Stability of Autonomous Decentralized Flow Control Schemes in High-Speed Networks

Masaki Aida
NTT Information Sharing Platform Labs
NTT Corporation
Musashino-shi 180-8585, Japan
aida@mieice.org

Chisa Takano
Traffic Research Center
NTT Advanced Technology Corporation
2-4-15, Naka-cho, Musashino-shi 180-0006, Japan
chisa@m.ieice.org

Abstract

This paper focuses on flow control in high-speed networks. Each node in the networks handles its local traffic flow only on the basis of the information it knows, but it is preferable that the decision-making of each node leads to high performance of the whole network. To this end, we investigated the behavior of packet flow when a node is congested, and show an appropriate flow control model through simulation results.

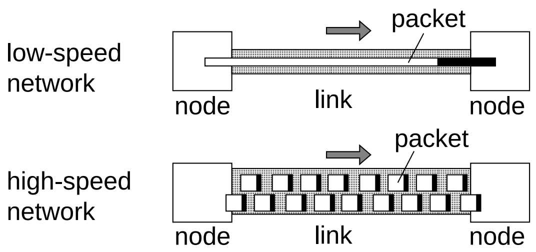

Figure 1. Large delay-bandwidth product.

1. Introduction

In a high-speed network, propagation delay becomes the dominant factor in the transmission delay because the speed of light is a non-scaling factor and is an absolute constraint. Therefore, at a given time, a large amount of data is being propagated on links in the network (Fig. 1). The amount of this data is characterized by the *delay-bandwidth product*, *i.e.*, the propagation distance times transmission rate. Therefore, in high-speed and/or long-distance transmission, there is more data on links than in nodes.

Figure 2 shows an example of how much data can be on a link. Let us consider the situation involving data transmission between two nodes, a distance of 1 km apart with a link speed of 1 Mbps. If the transmission speed is increased to 1 Gbps, the amount of data on the link is equivalent to that on 10^3 km of a 1-Mbps link. And, if the transmission speed is increased to 1 Tbps, the data amount is equivalent to 10^6 km of a 1-Mbps link. This distance is about 2.5 times the distance between the earth and the moon. Consequently, it is impossible to exert time-sensitive control based on collecting global information about the network. So, in a high-speed network, the frameworks of time-sensitive control are inevitably autonomous decentralized systems [1, 2, 3].

This paper focuses on back-pressure type flow control in networks in which nodes handle their local traffic flow themselves based only on the information they know. It is, of course, preferable that the decision making of each node leads to high performance of the whole network. In flow control, we use the total throughput of a network as a global performance measure. We investigate the behaviors of local packet flow and the global performance measure when a node is congested, and show an appropriate flow control model through simulation results.

Related issues on global optimization of flow control by using local information have been studied. For window-based flow control, the optimization problems of some aggregated utility function have been studied in [5, 6]. For the connection-oriented networks, bandwidth assignment problems for each source by optimizing some end-end utility function have been studied in [2, 4]

2. Models

2.1. Performance Measure

Each packet in a network is either in a node or on a link. Since the packets stored in nodes at the present are not being transmitted by the networks, it is natural to define the total throughput of the network as a global performance measure

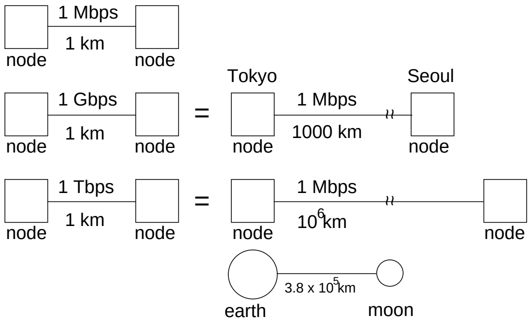

Figure 2. Example of delay-bandwidth product.

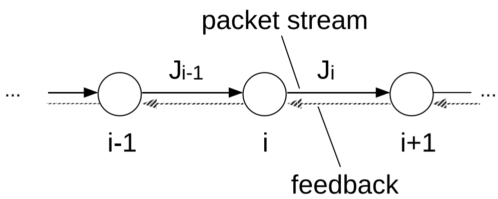

Figure 3. Network model.

as follows. We define the total throughput of a network at time t as being the amount of data being propagated on the network [1, 3, 5]. In other words, it is the number of packets being propagated on all links in the network at time t.

On the other hand, the only packets we can control are ones stored in nodes and not ones being propagated. Thus, higher performance of all the networks involves many uncontrollable packets being propagated on links. Therefore, inappropriate flow control cannot produce the state that has high performance and stability.

2.2. Network and Flow Control Models

Our network model has a simple 1-dimensional configuration (Fig. 3). All nodes have two incoming links and two outgoing ones for a one-way packet stream and feedback, that is, node i ($i = 1, 2, 3, \ldots$) transfers packets to node $i+1$ and node $i+1$ sends feedback information (node information) to node i. For simplicity, we assume that packets have a fixed length in bits.

All nodes are capable of receiving and sending node information from/to adjacent downstream and upstream nodes, respectively. Node information is sent from downstream node $i+1$ to upstream node i. In addition, each node i can send its node information to the upstream node $i-1$.

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

- Each node i autonomously determines the transmission rate J_i based only on information it knows, *i.e.*, the node information obtained from the downstream node $i+1$ and its own node 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 node information according to a predefined rule and sends it to the upstream node $i-1$.
- The rule for creating the node information is the same for all nodes.
- Packets and node 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 achieve centralized control in a high-speed network environment. Hereafter, we investigate the behavior of the total network performance driven by two different flow control schemes, applied for different processes used to determine the transmission rate.

3. Preliminary Description of Flow Control

3.1. Packet Flow

In this paper, we focus on the stability of flow control in the congested state, and we consider packet flow in a heavy-traffic environment. In this situation, we let the packet flow be J_i if the transmission rate specified by node i is J_i. This is because node i has sufficient packets to transfer. Hereafter, we identify the packet flow with the transmission rate specified by the node.

We define the packet flow as

$$J_i := r_i - D\,(n_{i+1} - n_i), \qquad (1)$$

where $n_i(t)$ denotes the number of packets in node i at time t, r_i is the rate sent by downstream node $i+1$ as node information, and D ($D \geq 0$) is a constant. In addition, we call the first and second terms of the right hand side of Eq. (1) drift and diffusion terms, respectively.

If there is no packet loss in the network, the temporal variation of $n_i(t)$ is expressed as

$$n_i(t+1) - n_i(t) = J_{i-1}(t-1) - J_i(t). \tag{2}$$

Here, we let the propagation delay between adjacent nodes be 1, for simplicity.

To estimate the temporal variation roughly, we replace i by x and apply continuous approximation. The packet flow is expressed as

$$J = r(x) - D\,\frac{\partial n}{\partial x}, \tag{3}$$

and the temporal variation of the number of packets at x is expressed as

$$\frac{\partial n}{\partial t} = -\frac{\partial r}{\partial x} + D\,\frac{\partial^2 n}{\partial x^2}, \tag{4}$$

by using the continuous equation

$$\frac{\partial n}{\partial t} + \frac{\partial J}{\partial x} = 0. \tag{5}$$

That is, our method aims at performing a flow control by the analogy of a diffusion phenomenon.

Hereafter, we consider two types of flow control and compare them. One type handles the drift term and the other controls the diffusion term of J_i,

3.2. Drift-Term Driven Flow Control and Stability

In this subsection, we set $D = 0$ in Eqs. (1) and (3), and investigate the characteristics of the flow control scheme whose packet flow is determined only by the drift term.

Let the number of packets in the network be N. To obtain higher network performance, flow control should enable the state in which many packets are being propagated on links. This state corresponds to the state that has fewer packets in nodes.

The simplest strategy for achieving this state is for each node to attempt to decrease the number of packets in it. Therefore, the temporal variation of $n_i(t)$ should be

$$n_i(t+1) - n_i(t) < 0. \tag{6}$$

From Eq. (2) and $D = 0$, this strategy means that node i notifies a smaller rate to the upstream node $i-1$ than the rate notified by the downstream node,

$$r_{i-1} < r_i. \tag{7}$$

However, if all nodes use this strategy, then the total throughput decreases with respect to time as a result. Therefore, the strategy described by Eq. (7) cannot be used continuously.

Conversely, if we use the rate specified to the upstream node as

$$r_{i-1} > r_i, \tag{8}$$

then $n_i(t)$ increases with respect to time (when there are many packets in the upstream node). But the buffer in each node has a finite capacity, so this strategy described by Eq. (8) cannot be used continuously either.

If we set the rate specified to the upstream node as

$$r_{i-1} = r_i, \tag{9}$$

then $n_i(t)$ does not change with respect to time under the heavy traffic condition. This means that the strategy described by Eq. (9) does not diminish the total performance of the network. However, when some node is congested, its restoration requires a long time. Thus, the strategy described by Eq. (9) cannot be used continuously either.

From the above considerations, we choose the following strategy. The rate specified from node i to the upstream node $i-1$ is determined by according to the state of node i. Let the objective of n_i be n_s. If $n_i > n_s$, then r_{i-1} is specified by using Eq. (7); if $n_i < n_s$, then r_{i-1} is specified by using Eq. (8); and if $n_i = n_s$, then r_{i-1} is specified by using Eq. (9).

Since the above flow control do not use the diffusion term, we call it drift-term driven flow control in this paper.

3.3. Diffusion-Term Driven Flow Control and Stability

In this subsection, we set $D > 0$ in Eqs. (1) and (3), and investigate the characteristics of the flow control scheme whose packet flow is determined by both drift and diffusion terms.

From the stability of the drift-term driven flow control, we set

$$r_{i-1} = J_i, \tag{10}$$

as in Eq. (9). Then, Eq. (2) is written as

$$n_i(t+1) - n_i(t) = D\,(n_{i+1} - 2n_i + n_{i-1}). \tag{11}$$

Applying the continuous approximation to this equation, we have

$$\frac{\partial n}{\partial t} = D\,\frac{\partial^2 n}{\partial x^2}. \tag{12}$$

This is the diffusion equation. It implies the feasibility of the following flow control. In the drift-type control $D = 0$, the flow control described by Eq. (9) cannot change $n_i(t)$ and it is difficult to recover from congestion. However, for $D > 0$, since we can control the diffusion term, we expect packets in the congested node to be distributed to the whole network and normal network conditions to be restored after some time.

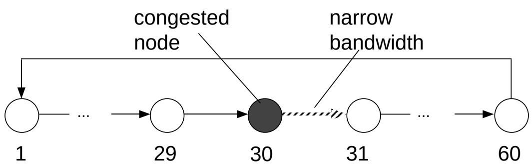

Figure 4. Network model with a bottlenecked link.

In this control scheme, node i's packet transmission rate to the downstream node $i+1$ is determined as

$$J_i = r_i - D\,(n_{i+1} - n_i), \tag{13}$$

and the node information of node i sent to the upstream node $i-1$ is determined as

$$r_{i-1} = J_i. \tag{14}$$

In this framework, node information of i specified to the upstream node $i-1$ is a pair of values $(r_{i-1},\, n_i)$.

Since the above flow control uses the diffusion term, we call it the diffusion-term driven flow control in this paper.

4. Evaluations

In this section, we consider a simple network model with a bottlenecked link and compare the performance of the two different flow controls described in the previous sections.

4.1. Simulation Model

Figure 4 shows our network model, which is a closed network with a 1-dimensional configuration and toroidal boundary. The network has a congested node and a bottlenecked link. All the other nodes and links are in the same condition. This model simulates the situation when congestion occurs at a certain node. We are interested in the behavior of the local congestion, whether

- it causes deterioration of the total network performance through interaction among nodes, or
- it diminishes with time.

Detailed conditions of our network model are listed below.

- Number of nodes: $m = 60$
- Propagation delay between adjacent nodes: 1 (unit time)
- Index of the congested node: $i = 30$

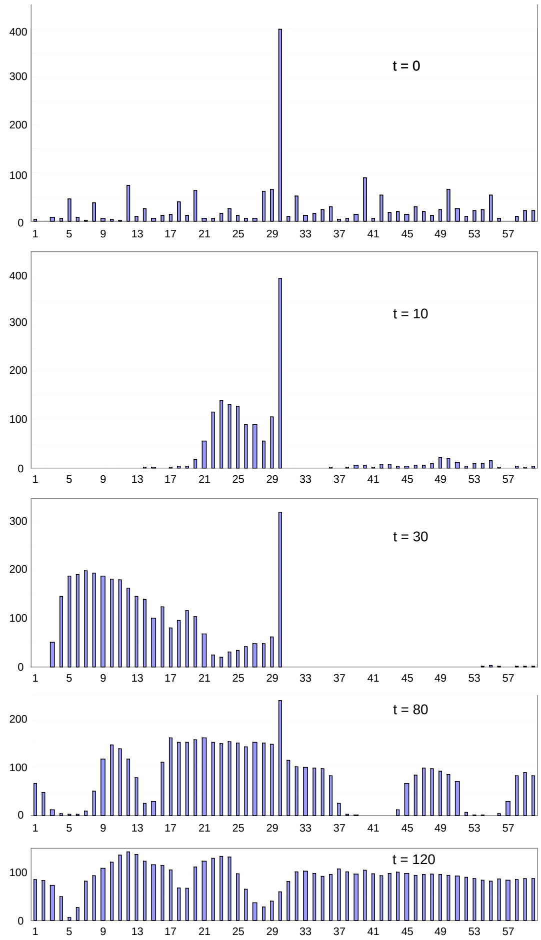

Figure 5. Temporal variation of the number of packets in each node for a drift-term driven flow control scheme.

- Total number of packets in the network: $N = 6000$
- Maximum number of packets on a link (except the bottlenecked link): 100
- Maximum number of packets on the bottlenecked link (between nodes $i = 30$ and 31): 50
  (that is, half the capacity of other links and the same distance)

To investigate the stability under congestion, in addition to the above conditions, we set the initial condition for congested node $i = 30$ as follows.

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

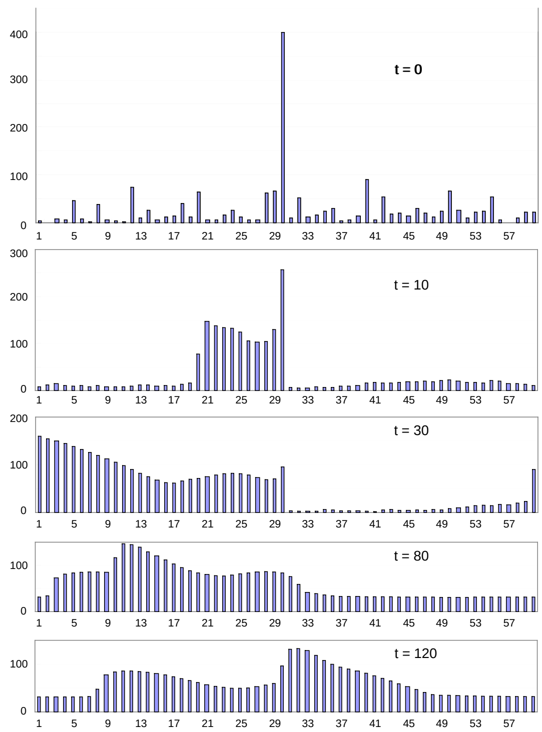

Figure 6. Temporal variation of the number of packets in each node for a diffusion-term driven flow control scheme.

4.2. Drift-Term Driven Flow Control Scheme

As a model for the drift-term driven flow control, we set an objective for the number of packets in a node to be $n_s = 60$, and set the following transmission rate and node information.

$$J_i = \min(r_i, L_i), \tag{15}$$

$$r_{i-1} = \begin{cases} J_i \times 0.9 & (n_i > n_s), \\ J_i \times 1.0 & (n_i = n_s), \\ J_i \times 1.1 & (n_i < n_s), \end{cases} \tag{16}$$

where L_i denotes the link capacity between nodes i and $i+1$.

Figure 5 shows the simulation result for the drift-term driven flow control model. The horizontal axis of each graph denotes node id and the vertical axis denotes the number of packets stored in the node. In addition, t denotes the simulation time and initially $t = 0$.

The number of packets in congested node $i = 30$ decreases with time, and all nodes have around 100 packets at $t = 120$. Each node has the objective $n_s = 60$, so the strategy of each node is a failure as a result.

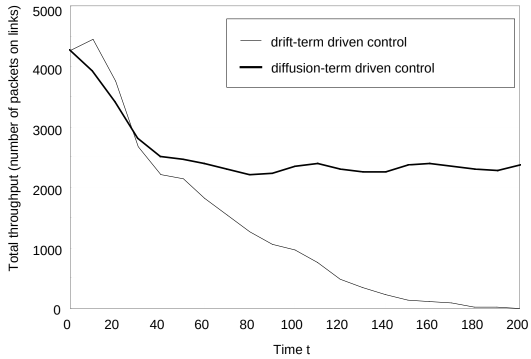

Figure 7. Temporal behavior and stability of the total throughput of the network for two different flow control schemes.

4.3. Diffusion-Term Driven Flow Control Scheme

We set $D = 0.1$ in Eqs. (13) and (14), and use the following flow control model.

$$J_i = \begin{cases} L_i, & (r_i - D\,(n_{i+1} - n_i) > L_i), \\ 0, & (r_i - D\,(n_{i+1} - n_i) < 0), \\ r_i - D\,(n_{i+1} - n_i), & (\text{otherwise}), \end{cases} \tag{17}$$

$$r_{i-1} = J_i. \tag{18}$$

Figure 6 shows the simulation result for the diffusion-term driven flow control model using the same initial condition as described in Fig. 5. The horizontal axis of each graph denotes the node id and the vertical axis denotes the number of packets stored in the node.

The number of packets in congested node $i = 30$ decreases with time, and most of the nodes have fewer packets than in the drift-term driven flow control model at $t = 120$.

4.4. Stability of Flow Control and Network Performance

From the simulation results for the drift- and diffusion-term driven control models, we compare the total throughput of the network. Figure 7 shows the total throughput for both models. The horizontal axis denotes the simulation time and the vertical axis denotes the total throughput (the total number of packet being propagated on links).

For the drift-term driven control models, the total throughput decreases with time. This means that the flow control model inappropriately influences the global performance of the network. If all nodes achieve their objective of $N_s = 60$, the total throughput should be 2400 (total of 6000 packets; and (60 packets/node $\times$ 60 nodes) packets stored in the nodes).

On the other hand, for the diffusion-term driven control models, the total throughput decreases with time but becomes stable around 2400. From the link capacity of the bottlenecked link, the maximum value of the sustainable total throughput (the number of packets being propagated on links stably) is 3000, *i.e.*, 50 packets/link $\times$ 60 links. Thus, the drift-term driven flow control achieves 80total throughput and its value is stable.

If we can choose an appropriate value of the objective n_s for the drift-term driven flow control, the total throughput maybe stable. However, the value should depend on the bandwidth of the bottlenecked link. Since nodes cannot know information about the bandwidth in a high-speed network environment, the drift-term driven control cannot achieve high performance. In the diffusion-term driven control model, although no nodes know the bandwidth of the bottlenecked link, high performance is achiieved.

5. Conclusions

This paper presented a framework for flow control in high-speed networks as an autonomous decentralized system. We showed two typical back-pressure type flow control models based on the framework. The drift-term driven flow control handles the drift term of the packet flow and the diffusion-term driven flow control handles the diffusion term of the packet flow. For both controls, nodes handle their local traffic flow themselves based only on the information they know.

To investigate the behaviors of local packet flow and the global performance measure when a node is congested, we compared two models through simulations. For comparison, we useed the total throughput for the flow control performance measure. The diffusion-term driven flow control adaptively achieves high performance and it is stable in congested situations.

References

[1] 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*, E82-B(12):2095–2106, December 1999.

[2] 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.

[3] 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*, pages 111–117, 1996.

[4] S. H. Low and D. E. Lapsley. Optimization flow control-i: basic algorithm and convergence. *IEEE/ACM Transactions on Networking*, 7(6):861–874, 1999.

[5] D. Mitra. Optimal design of windows for high speed data networks. In *Proc. of INFOCOM '90*, pages 1156–1163.

[6] J. Mo and J. Walrand. Fair end-to-end window based congestion control. *IEEE/ACM Trans. Networking*, 8(5):556–567, October 1999.