論 文

Window Flow 제어기능을 가진 음성/데이타 패킷통신망의 성능해석

正會員 孫 壽 炫* 正會員 殷 鍾 官**

Performance Analysis of an Integrated Voice/Data Packet Communication Network with Window Flow Control

Soo Hyeon SON* Chong Kwan UN** *Regular Members*

要 約 본 논문에서는 window flow 제어기능을 가진 음성과 데이타가 집적된 패킷 통신망을 폐쇄 multichain queueing 시스템으로 modeling하고 평균치해석 방법으로 성능을 분석하였다. 서로 다른 priority 등급을 가진 여러가지 메세지를 전송하기 위한 패킷망 성능해석을 위해서 본 논문에서는 평균치해석과 effective capacity 개념에 의한 성능해석 방법을 사용하였다. 구체적으로 각 노드에서의 평균 buffer 점유율, virtual channel의 link throughput의 이용도, 각 메세지의 평균지연시간등 망의 통계적 특성들을 이론으로 분석하고 simulation으로 검증하였다. 제안된 해석방법을 사용할 경우 link의 데이타의 상태를 10% 이내의 정확도로 예측할 수 있고 음성 메시지와 외부 데이타의 상태는 5% 이내의 정확도로 예측할 수 있다.

ABSTRACT In this paper, an integrated voice/data packet network with window flow control is modeled by a closed multichain queueing system, and its performance is analyzed by the mean value analysis method. Particularly, for the analysis of a packet network having various kinds of messages with different priority classes, we introduce an approach based on the mean value analysis and the concept of effective capacity. By the mathematical analysis and computer simulation, we obtain the following network statistics in the steady state: Mean buffer occupancy at each node, utilization of link throughput of a virtual channel, and the mean delay time of each message. Our iterative analysis method can predict the link data status in most cases within about 10 percent of accuracy, and the statistics of voice messages and external data within 5 percent as compared to simulation results.

I. INTRODUCTION

The store and forward concept of packet-switched networks was originally introduced in order to increase the efficiency by sharing network resources. This efficiency decreases rapidly when the network becomes congested. Routing and flow control must be done to reduce such congestion [1],[2]

* 金星電氣(株) 研究員
GOLD STAR Electronics Co., LTD.
** 韓國科學技術院 通信工學研究室
Communications Research Laboratory, KAIST
論文番號 : 86-24(接受 1986. 5. 10)

Routing increases the efficiency by avoiding queue deadlock. It increases the actual service rate by distributing the incoming data streams from many users. Internal network congestions may be alleviated by changing the route of some traffic from heavily loaded paths to underutilized paths.

On the other hand, flow control gives a mechanism by which flow of data is controlled so that data cannot be transmitted at a rate faster than the receiver can handle it. This end-to-end flow control concept can be applied to prevent from congesting externally. Since data traffic is bursty in general, flow control must have 'the function of speed matching between the network and its attached users. One effective method for this purpose is the sliding window scheme. This entry-to-exit flow control is done for a source-destination pair. The flow of data between each pair of source and destination is individually controlled by a window. This scheme allows up to W sequential messages to be outstanding in the source-destination pair path when the window size is W. When an acknowledgement (ACK) signal is received from a destination node, a next sequential message will be transmitted from the source node.

It is noted that the restriction on the total number of outstanding messages causes network delay which may have undesirable effects for packetized voice. For voice messages, some transmission errors can be permitted, but long delay cannot be allowed. The reason is that, if voice packets are delayed too long, the received speech assembled at a destination node can be different from the original speech and also perceptually unnatural. Therefore, there cannot be acknowlledgment messages for voice packets from the destnation node. And voice packets must be served prior to data packets to reduce the delay time.

In modeling a computer network, each node is modeled by a single queue. And the network is regarded as tandem links in which all queues are linked each other [3]. As a result of the restriction on the total number of outstanding messages for data, window flow control transforms an open queueing system into a closed one [4]. Therefore, queueing analysis of data packets controlled by a window scheme can be done using a closed chain model. Analysis of queueing networks with many closed chains have been done by Reiser and Lavemberg [5]. However, analysis of a queueing network that has closed chains mixed with open queueing systems has not been done yet.

In this work, we analyze the performance of a voice/data integrated queueing network by using the modified method of mean value analysis [4], [5]. In our analysis a packet switching node is modeled as a single server queueing system. The queueing discipline for the same kind of message is based on the "first-in, first-out" (FIFO) rule. There are many kinds of messages that can be transmitted through a packet communication network. As a result, these messages can have different levels of priority for service. Short control messages, acknowledgement messages, and voice packets are often given priority over normal data messages. Such messages with high priority are commonly served by the nonpreemptive priority queueing discipline [6].

In this study we analyze the performance of the network for the two cases; the case of having acknowledgments with the same priority as data messages, and the case of having acknowledgments with priority higher than data or voice messages.

For analysis of a network having voice and external data as well as link data, we use the M/G/1 queueing theory with different priority classes. With these methods, we obtain the mean buffer occupancy at each node, utilization of links, throughput of a virtual channel, and the mean delay time of each message. In addition, we investigate the network performances for different window sizes, thereby obtaining the optimal window size that maximizes the network efficiency for different traffic intensities.

Following this introduction, in section II we analyze the performance of a voice/data packet network. In this section we first consider a queueing model for the virtual path in a packet network with window flow control. We analyze the network performance for two cases of transmitting data only and transmitting both voice and data. In section III we present and discuss simulation results, and compare them to analytical results. Finally, in section IV we make conclusions.

II. PERFORMANCE ANALYSIS OF A VOICE/ DATA PACKET COMMUNICATION NETWORK

A. Queueing Model of a Virtual Path with Window Flow Control

When a user transmits messages to some destination nodes along a fixed path, the node in the path (commonly called the logical line) is regarded as a queue, and the communication line as a server. The message generated by the user of this logical link (called the link customer) enters the first queue. If its server is busy, it waits to be served. When its service turn arrives, it will be served and then move to the second queue. This procedure continues until it is finally served by the last server. In addition to link customers, messages from other users also enter the queue, and are served. To distinguish those customers from the link customers, they will be called the external customers. The sources of these external customers are the preceding link queues, other sources within the network, or those outside the network. Note that the number of data being transmitted is the same as the window size in the path. Therefore, the path can be regarded as a closed chain in which the number of moving packets is constant. Assuming that there is some acknowledgment delay, the network under consideration can be modeled as a single closed chain as shown in Fig. 1.

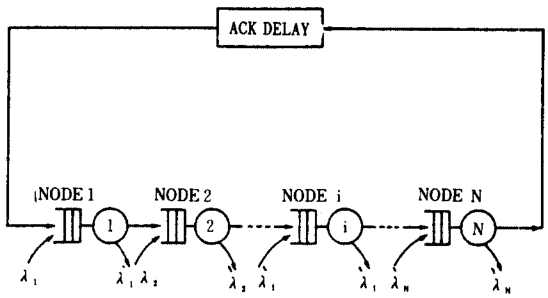


Fig. 1 A single closed-chain queueing network model with external customers.

To make our analysis mathematically tractable, we make the following assumptions:

(i) The arrival process of external customers at each node is independent of the flow of link customers and also independent from node to node. Assuming that the departure time of customers is distributed exponentially, arrivals at each node can be considered independently. Therefore, the arrivals of the external customers at each node are assumed to occur according to an independent Poisson process. However, because the link customers at the first queue arrive only when an acknowledgement message is received from the last queue, their arrivals cannot be a Poisson process.

(ii) Assuming that the buffer size at each node is large, the blocking probability of messages at each node is ignored.

(iii) The routing policy employed in this model is fixed routing. The virtual path between the source and the destination cannot be changed since it is predetermined. Therefore, the steady state of the network can be assumed.

(iv) The number of link messages is less than its window size. When the number of the link messages in the virtual path is equal to its window size, the arrival process of the link messages is halted.

(v) The link messages in the virtual path are individually acknowledged. We assume that each message having arrived at a destination node generates an acknowledgment message which is transmitted over the reverse route. For simplicity, we assume that each acknowledgment is not piggybacked with data messages, but is transmitted by itself. Also, we assume that it has equal or higher priority than data messages.

(vi) Voice messages have higher priority than data. If a data message is not in the process of being served, voice messages are always served prior to data messages.

B. Performance Analysis

Before analysis, we define the following notations:

N	Number of nodes in the virtual path,
W	Window size,
C_i	Link capacity at node i,
E(W)	Mean waiting time,
ℓ	Mean message length,
τ_i	Mean service time at node i, i$(=\tilde{\ell}/C_i)$,
n_i	Mean queue size at node i,
t_i	Mean delay time at node i,
λ_i	Throughput of each message at node i,
ρ_i	Traffic intensity of message at node i,
ℓ_a	Length of acknowledgment message.

For the above notations we also use subscripts "v", "d", "ℓ", "e", "a" representing voice, data, link and external messages, and acknowledgment, respectively. In what follows, we will consider the following two categories of acknowledgement strategies:

(i) Stand-alone acknowledgment messages with the same priority as data messages,

(ii) Stand-alone acknowledgment messages higher priority than data messages.

(1) Analysis in the Case of Acknowledgments without Priority

After having been successfully absorbed at the destination node, each message generates an acknowledgment that is transmitted as an independent packet over the reverse route. Acknowledgment messages queue in buffers in order of arrival since they have the same priority as data messages in this case.

(a) The Case of Transmitting Data Only

In this case we divide the input messages of a queue into three types. There are link data transmitted along the virtual path and acknowledgments as the responses to them. In addition, there may be external data arrived from various other nodes within or outside the network.

Since acknowledgment messages are normally much shorter than data messages (i.e., $\ell_a \ll \tilde{\ell}_d$), we can easily obtain the effective capacity C_i of link data as

$$\overline{C}_i = C_i\,(1-\rho_{ei}), \tag{1}$$

and the mean service time as

$$\tau_i = \tilde{\ell}_d \,/\, \overline{C}_i \,. \tag{2}$$

Also, since there are some acknowledgment delay in a network, data packets stay on the virtual path for a time duration that is equal to the sum of data transmission delay and acknowledgment delay. Therefore, the actual average number of packets in the virtual path, $\overline{W}$, is given by

$$\overline{W} = W \cdot \frac{\text{Data transmission time}}{\text{Data transmission time} + \text{ACK delay}} \,. \tag{3}$$

Once we get the mean service time and the total average queue length from (2) and (3), we can obtain various parameter values of the link data by using the mean value analysis method. Also, we can easily obtain the average delay time of the external data by using the Pollaczek-Khinchine formula [6] as

$$t_{ei} = [\text{Waiting time at queue i}] + [\text{Service time}]$$
$$= \frac{(\rho_{li} + \rho_{ei}) \cdot E(\tau_{di})}{(1 - \rho_{li} - \rho_{ei})} + E(\tau_{di}) \qquad (4)$$

where $E(\tau_{di})(=\tilde{l}_d/C_i)$ is the mean service time of data packets at the queue i. In addition, the average queue length of the external data can be obtained from Little's formula as

$$n_{ei} = \lambda_{ei}\, t_{ei} \cdot \qquad (5)$$

Since there are packets at a queue whose length is the same as the sum of n_{ei} and n_{li}, the delay time due to acknowledgment delay is

$$t_{ai} = (n_{li} + n_{ei})\, E(\tau_{di}) + l_a/C_i \cdot \qquad (6)$$

(b) **The Case of Transmitting Voice and Data**

Since voice packets have higher priority than data packets for service, we shall simply use the M/G/1 queueing theory with nonpreemptive priority in this case. We assume that voice and data packets have their own buffers.

Let us consider the capacities consumed by each user at an M/G/1 queue with different priority. When there is no priority scheme, and messages are served only by the FIFO rule, the average waiting time for any message is given by the Pollaczek-Khinchine formula as

$$E(w) = \frac{\frac{1}{2}\sum_{k=1}^{K} \lambda_k E(\tau_k^2)}{1 - \sum_{k=1}^{K} \rho_k} \cdot \qquad (7)$$

And, in an M/G/1 queue with K multiple classes of users, each with different priority, the average waiting time for the eessage of priority, $E(w_P)$, is given by

$$E(w_P) = \frac{\frac{1}{2}\sum_{k=1}^{K} \lambda_k E(\tau_k^2)}{(1 - \sigma_{P-1})(1 - \sigma_P)} \qquad (8)$$

with $\sigma_P \triangleq \sum_{k=1}^{P} \rho_k$,

where $E(\tau_k^2)$ is the second moment of the service time for messages of priority k, λ_k is its arrival rate, and ρ_k is its traffic intensity [7]. From (1) and (2) it can be stated that messages with the lowest priority r see the "actual capacity" as

$$\overline{C}_i = (1 - \sigma_{r-1})\, C_i \qquad (9)$$
$$= (1 - \sum_{k=1}^{r-1} \rho_k)\, C_i \cdot$$

Consequently, from (9) the waiting time of voice at a node is given by

$$E(w_v) = \frac{[(\rho_{li} + \rho_{ei})\, E(\tau_{di}) + \rho_{vi} E(\tau_{vi})]}{(1 - \rho_{vi})} \qquad (10)$$

where $E(\tau_{vi})$ $(\triangleq \tilde{l}_v/C_i)$ is the mean service time of the voice message at the node i, and $E(\tau_{di})$ is similarly defined for data. Considering the priority effect of voice messages, the capacity used by voice messages is

$$C_v = \rho_{vi}\, (1 - \rho_{vi})\, C_i \cdot \qquad (11)$$

Therefore, the effective capacity available for link data is given as

$$\overline{C}_i = (1 - \rho_{vi})\, C_i\, (1 - \rho_{ei}) - C_v \cdot \qquad (12)$$

they have priority higher than data messages. The results are shown in Fig. 8. As seen in the figure, since link data can receive a response from the destination node quickly when acknowledgments have priority, the throughput for the case of having priority is larger than that without priority.

Also, Fig. 9 shows the performance of the network with data traffic only for different window sizes. Ir is noted that the network

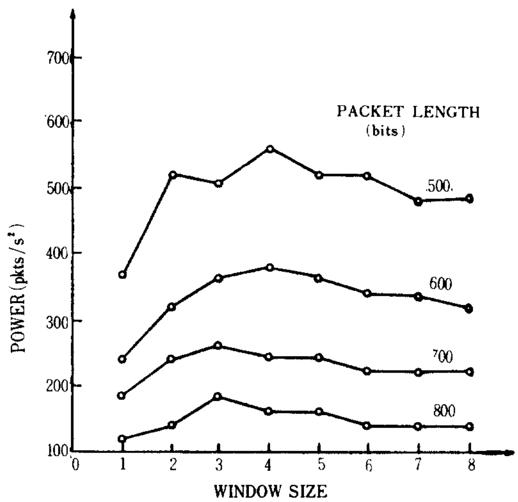


Fig. 9 Power vs. window size for different packet lengths.

parameters that are directly influenced by the flow control mechanism are the throughput and the average delay. But, these two parameters are contradictory in that in general the increase of throughput results in the undesirable effect of having a larger delay. Therefore, the ratio of these two parameters (termed the power) gives some measure of the network performance. That is, the power P is defined by

$$P \triangleq R/T \tag{30}$$

where R is the throughput and T is the average network delay. As seen in Fig. 9, the performance of a low loaded network is better than that of highly loaded network. When the traffic of a network is scarce, the optimal window size is close to the number of nodes through which messages traverse. But, to maximize the efficiency of the network, the window size must be decreased as the traffic increases.

So far, we have discussed performances of the packet network when the traffic is data only. We now discuss the performance of the network when the traffic is both voice and data. In an integrated voice/data network, voice messages have priority higher than data messages so that voice delay can be minimized. Here we assume that each voice packet has a constant length of 1000 bits. First, let us consider the case that no priority is given to acknowledgment messages. In Table 3, both analysis and simulation results for throughput, queue length and delay time of link data are shown. In this case, the delay time of acknowledgment messages increases. Therefore, the results of analysis is less accurate. The throughputs and the average delay time in this

Table 3. Throughput, queue length and delay time of link data when the traffic is data and voice. (No priority is given to acknowledgment messages. Traffic rate=10 pkts/sec. The value in parethesis is simulation result.)

Data Length (bits)	Throughput (pkts/s)	Queue Length (pkts/node)	Delay time (ms)
500	29.99 (32.86)	0.80 (0.83)	107.2 (99.3)
600	24.65 (28.04)	0.82 (0.84)	133.0 (118.1)
700	20.75 (24.4)	0.83 (0.86)	160.1 (139.9)
800	17.79 (20.57)	0.84 (0.87)	188.7 (165.1)
900	15.47 (18.68)	0.85 (0.87)	218.9 (180.1)
1000	13.60 (16.98)	0.85 (0.88)	250.7 (202.1)

network are shown in Figs. 10 and 11, respectively. We note that compared with the previous results, the results of analysis is more accurate when the network traffic is small. Also, throughputs, queue length and delay time of voice messages at a node are shown in Table 4, and the same for external data are give in Table 5. In addition, the network performance represented by the power is shown for different window sizes in Fig. 12. It is seen that the window size in the network

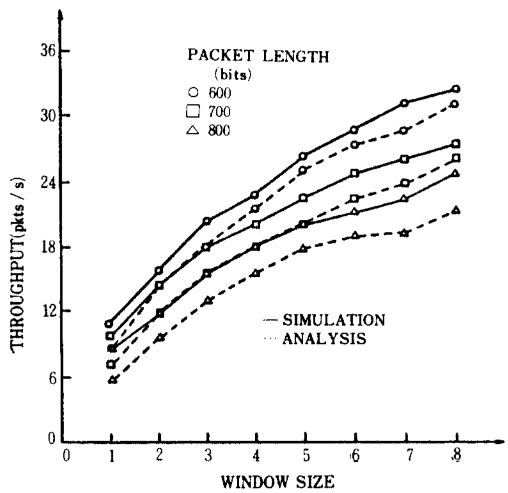


Fig. 10 Throughput vs. window size in the case of acknowledgment without priority in a voice/data integrated network(Traffic intensity=0.68).

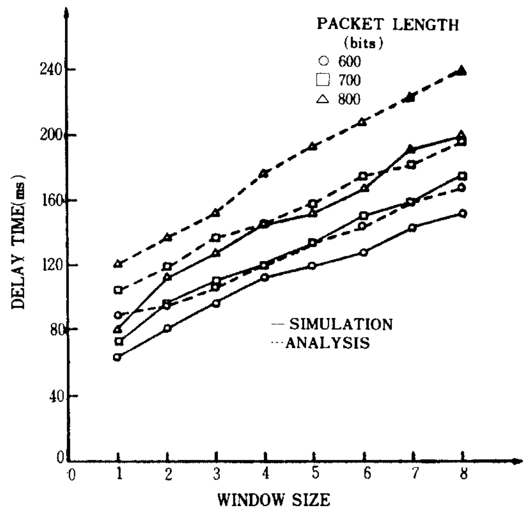


Fig. 11 Delay time vs. window size in the case of acknowledgment without priority in a voice/data integrated network (Traffic intensity = 0.68).

Table 4. Throughput, queue length and delay time of voice messages in an integrated voice/data network. (No priority is given to acknowledgment. Traffic rate=10 pkts/sec. The value in parethesis is simulation result.)

Data Length (bits)	Throughput (pkts/s)	Queue Length (pkts/node)	Delay Time (ms)
500	10.00 (10.21)	0.26 (0.27)	25.7 (26.9)
600	10.00 (9.64)	0.27 (0.28)	26.7 (28.6)
700	10.00 (9.89)	0.28 (0.28)	27.7 (28.7)
800	10.00 (10.36)	0.29 (0.33)	28.7 (31.5)
900	10.00 (9.77)	0.30 (0.31)	29.8 (31.5)
1000	10.00 (9.89)	0.31 (0.33)	31.0 (33.0)

Table 5. Throughput, queue length and delay time of external data in a voice/data integrated network. (No priority is given to acknowledgment. Traffic rate=10 pkts/s. The value in parethesis is simulation result.)

Data Length (bits)	Throughput (pkts/s)	Queue Length (pkts/node)	Delay Time (ms)
500	10.00 (10.04)	0.26 (0.27)	25.7 (27.1)
600	10.00 (10.12)	0.30 (0.34)	30.3 (33.6)
700	10.00 (9.62)	0.35 (0.36)	35.0 (37.4)
800	10.00 (9.99)	0.40 (0.42)	39.9 (42.1)
900	10.00 (9.90)	0.45 (0.50)	45.2 (50.2)
1000	10.00 (10.03)	0.51 (0.57)	50.7 (57.1)

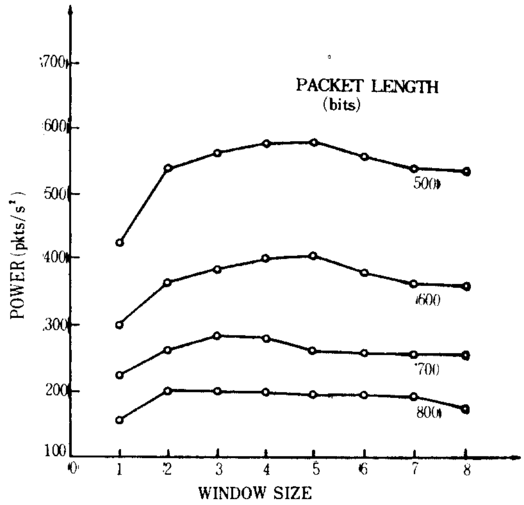


Fig. 12 Power vs. window size for different data packet lengths in a voice/data integrated network (Traffic intensity = 0.68).

of voice and data is less than that of the network where the traffic is data only.

Next, we investigate the performance of an integrated voice/data network where acknowledgments have the highest priority. In Table 6, both analysis and simulation results for throughput, queue length and delay time of link data in the network are shown. In this case, we find that throughputs and the average delay time are larger as compared with the network in which acknowledgments have no priority. And, in Figs. 13 and 14 the analytical solutions for throughput

Table 6. Throughput, queue length and delay time of link data in a voice/data integrated network. (Acknowledgments have the highest priority. Traffic rate=20 pkts/s. The value in parethesis is simulation result.)

Date Length (bits)	Throughput (pkts/s)	Queue Length (pkts/node)	Delay Time (ms)
500	15.41 (16.94)	0.86 (0.87)	222.3 (205.1)
600	12.01 (14.10)	0.88 (0.87)	292.4 (249.0)
700	9.53 (12.18)	0.89 (0.91)	375.4 (290.5)
800	7.64 (9.46)	0.91 (0.92)	475.3 (381.2)
900	6.15 (8.04)	0.92 (0.94)	597.7 (452.5)
1000	4.95 (6.73)	0.93 (0.86)	751.2 (542.7)

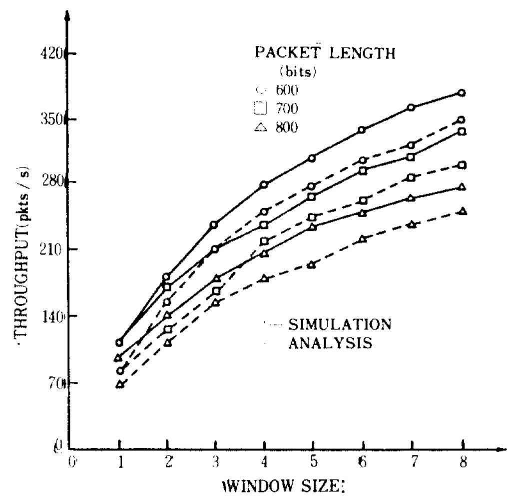


Fig. 13 Throughput vs. window size in the case of acknowledgment with priority in a voice/data integrated network (Traffic intensity = 0.68).

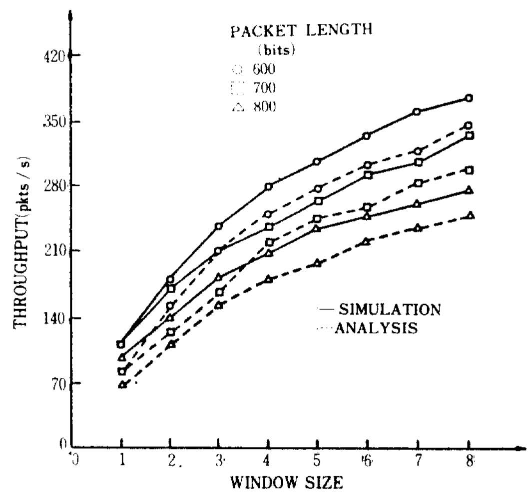


Fig. 14 Throughput vs. window size in the case of acknowledgment with priority in a voice/data integrated network (Traffic intensity = 0.68).

and delay time, respectively, are compared to the results obtained from simulation. It is seen that there is less deviation between analytical and simulation results as compared to the results of the network in which the traffic is data only. But, in contrast with link data, the analytical solutions of voice messages and external data are by far closest to the results of simulation.

IV. CONCLUSIONS

We have investigated the performance of a packet communication network with the end-to-end window flow control on a virtual channel when the network has both voice and data traffics. After formulating a queueing model, a computationally efficient solution has been obtained. By the mathematical analysis and the computer simulation, we have obtained the mean buffer occupancy for each node, utilizations of links, the throughput of a virtual channel, and the mean delay time of each message. Our iterative analysis method can predict the link data status in most cases within about 10 percent of accuracy, and the statistics of voice messages and external data within 5 percent as compared to simulation results. It has been observed that results of the analysis are closer to those obtained from simulation as the traffic intensity decreases because the delay time of acknowledgment message has a significant effect on the results of the analysis. As the delay time of acknowledgment messages increases, their numbers in a virtual path increase. Therefore, the traffic of acknowledgment messages in this case cannot be neglected. From the results of both analysis and simulation, it has been found that, if the traffic intensity is less than about 0.7, the results of the analysis approache to that obtained from simulation within about 10 percent of accuracy. If the traffic intensity of a virtual path is more than 0.7, then we should use the multi-users priority queueing theory with three classes in which acknowledgment messages are included.

In addition, we have studied the network performances for the two cases; the case of having acknowledgements with the same priority as data messages and the case of having acknowledgments with priority higher than other messages. In the case of acknowledgments with priority, it has been found that, although the delay time of other messages increases slightly, their throughputs increase. The priority effect of acknowledgment messages results in the increase of the network performance.

Also, it has been found that the optimal window size to maximize the network performance should be reduced as the delay time of acknowledgment messages increases. In other words, when the network traffic increases, the number of messages accepted by the network should be reduced.

參 考 文 獻

(1) L. Kleinlock, Queueing Systems, vol. 2, New York, Wiley 1976.

(2) M. Schwartz, Computer Communication Network Design and Analysis, New York, Prentice Hall(1977).

(3) M. C. Pennoti and M. Schwartz, "Congestion control in store and forward tandem links," IEEE Trans. Comm., vol. COM-23, pp. 1434-1442, Dec. 1975.

(4) M. Reiser, "A queueing network analysis of computer communication networks with window flow control," IEEE Trans. Comm., vol. COM-27, pp. 1199-1209, Aug. 1979.

(5) M. Reiser and S. S. Lavemberg, "Mean value analysis of closed multichain queueing networks," IBM Research Report RC 7023, Yorktown Heights, N. Y., 1978.

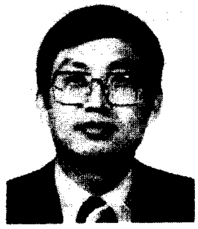
(6) H. Kekre, et al., "Buffer behavior for mixed arriuals and single server with random interruptions," IEEE Trans. Comm., vol. COM-28, pp. 59-64, Jan. 1980.

(7) J. W. Cohen, The Single server Queue, New York, Wiley(1969).

(8) General Purpose Simulation System, User's Manual, IBM Corp., 1977.

(9) S. Park and C. K. Un, "A simulator for integrated voice /data packet communication networks," J. of Korea Inst. of Commun. Sci., vol. 11, no. 2, pp. 108-121, Apr., 1986.

孫 壽 炫(Soo Hyeon SON)　正會員
1960年 5月 3日生
1982年：서울大學校 電子工學科 卒業 (B. S)
1984年：韓國科學技術院 電氣및 電子工學科 卒業(M. S)
1984年～現在：金星電氣(株) 研究所 研究員



殷 鍾 官(Chong Kwan UN)　正會員
1940年 8月25日生
1964. 6：미국University of Delaware 電子工學科卒業, 電子工學學士 學位
1966. 6：同大學院卒業, 電子工學碩士學位
1969. 6：同大學院卒業, 電子工學博士學位
1969. 9～1973. 5：美國University of Maine電子工學助教授
1973. 5～1977. 6：美國스텐포드研究所(SRI) 責任研究員
1977. 6～現在：韓國科學技術院電氣및電子工學科 教授, 本學會 理事
研究分野：디지탈通信, 디지탈信號處理