nodes is high enough, propagation delay becomes the dominant factor in the total delay.

In a high-speed network, it is impossible to implement time-sensitive control based on collecting global information about the whole network because the state of a node varies rapidly in accordance with its processing speed although the propagation delay is constant. If we allow sufficient time to collect network-wide information, the data so gathered is too old to apply to time-sensitive control. In this sense, each node in a high-speed network is isolated from information about the state of other nodes or of the overall network.

In addition, when the propagation delay is dominant, at any instant a large amount of data is being propagated over the links in a high-speed network. Since we can only control the packets stored in nodes, and not those being propagated, high-speed networks contain many packets which cannot be controlled because they are being propagated over links. In this situation, since the control delay greatly affects the network performance, a very rapid control mechanism is required, in which the control delay is as short as possible.

This paper focuses on a flow control mechanism for high-speed networks. From the above considerations, the technique used for our flow control method should satisfy the following requirements:

- With regard to the collection of information, it must be possible to collect the information used in the control method.
- With regard to the delay in applying control, the control should take effect immediately.

There are many other papers reporting studies on the optimization of flow control problems in a framework of solving linear programs [1, 2, 3, 4, 5]. These studies assume the collection of global information about the network, but

Proceedings of the 24th International Conference on Distributed Computing Systems Workshops (ICDCSW'04)

it is impossible to realize such a centralized control mechanism in high-speed networks. In addition, solving these optimization problems requires enough time to be available for calculation, and so it is difficult to apply these methods to decision-making in a very short time-scale. So, in a high-speed network, the principles adopted for time-sensitive control are inevitably those of autonomous decentralized systems [6, 7].

Decentralized flow control by end hosts, including TCP, is widely used in current networks, and there is much research in this area [4, 5, 8]. However, since end-to-end or end-to-node control cannot be applied to decision-making in a time-scale shorter than the round-trip delay, it is inadequate for application to support decision-making in a very short time-scale. In low-speed networks, a control delay of the order of the round-trip time (RTT) has a negligible effect on the network performance. However, in high-speed networks, the control delay greatly affects the network performance. This is because the RTT becomes large relative to the unit of time determined by node's processing speed, although the RTT is itself unchanged. This means that nodes in high-speed networks experience a relatively larger RTT, and this causes an increase in the sensitivity to control delay. To achieve rapid control in a shorter time scale than the RTT, it is preferable to apply control by the nodes rather than by the end hosts.

We therefore have considered a control mechanism in which the nodes in a network handle their local traffic flows themselves, based only on the information they are aware of. This mechanism can immediately detect a change in the network state around the node and exert quick decision-making. However, decision-making at a local node leads to action suitable for the local performance of the networks, but it is not guaranteed that the action is appropriate for the overall network-wide performance. So, the implementation of decision-making at each node cannot lead to optimum performance for the whole network.

Bartal *et al.* [1] studied global optimization of flow control using local information. The motivation for their work was to enable the distributed routers in high-speed networks to make decisions on flow control as quickly as possible, and they studied the problem in a framework of solving linear programs by distributed agents. Although this motivation is similar to that of our work, their study assumed that the distributed agents can obtain detailed information about networks if we allow them sufficient time to gather it. As stated above, our standpoint is based on the fact that we cannot obtain detailed, useful and up-to-date information about the whole network in a high-speed network environment.

In our previous studies, we investigated the behavior of local packet flows and the global performance when a node is congested, and proposed a diffusion-type flow control [9]. In addition, we investigated the stability and adaptability of the network performance when the capacity of a link is changed, by using network models with homogeneous [10] and inhomogeneous [11] configurations. Diffusion-type flow control provides a framework in which the implementation of decision-making of each node leads to high performance for the whole network. The principle of our flow control model can be explained through the following analogy [12, 13].

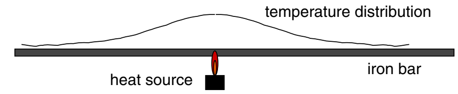

Figure 1. Example of thermal diffusion phenomena.

When we heat a point on a cold iron bar, the temperature distribution forms a normal distribution and heat spreads through the whole as a diffusion phenomenon (Fig. 1). In this process, the action in a minute segment of the iron bar is very simple; heat flows from the higher temperature side towards the lower temperature side. The rate of heat flow is proportional to the temperature gradient. There is no communication between two distant segments of the iron bar. Although each segment acts autonomously, based on its local information, the temperature distribution of the whole iron bar exhibits orderly behavior. In diffusion-type flow control, each node controls its local packet flow, which is proportional to the difference between the number of packets in the node and those in an adjacent node. Then the distribution of the number of packets in all the nodes in the network becomes uniform over time. In this control mechanism, the state of the whole network is controlled indirectly through the autonomous action of each node.

In our previous research on diffusion-type flow control [9, 10, 11], we used a closed network model because with this model the number of packets in the network is unchanged and it is appropriate for comparison with other control mechanism under the same conditions. In this paper, we use an open network model for evaluation in order to show that diffusion-type flow control is applicable to more realistic networks. We identify the key issues in applying our flow control to an open network and show a solution for them.

2 Preliminary Description of Diffusion-Type Flow Control

We assume that a target flow has a static route. In the case of Internet-based networks, to guarantee end-to-end Quality of Service (QoS) of a flow, the QoS sensitive flow has a static route (e.g. RSVP). In addition, we assume all routers in the network can employ per-flow queueing for the all target flows.

Figure 2 shows the interactions between nodes (routers)

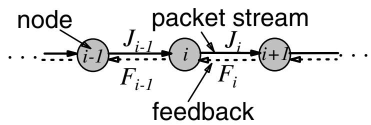

Figure 2. Node interactions in our flow control model.

in our flow control method, using a network model with a simple 1-dimensional configuration. All nodes have two incoming and two outgoing links, for a one-way packet stream and for feedback information, that is, node i (i = 1, 2, ...) transfers packets to node $i+1$ and node $i+1$ sends feedback information to node i. For simplicity, we assume that packets have a fixed length in bits.

All nodes are capable of receiving feedback information from, and sending it to, adjacent downstream and up-stream nodes, respectively. Each node i can receive feedback information sent from the downstream node $i+1$, and can send feedback information about node i itself to the upstream node $i-1$.

When node i receives feedback information from downstream node $i+1$, it determines the transmission rate for packets to the downstream node $i+1$ using the received feedback information, and adjusts its transmission rate towards the downstream node $i+1$. The framework for node behavior and flow control is summarized as follows:

- Each node i autonomously determines the transmission rate J_i based only on information it is aware of, that is, the feedback information obtained from the downstream node $i+1$ and its own feedback information.
- The rule for determining the transmission rate is the same for all nodes.
- Each node i adjusts its transmission rate towards the downstream node $i+1$ to J_i.
(If there are no packets in node i, the packet transmission rate is 0.)
- Each node i autonomously creates feedback information according to a predefined rule and sends it to the upstream node $i-1$. Feedback information is created periodically with the fixed interval τ_i.
- The rule for creating the feedback information is the same for all nodes.
- Packets and feedback information both experience the same propagation delay.

As mentioned above, the framework of our flow control model involves both autonomous decision-making by each node and interaction between adjacent nodes. There is no centralized control mechanism in the network. More precisely, it is impossible to realize centralized control in a high-speed network environment.

Next, we will explain the details of the diffusion-type flow control.

The transmission rate $J_i(\alpha, t)$ of node i at time t is determined by

$$J_i(\alpha,t) = \max(0, \min(L_i(t), \tilde{J}_i(\alpha,t))), \quad \text{and} \qquad (1)$$

$$\tilde{J}_i(\alpha,t) = \alpha\, r_i(t-d_i) - D_i\,(n_{i+1}(t-d_i) - n_i(t)), \qquad (2)$$

where L_i denotes the value of the link capacity from node i to node $i+1$, $n_i(t)$ denotes the number of packets in node i at time t, $r_i(t-d_i)$ is the target transmission rate specified by the downstream node $i+1$ as feedback information, and d_i denotes the propagation delay between node i and node $i+1$.

In addition, $r_i(t-d_i)$ and $n_{i+1}(t-d_i)$ are notified every fixed period τ_{i+1} from the downstream node $i+1$ with a propagation delay d_i. Parameters α (≥ 1) and D_i (> 0) are constants and are the flow intensity multiplier and the diffusion coefficient, respectively.

The diffusion coefficient is chosen to be inversely proportional to the propagation delay [11] as follows:

$$D_i = D\,\frac{1}{d_i} \propto (d_i)^{-1}, \qquad (3)$$

where D is a positive constant.

The feedback information, $\mathbf{F}_i(t)$, created every fixed period τ_i by node i consists of two quantities as follows:

$$\mathbf{F}_i(t) = (r_{i-1}(t), n_i(t)). \qquad (4)$$

Node i notifies this to the upstream node $i-1$ with a period of $\tau_i = d_{i-1}$. Here, the target transmission rate is determined as

$$r_{i-1}(t) = J_i(1,t). \qquad (5)$$

Moreover, the packet flow $J_i(t)$ in node i is renewed whenever feedback information arrives from the downstream node $i+1$ (with a period of $\tau_{i+1} = d_i$).

To assist an intuitive understanding, we can briefly explain the physical meaning of the diffusion-type flow control. We replace i with x and apply continuous approximation. Then the propagation delay becomes $d_i \to 0$ for all i and the packet flow (2) is expressed as

$$\tilde{J}(\alpha,x,t) = \alpha\, r(x,t) - D\,\frac{\partial n(x,t)}{\partial x}, \qquad (6)$$

and the temporal variation of the packet density $n(x,t)$ may be expressed as a diffusion-type equation,

$$\frac{\partial n(x,t)}{\partial t} = -\alpha\,\frac{\partial r(x,t)}{\partial x} + D\,\frac{\partial^2 n(x,t)}{\partial x^2}, \qquad (7)$$

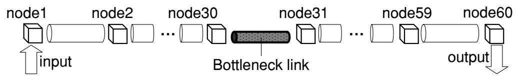

Figure 3. Simulation Model

using the continuous equation

$$\frac{\partial n(x,t)}{\partial t} = -\frac{\partial \tilde{J}(\alpha, x, t)}{\partial x}. \tag{8}$$

That is, our method aims to perform flow control using the analogy of a diffusion phenomenon. We can expect excess packets in a congested node to be distributed to the whole network and normal network conditions to be restored after some time.

3 Applying Diffusion-Type Flow Control to an Open Network Model

We have investigated the key issues involved in applying the diffusion-type flow control mechanism described in the previous section to an open network model. Figure 3 shows the open network model, with 60 nodes, which was used in the simulations. Although the 1-dimensional model looks simple, it represents a part of a network and describes the target end-to-end flow for a path extracted from the whole network. We represent the lengths of links as their delays, and determine the length of each link according to a log-normal distribution, in advance, where the mean delay in the network model is 1.0 [unit of time] and the variance is 5.2 [unit of time^2].

The capacity of all normal links is $L_i = 100$ [packets/unit of time] $(i \neq 30)$ and the capacity of the bottleneck link (between nodes 30 and 31) is $L_{30} = 50$ [packets/unit of time] (that is, this link has half the bandwidth of the other links). To investigate the stability under congestion, in addition to the above conditions, we set the initial condition for the congested node 30 as follows.

- Number of packets in node 30 at time $t = 0$: 400.
- The other 5600 packets are distributed randomly in the other nodes and on other links.

We set the values of parameters as:

$$D = 0.1, \quad \text{and} \quad \alpha_i = 1.0. \tag{9}$$

At the ingress point of the network, the rate of packet flow is regulated by $r_0(t)$ which is specified by node 1.

The left and right sides of Fig. 4 show the calculated results for the total number of packets being propagated on all links and the total number of packets stored in all nodes

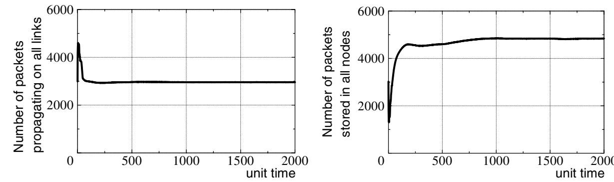

Figure 4. Total number of packets propagating on all links and stored in all nodes for the diffusion-type flow control described in Sec.3.

of the network. The horizontal axes denote the simulation time and the vertical axes denote the total number of packets propagated on links and stored in nodes, respectively. The total number of packets propagating on links indicates the transmission efficiency of the network, and we call this the total throughput.

From the quantitative point of view, for the case where the link capacity of the bottleneck links $L_{30} = 50$, the maximum value of the sustainable total throughput (the maximum number of packets propagated stably on the links) is about 3000, i.e., 50 packets/link $\times$ 59 links. Thus, the diffusion-type flow control achieves about 100% of the maximum value of the total throughput and its value is stable, but the total number of packets stored in nodes is large. A large number of packets stored in nodes is undesirable because it causes an increase in delay and packet loss.

This phenomenon was not observed in the closed network model. We can see from these results that it is necessary to regulate the packet flow carefully, especially in the open network model in which the total number of packets in the network changes with time.

4 Diffusion-type Flow Control Model Adapted to Open Network Model

4.1 Extension of Flow Control model

In this section, we consider two extensions to the diffusion-type flow control model, to regulate the packet flow in the network. In the first model, only the ingress node uses its minimum value when calculating the packet flow rate. On the other hand, in the second model, all nodes regulate the packet flow rate less than or equal to the minimum value of link capacity among all the downstream links.

In both models, node i generates feedback information $\mathbf{F}_i(t)$ as

$$\mathbf{F}_i(t) = (r_{i-1}(t), n_i(t), \ell_i(t)), \tag{10}$$

and notifies the information to the upstream node $i-1$, where $\ell_i(t)$ denotes information about the maximum value

Proceedings of the 24th International Conference on Distributed Computing Systems Workshops (ICDCSW'04)
0-7695-2087-1/04 $20.00 © 2004 IEEE

of available capacity as

$$\ell_i(t) = \min(L_i, \ell_{i+1}(t - d_i)). \qquad (11)$$

Model 1 When node i receives feedback information from downstream node $i+1$, it determines the transmission rate for packets to the downstream node $i+1$, using the received feedback information, and adjusts its transmission rate towards the downstream node $i+1$. Then, the packet transmission rate $J_i(\alpha, t)$ is determined by Eq. (1) and (2) in Model 1. The feedback information $r_{i-1}(t)$ is determined as Eq. (5).

This model calculates the transmission rate $J_0(\alpha, t)$ for packets to flow at the ingress to the network as follows:

$$J_0(\alpha, t) = \min(\ell_0(t), r_0(t)). \qquad (12)$$

Model 2 In Model 2, the packet transmission rate $J_i(\alpha, t)$ and the feedback information $r_{i-1}(t)$ are determined as

$$J_i(\alpha, t) = \max(0, \min(\ell_i(t), \tilde{J}_i(\alpha, t))), \qquad (13)$$

$$r_{i-1}(t) = \max(0, \min(\ell_i(t), \tilde{J}_i(1, t))) \quad (i > 0), \qquad (14)$$

where the transmission rate $J_0(\alpha, t)$ is determined as Eq. (12).

A remarkable feature of Model 2 is that the maximum values of $J_i(\alpha, t)$ and $r_{i-1}(t)$ are bounded by $\ell_i(t)$.

The difference between Model 1 and 2 is that the maximum values of $J_i(\alpha, t)$ and $r_{i-1}(t)$ are bounded by $\ell_i(t)$ at all nodes in the network in Model 2, although they are bounded by $\ell_i(t)$ only at the ingress point in Model 1.

4.2 Evaluation for the two Extended Flow Control Models

We compare the network performance of the diffusion-type flow control models 1 and 2 described in Sec. 4.1 using the same network model as was used in Sec. 3, and we discuss the stability of flow control models in a high-speed network environment.

Figures 5 and 6 show results for the flow control models 1 and 2, where the horizontal axes denote the simulation time and the vertical axes denote the total number of packets in all links or nodes in the network.

We can see in Figs. 5 and 6 that the total number of packets in all links becomes stable at a high level of performance in both models.

However, there is a difference in the number of packets stored in nodes. In Model 2, the total number of packets in all nodes is smaller than in the previous flow control model but become stable and does not decrease below 3000. In Model 1, on the other hand, the total number of packets in all nodes decreases rapidly and falls to zero with time.

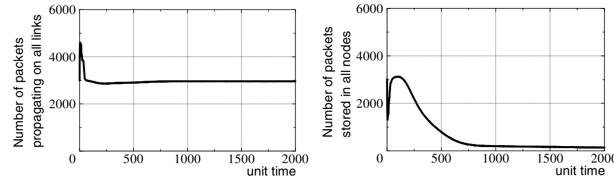

Figure 5. Total number of packets propagating on links and stored in nodes for Model 1.

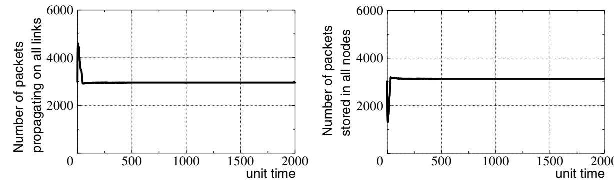

Figure 6. Total number of packets propagating on links and stored in nodes for Model 2.

Next, we compare the behavior of the packet distribution of the each node, in order to investigate the performance of Model 1 and 2. Figures 7 and 8 show the simulation results for Model 1 and Model 2, respectively. The horizontal axis of each graph denotes the node ID, the vertical axis denotes the number of packets stored in the node, and t denotes the simulation time (initially $t = 0$).

We can see from Fig. 7 that in Model 1 the number of packets stored in nodes decreases with time and is smoothly distributed over the whole network. For Model 2, on the other hand, the number of packets stored in each node is smoothly distributed over a part of the network with time and remains uneven over the whole network.

This difference arises because in Model 2 all nodes calculate the packet transmission rates using the minimum value of link capacity from among the downstream links. This restriction, which is imposed on calculating the rate, is too severe in Model 2. So, the transmission rates (determined by Eq. (13)) calculated in Model 2 differ greatly from the ideal rates (determined by Eq. (2)) which cause the diffusion phenomenon. Our results also show that the rule for determining the packet rate at a node, which is the most important feature of our flow control, is crucially important for stability or adaptability of the whole network performance.

5 Conclusions

This paper has investigated the performance of diffusion-type flow control applied to an end-to-end flow. The diffusion-type flow control proposed in our previous studies is an autonomous decentralized control, in which each node in the networks handles its local traffic flow itself, based

Proceedings of the 24th International Conference on Distributed Computing Systems Workshops (ICDCSW'04)
0-7695-2087-1/04 $20.00 © 2004 IEEE

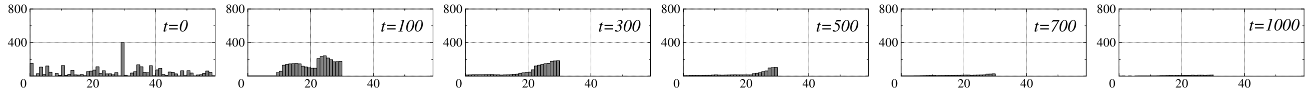

Figure 7. Distribution of the number of packets stored in each node for Model 1.

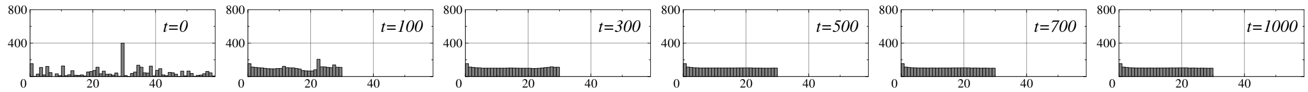

Figure 8. Distribution of the number of packets stored in each node for Model 2.

only on the information it is aware of. Our previous studies showed that, in the case of a closed network model, the action of each node leads to desirable performance of the whole network.

In order to show that our flow control method is effective for an end-to-end flow as well, we simulated the application of our flow control to an open network, and identified an issue. As a result of simulation, we found that regularization of the flow at the ingress to the network is important. As a solution, we introduced an additional item of feedback information and use it for flow regularization at the ingress point, and simulated this using the two extended models.

Our results show that the rule for determining the packet rate at a node, which is the most important feature of our flow control method, is also crucially important for the stability and adaptability of the whole network performance as well.

In Model 1, the half-life period of the peak for the number of packets stored in nodes is about only 400 units of time. For example, if the mean propagation delay of the links is 50 μs (corresponding to about 10 km in length), the state of the whole network is completely restored in only 20 ms. As for the number of packets in a congested node, the half-life period of the peak is about 1.5 ms.

We are interested in the relationship between the length of the period for congestion restoration and the values of parameters D and α in diffusion-type flow control. This issue will be the subject of further study.

References

[1] Y. Bartal, J. Byers, and D. Raz, "Global optimization using local information with applications to flow control," in *Proc. of the 38th Ann. IEEE Symp. on Foundations of Computer Science*, October 1997.

[2] S. H. Low and D. E. Lapsley, "Optimization flow control-I: basic algorithm and convergence," *IEEE/ACM Transactions on Networking*, vol. 7, no. 6, pp. 861–874, 1999.

[3] K. Kar, S. Sarkar, and L. Tassiulas, "A simple rate control algorithm for maximizing total user utility," in *Proc. of IEEE INFOCOM 2001*, pp. 133–141, 2001.

[4] J. Mo and J. Walrand, "Fair end-to-end window based congestion control," *IEEE/ACM Trans. Networking*, vol. 8, no. 5, pp. 556–567, October 1999.

[5] S. Kunniyur and R. Srikant, "A decentralized adaptive ECN marking algorithm," in *Proc. of IEEE GLOBECOM '00*, pp. 1719–1723, 2000.

[6] M. Aida and K. Horikawa, "Stability analysis for global performance of flow control in high-speed networks based on statistical physics," *IEICE Transactions on Communications*, vol. E82-B, no. 12, pp. 2095–2106, December 1999.

[7] K. Horikawa, M. Aida, and T. Sugawara, "Traffic control scheme under the communication delay of high-speed networks," in *Proc. of Int. Conf. on Multi-Agent Systems (ICMAS) '96*, pp. 111–117, 1996.

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

[9] M. Aida and C. Takano, "Stability of autonomous decentralized flow control schemes in high-speed networks," in *Proc. of IEEE ICDCS 2002 Workshop (ADSN 2002)*, pp. 63–68, July 2002.

[10] C. Takano and M. Aida, "Stability and adaptability of autonomous decentralized flow control in high-speed networks," *IEICE Transactions on Communications*, vol. E86-B, no. 10, 2003.

[11] C. Takano, M. Aida and S. Kuribayashi, "On the stability of autonomous decentralized flow control in high-speed networks with asymmetric configurations," in *Proc. of IEEE ICDCS 2003 Workshop (ADSN 2003)*, pp. 56-62, Providence, RI, May 19, 2003.

[12] M. Aida and C. Takano, "Principle of autonomous decentralized flow control and layered structure of network control with respect to time scales," in *Supplement of the ISADS 2003 Conference Fast Abstracts*, pp. 3-4, Pisa, Italy, April, 2003.

[13] M. Aida and C. Takano, "Principle of Diffusion-Type Autonomous Decentralized Flow Control," *IEICE Society Conference*, B-7-13, 2003.

0-7695-2087-1/04 $20.00 © 2004 **IEEE**