Congestion Avoidance in Interconnected Local Area Networks

H. S. Hassanein

Department of Electrical and Computer Engineering
University of Waterloo
Waterloo, Ontario N2L 3G1, Canada

A. A. Elfeitori and P. H. Roe

Department of Systems Design Engineering
University of Waterloo
Waterloo, Ontario N2L 3G1, Canada

ABSTRACT

The interconnection of dissimilar LANs and devices over a high-speed backbone network can creat congestion at a destination LAN (output) gateway when more than one device or LAN attempt to simultaneously communicate with the same device or LAN. We address the issue of congestion avoidance in LANs interconnected by a high speed backbone. A congestion-avoidance scheme that requires input gateways to secure permits to the output gateway before transmitting is proposed. An implementation of the proposed congestion-avoidance scheme in an ATM switch is shown. Performance results indicate that packet dropping rates are greatly reduced.

1. Introduction

The need to support distributed applications such as, file transfers and distributed database, coupled with the availability of low-cost Local Area Network (LAN) equipment has created a profusion of LANs in the last decade. The ability to interconnect LANs (via bridges and routers) made LANs even more attractive.

Bridges and routers, however, may not be sufficient to support the newer applications that require much higher bandwidth. Such applications include, multimedia and medical images. High-speed backbone networks may offer the suitable interconnecting media for LANs supporting traffic-intensive applications (see Figure 1). Such high speed networks include FDDI [1], DQDB [2] and the emerging Local ATM networks [3, 4].

While a high-speed backbone may offer the required bandwidth for traffic-intensive applications, the interconnection of dissimilar LANs and devices over a high-speed backbone has a number of inherent problems. First, a high-speed device such as a file server or a traffic-intensive host can be sending a traffic stream to a low-speed LAN, e.g., Ethernet. Second, more than one device or LAN may attempt to simultaneously communicate with the same device or LAN. Since the speed of the backbone is much higher than that of the individual LANs, packets may be admitted to the backbone network, but could be dropped later at the destination LAN. Without an effective congestion control mechanism, this can lead to performance degradation.

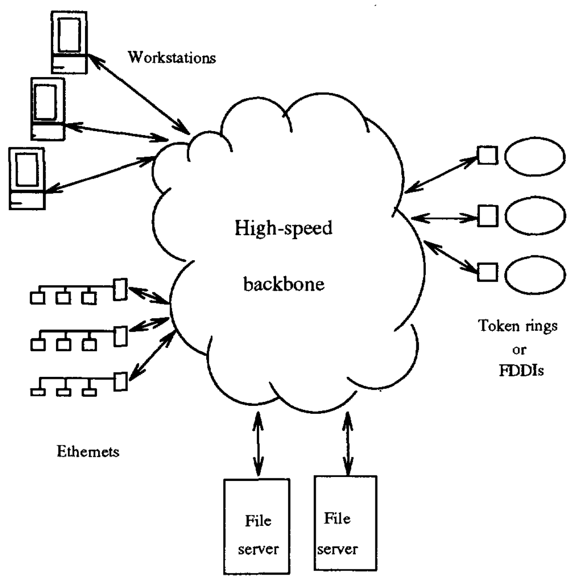


Figure 1
LAN Interconnection via a high-speed backbone

This paper presents a scheme for avoiding congestion at destination LANs. The proposed scheme operates under the provision that the backbone network should assume responsibility of ensuring delivery of all packets it admits. This scheme is described in Section 2.

In Section 3, we illustrate an implementation of the proposed congestion-avoidance scheme in an ATM switch environment. We also show numerical results on the effectiveness of using the proposed scheme to avoid congestion in interconnected LANs. Finally, Section 4 summarizes our findings.

2. Congestion-Avoidance Scheme

Interconnection of LANs via a backbone network requires the use of gateways (or interface modules) to handle the transmit and receive functions on behalf of a LAN. A gateway that handles the transmission of packets from a LAN will be referred to as an *input* gateway and a gateway that handles transmission of traffic from the backbone to a LAN will be referred to as an *output* gateway.

Congestion at the output gateway occurs when more than one LAN attempt to simultaneously communicate with a particular LAN. Such a situation leads to packet dropping at the output gateway due to buffer overflow. This problem, as mentioned earlier, is more apparent when LANs are interconnected via a high speed backbone, since the rate of the backbone is much higher than that of individual LANs. Therefore, packets may be admitted to the backbone network, but could be later dropped at the output gateway, since the destination LAN is of much lower speed.

The proposed scheme operates under the provision that the backbone network should assume responsibility of ensuring delivery of all packets admitted into the input gateway buffer. This is indeed the case for public networks, e.g., SMDS [5]. The problem we are addressing, therefore, is different from classical flow control problems that are performed end-to-end.

The scheme is an admission control one that is an extension of the isarithmic congestion control technique originally proposed by Davies [6]. In the isarithmic technique a limit is placed on the total number of messages in the network (the backbone in our case). This is achieved by circulating a fixed number of *permits* in the network and requiring a source to secure a permit before a message from that source can be admitted to the network. Note that there is no discrimination on basis of source or destination. The isarithmic technique and its variations thereby do not address the problem of simultaneous transmissions to a particular destination.

To alleviate this problem, we define the concept of permit *classes*, where permit class C_i is associated with output gateway G_i. A packet destined for gateway G_i at an input gateway must secure a C_i type permit before it can to transmitted. The number of permits of class C_i and the rate by which these permits are generated can be controlled such that the capacity of the buffer at output gateway G_i is never exceeded. This then should guarantee no packet dropping at the output gateway.

The output gateway is responsible for maintaining the appropriate number of permits. We propose that the total number of permits of class C_i be made equal to the capacity (in packets or cells in ATM networks) of the output buffer of gateway G_i. Therefore, if the buffer occupancy at G_i is n_i, then the number of available C_i permits must exactly be $B_i - n_i$, where B_i is the length of the output buffer of G_i. To achieve this, an output gateway receiving a packet does not recycle the permit into the backbone network until it forwards this packet (or another packet) to the (destination) LAN.

The permit congestion-avoidance scheme, therefore, only requires a controller to monitor permits. This controller is located at the output gateway and should perform the following functions:

(1) maintain the correct number of permits (of its class) in the system.

(2) Circulate the permits equally and fairly among the other gateways.

The latter function is dependent on the interconnecting media (backbone network) and on the fairness criteria required. So, while the congestion-avoidance scheme above is general, its implementation is network- and application-dependent. In the following section we show an ATM switch implementation of our proposed permit congestion-avoidance scheme.

3. ATM Switch Implementation

Asynchronous Transfer Mode (ATM) has been chosen as the switching and multiplexing technique for Broadband Integrated Services Digital Networks (B-ISDN). Bandwidth is allocated on demand using 53-byte data units called cells. ATM supports a wide range of services, including voice, data and video, and provides a flexible means to multiplex these traffic types on the same physical network.

Compared to shared media technology, ATM has a higher aggregate throughput because it allows multiple simultaneous data transfers through a switch. Moreover, switches may be interconnected to further increase the aggregate capacity. ATM switches, therefore, are being proposed as interconnecting media for LANs [3 and 4].

Figure 2 depicts a number of LANs interconnected by an ATM switch. Packets arriving at an (input) interface module are segmented into ATM cells before being transmitted onto the ATM switch. Cells are switched to the appropriate destination within the ATM switch. Packets are then reassembled at the (output) interface modules. The interface modules in ATM networks also assume the role of the gateways described above. In the remainder of this paper, we use the term packet to indicate a LAN packet that has already been segmented into a number of ATM cells.

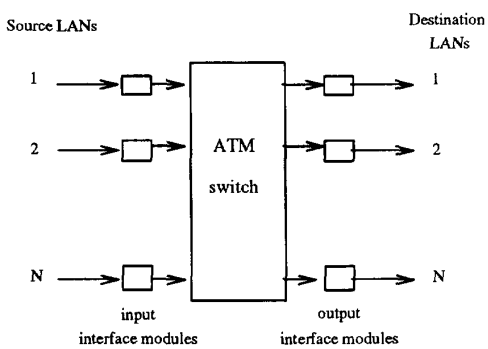


Figure 2
LAN Interconnection using an ATM switch

An ATM switch implementation of the permit congestion-avoidance scheme offers the following advantages:

(1) Input and output buffers are located in the same vicinity. This allows the use of a centralized controller for permit distribution. As well, the information at the output buffer can be made available at the input side with no delays.

(2) Fair and fast permit distribution is realizable. A simple look up table with N entries corresponding to the number of output buffers can be used for this purpose. The values of the table's entries would then reflect the remaining buffer space at the corresponding output buffer. A value of 0 indicates no more cell forwarding to the output buffer.

(3) Since input buffers are all within the ATM switch, buffer management techniques may be used to further improve the performance of the system. This is illustrated later in Section 3.3.

3.1. Switch Architecture

We do not propose a new ATM switch fabric. Rather we provide a generic implementation of the permit congestion-avoidance scheme on a class of input- and output-buffered ATM switches. Detailed descriptions of switching fabrics is beyond the scope of this paper.

Figure 3 depicts an $N \times N$ switch with input and output buffering. The switch is composed of N packet sorters, input buffer space, a routing network and N output buffers. Packet sorters separate the input stream into different streams according to the destination. Input buffers are dynamically allocated. A separate buffer is allocated from the input buffer pool for every source-destination pair. (In this paper, and for simplicity, we assume a same buffer size for each communication session.) The routing network is responsible of routing traffic to the appropriate output buffer.

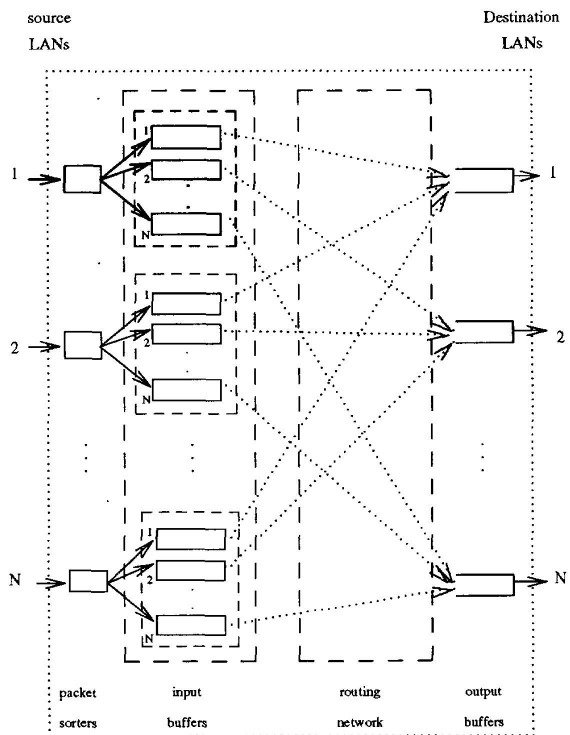


Figure 3
ATM switch architecture

Packets arriving at the input port from LAN i are separated into L_i different streams according to the destination, where $L_i \le N$ is the number of simultaneous transmissions from LAN i. Each of the traffic streams is then allocated an input buffer of size B such that $\sum_{i=1}^{N} B \cdot L_i \le B_{total}$, where B_{total} is the input pool capacity. Packets arriving at a full input buffer pool are dropped, as the switch cannot support more communication sessions.[1] As a communication session (call) is terminated, its associated buffer is made available for future calls.

A Cell destined to output port j is forwarded to the output buffer if the j^{th} entry of the permit table is positive. This entry is then decremented. Otherwise, the cell remains at the input buffer. The j^{th} entry of the permit table is incremented for every cell forwarded from output buffer j onto LAN j. Note that there is no cell loss at the output buffer, but rather at the individual input buffers.

3.2. Performance Model

We analyze the performance of the permit congestion-avoidance scheme using a queuing model in a generic type of environment similar to that in [7]. Without loss of generality, consider the case where K LANs are communicating with the same output LAN (see Figure 4). Packets arrive at the input buffers from the source LANs at a rate of λ_{LAN}. These packets are served according the backbone rate (μ_b), then they are routed to the appropriate output buffer and are served according to the destination LAN rate (μ_{LAN}).

We assume a symmetric system where K LANs are communicating with the same output LAN. All arrivals and service rates are poisson. We also assume that the backbone service rate (ATM switch speed) is much higher than that of individual LANs. Hence, we do not consider μ_b. Let each of the input buffers be of length B and the output buffer be of length B_o.

Note that both the input and output buffers reside on the switch. Therefore, cells arrive at the output queue as soon as they obtain a permit. The system can then be modeled as an M/M/1/K queue with queue-length dependent arrival rate, where a packet (cell) is randomly selected from one of the K input buffers.

[1] An optimal buffer size assignment, as well as buffer sharing schemes can be used to achieve a higher acceptance rate. This is the subject of further study.

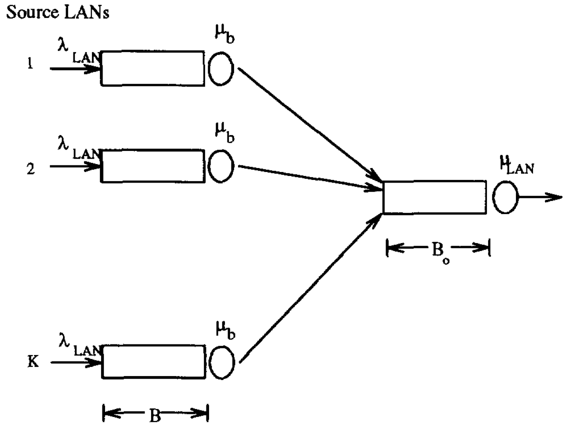


Figure 4
Queuing Model

Consider the problem of inserting n distinct balls into K urns each of length B [8]. The number of distinct occupancies, C(n, K), is given by

$$C(n,K) = \sum_{j=0}^{K} (-1)^j \binom{K}{j} \times \binom{n+K-j(B+1)-1}{K-1}$$

Assuming a uniform distribution of occupancy for $1 \le n \le KB$, the conditional joint probability of the individual occupancies given that the total number in the system is n is given by $\frac{1}{C(n,K)}$.

Now, the probability that input queue K is full is given by

$$P^{full}(n) = \begin{cases} 0 & 0 \le n < B \\ \frac{C(n-B,K-1)}{C(n,K)} & B \le n \le KB \end{cases}$$

The queue-length dependent arrival rate, taking into account the out buffer capacity (B_o), is given by

$$\lambda(n) = \begin{cases} K\lambda & 0 \le n < B+B_o \\ P^{full}(n-B_o)K\lambda & B+B_o \le n \le KB+B_o \quad (1) \\ 0 & otherwise \end{cases}$$

Let $P(n)$ be the stationary queue length distribution, the probability of cell loss is then given by

$$P_{loss} = \sum_{n=B+B_o}^{KB+B_o} P^{full}(n-B_o)P(n) \qquad (2)$$

3.3. Numerical Results

In this section, we present performance results of the permit congestion-avoidance scheme using the analysis above. We compare these results to those from an uncontrolled system, as well as to a system with shared input buffering, as explained below.

(a) *Uncontrolled System:*
In this case, cells move from the input buffers to the output buffer without any control. Since the switch speed is much higher than that of the LANs, it is fair to assume that packets are not queued at the input buffers, but are rather immediately forwarded onto the switch to the output buffer. The uncontrolled system with K simultaneous transmissions to the output buffer can, therefore, be modeled as an M/M/1 queue with finite buffer (B_o) and an arrival rate of $K\lambda$.

(b) *Controlled System:*
This is the system described in 3.1 above. The queue-length dependent arrival rate, and the loss probability were given in (1) and (2) above respectively.

(c) *Controlled System (shared destination buffers):*
Recall that the input buffer space is dynamically allocated. In this case, rather than allocating a separate buffer for each LAN source-destination pair, cells destined to the same output port (LAN) from K different source LANs share a same (KB) buffer. The probability of loss in this case can be obtained by setting the input buffer capacity to KB and setting K to 1 is equation (2) above.
Note that buffer sharing is only achievable when input buffers reside in the same vicinity, i.e., within an ATM switch. In the more general case where input gateways are remotely located such a configuration is unrealizable.

We set the output buffer size (B_o) to 10 and vary the input buffer size, B, the number of simultaneous transmissions, K, and the offered load. To simplify matters we use an arrival rate normalized to the LAN service rate, μ_{LAN} (or equivalently set μ_{LAN} to 1). The offered load is given by $K\lambda$. We, obviously, do not consider cases with an offered load greater than 1.

Figure 5-7 plot the total loss probability for the three systems above as a function of offered load, input buffer capacity and number of simultaneous transmissions, respectively. We first note that these figures show that, regardless of system parameters, the uncontrolled system has the worst performance and the controlled system with shared destination buffers is best. The performance of the uncontrolled system is obvious. On the other hand, the system with shared destination buffers outperforms the one with separate queues, since in the former cell loss can occur only if the total buffer occupancy exceeds KB. While in the latter, cell loss occurs when one or more input buffers (each of capacity K) are full.

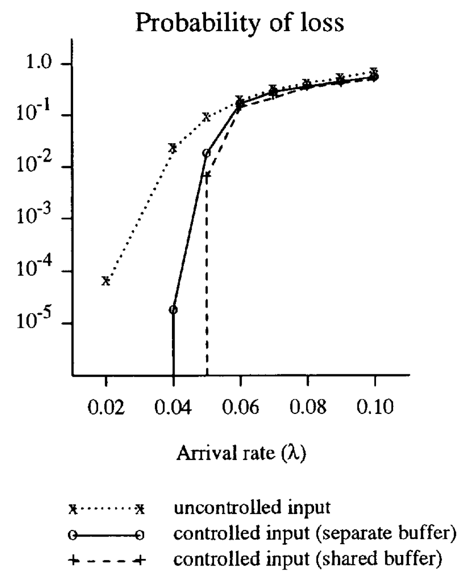
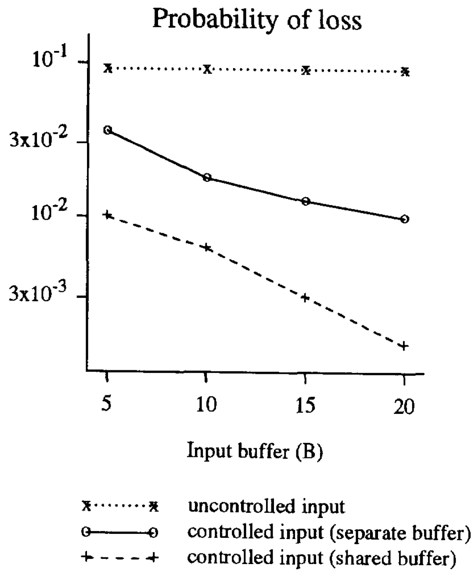


Figure 5
Probability of loss vs. the offered load

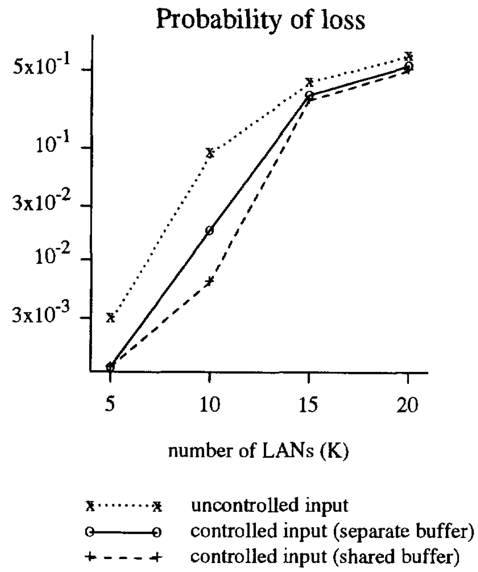
We make the following observations from Figures 5-7:

- As the offered load increases the performance of all three system becomes similar (see Figure 5). This is because controlled systems now cause an equivalent packet loss to the uncontrolled system but on the input side, as the system becomes unstable. In fact, no congestion-avoidance scheme can improve the performance of a system that is continuously under overload conditions.



(B_o = 10; K = 10; offered load = 0.5)

Figure 6
Probability of loss vs. the input buffer size



(B_o = 10; B = 10; offered load = 0.5)

Figure 7
Probability of loss vs. the number of LANs

- The advantage of using the permit congestion-avoidance scheme is clearly demonstrated in Figure 6. For instance, at B=20, the gain in loss performance is at least an order of magnitude.
- The advantage of buffer sharing becomes more evident as the number of simultaneous connections increase. For instance compare the loss probabilities for the separate and shared buffers systems at K=5 and K=10 in Figure 7. While this is generally true, further increasing K at a same LAN arrival rate causes a higher offered load. Eventually, both systems become equivalent.

4. Conclusions

The interconnection of dissimilar LANs Over a high-speed backbone has a number of inherent problems. First, a high-speed device such as a file server or a traffic-intensive host can be sending a traffic stream to a low-speed LAN, e.g., Ethernet. Second, more than one LAN may attempt to simultaneously communicate with the same device or LAN. Since the speed of the backbone is much higher than that of the individual LANs, packets may be admitted to the backbone network, but could be dropped later at the destination LAN. Without an effective congestion control mechanism, this can lead to performance degradation.

We have introduced a congestion-avoidance scheme to alleviate this problem. This scheme requires input gateways to secure permits to the output gateway before transmitting. The number of permits to any of the output gateways is made to be equal to the buffer capacity of the gateway. Therefore, there is no packet dropping at the output gateway, but rather at the individual input gateways. The proposed scheme does not require exchange of control packets, nor the need for an explicit switch from the uncontrolled to a controlled mode.

An implementation of the proposed congestion-avoidance scheme in an ATM switch has been demonstrated. An ATM switch implementation offers a number of advantages. First, as input and output buffers reside in the same vicinity, information from the output side can be made available at the input side without delay. Second, the distribution of permits is much simpler. A look-up table can be used to indicate the available buffer capacity on the output side. Finally, since input buffers are all within the ATM switch, buffer management techniques may be used to further improve the performance of the system.

Performance results indicate that the packet dropping rate is greatly reduced by using the permit congestion-avoidance scheme.

References:

[1] "Fiber Distributed Data Interface (FDDI) - Token Ring Media Access Control (MAC)", ANSI standard X3.193, 1987.

[2] "Distributed Queue Dual Bus (DQDB) Subnetwork of a Metropolitan Area Network (MAN)," IEEE Standard 802.6, Dec. 1990.

[3] J-Y Le Boudec, E. Port and H. L. Troung, "Flight of the FALCON," *IEEE Comm. Magazine*, Feb. 1993, pp. 50-56.

[4] V.J. Freisen, H.S. Hassanein, J.W. Wong and J.W. Mark, "A Performance Study of LAN Interconnection Using An ATM Switch," *To Appear, Int'l Conf. on Comp. Commun. Networks,* San Diego, Mar., 1993, 8pp.

[5] G. H. Clapp, "LAN Interconnection Across SMDS," *IEEE Network Magazine,* Sept. 1991, pp. 25-32.

[6] D. Davies, "The Control of Congestion in Packet Switching Networks," *IEEE Trans. on Commun., June, 1972, pp. 546-550.*

[7] D. Towsley, S. Fdida and H. Santaso, "Congestion Avoidance in High-Speed Interconnection Systems," *Computer Networks & ISDN Sys.,* Vol. 24, 1992, pp. 185-195.

[8] A. Ganz and I. Chlamtac, "A linear Solution to Queueing Analysis of Synchronous finite buffer networks," *IEEE Trans. on Commun., Vol. 38, No. 4, Apr. 1990, pp. 440-446.*