INTEGRATION OF IP-PACKET/ATM/CIRCUIT-SWITCHING TRAFFIC AND OPTIMUM CONDITIONS OF THE RELATIVE COST FUNCTION FOR TYPICAL TRAFFIC MODES

Noriharu Miyaho
NTT Service Integration Laboratories
miyaho@rdh.ecl.ntt.co.jp

Abstract An integrated switching-system architecture is proposed in which IP-packet, X.25-packet, ATM, and circuit-switching traffic can be handled simultaneously. It is based on a hierarchical memory system. The optimum boundary condition for frames sent over a TDM line is discussed from the viewpoint of minimum-cost requirements. Effectively utilizing transmission lines in an international-connection environment is particularly important when integrating various kinds of multimedia traffic to be sent over one line. Typical examples with key traffic-characteristic parameters are discussed and evaluated for determining optimum bandwidth allocation.

Keywords: Integrated switching, IP, ATM, Circuit switching, TDM boundary optimization Cost function

1. INTRODUCTION

A previously proposed mechanism [1] enables packet-switching (PS) and circuit-switching (CS) functions to be integrated in a switching-system architecture by using a hierarchical memory system. Various types of data traffic with different bearer rates can then be time division multiplexed (TDM) onto a single transmission line by setting an appropriate boundary in each frame to be sent over a TDM line to separate the two switching functions.

Communication services are generally specified as either guaranteed or best effort in terms of communication quality and routing mechanisms [2]-[3]. The QoS service classifications for circuit, packet, ATM, and frame-relay switching are shown in Table 1.

D. H. K. Tsang et al. (eds.), *Broadband Communications*

Table 1 QoS Service classification

Switching type		CO/CL	QoS guarantee	Delay characteristics
Circuit switching		CO	Guaranteed	Excellent (guaranteed to be constant and small)
Packet switching	X.25	CO	Guaranteed (virtual circuit)	Poor (not guaranteed)
	IP	CL	Best effort	Very poor (not guaranteed)
ATM switching	CBR, GFR	CO	Guaranteed	Good (not guaranteed)
	UBR,VBR, ABR	CO	Best effort	Poor (not guaranteed)
Frame relay		CO	Best effort	Good (not guaranteed)

CBR:Constant Bit Rate GFR:Guaranteed Frame Rate UBR:Unspecified Bit Rate VBR:Variable Bit Rate ABR:Available Bit Rate
CO: Connection-Oriented CL:Connectionless

Connection-oriented packet switching has traditionally been used to provide reliable communication services, but with the introduction of optical fibers, conventional packet switching is being replaced by frame-relay switching, particularly for transmitting large files and interconnecting LANs. Using frame-relay switching eliminates the need for retransmission control and hence achieves more efficient packet-data transmission. In Japan, the number of subscribers to packet-switching services, excluding DDX-TP (which is available via conventional telephone networks), was more than 300,000 at the end of 1998, while that to frame-relay services was only 50,000. However, the latter increased more than twice from the previous year, while demand for the former is subsiding.

Connection-less (CL) IP packet-switching service is deemed a best-effort service because it does not handle retransmission in the network layer when an error occurs; instead, error recovery is handled in an upper transport layer, such as the transmission-control-protocol (TCP) layer. Therefore, the CL function is considered to be useful for accommodating the multimedia databases that will be used in access networks.

The demand for Internet protocol (IP) services is increasing. These services currently include non-real-time and comparatively low-speed communication services, such as electronic mail (using SMTP), file transfer (using SMTP), and WWW access (using HTTP). Higher-speed and QoS-guaranteed communication services using IP should become available in the near future. In Japan, the number of subscribers to IP services over the Internet now exceeds ten millions; however, service quality has been degraded by the increasing traffic volume and limited router-processing capacities. To handle multimedia traffic with a wide range of communication speeds with QoS control, ATM switching is generally recognized as an efficient solution;

however, when ATM switching is used, it is a little bit difficult to ascertain the real-time traffic characteristics, such as for high-quality-voice and hi-fi stereophonic-sound traffic. This means that even if most multimedia traffic is handled using packet, ATM, or frame-relay switching to support efficient utilization of the transmission line and to ensure sufficient communication quality, genuine real-time traffic should sometimes be handled as such by using comparatively narrow-band circuit switching. This narrow-band switching should be appropriately applied to international connections because of the high cost of using international transmission lines. High-quality hi-fi stereophonic sound and highly bandwidth compressed voice are considered to be examples of genuine real-time traffic.

Taking the above discussion into account, we should handle non-real-time traffic, such as packet-switching (PS) traffic, and strict real-time traffic, such as high-quality voice and sound traffic, simultaneously. While ATM-mode traffic, such as real-time variable-bit-rate traffic, can currently be handled as real-time traffic, ATM mode is still inferior to CS mode.

In this paper I propose an integrated switching-system architecture in which PS, ATM, and CS traffic are integrated based on a hierarchical memory system. I also present a method for determining the optimum boundary conditions for a TDM transmission line. I focus on the conditions for a 2-Mb/s TDM line from the viewpoint of cost.

Integration of the IP-packet processing with layer 4 control and of ATM traffic processing with QoS control by using a hierarchical memory system concept is presented first. This paper also discusses optimum traffic multiplexing conditions from the viewpoint of minimum cost requirements.

2. PROPOSED INTEGRATED SWITCHING SYSTEM ARCHITECTURE

The basic concept of the proposed integrated switching-system architecture is presented in Fig. 1. The ATM traffic-handling function is included in addition to the IP-packet (including X.25)/CS traffic-handling functions by defining the following seven features.

(1) The boundary in a TDM packet is initially set assuming circuit-switching and other types of traffic. It is adjusted to match the variations in traffic types by using a movable boundary scheme [4]-[5].

(2) IP/X.25-packet/ATM-cell and circuit-switching functions are performed by sharing the same hierarchical memory structure that is conventionally used for program storage and execution in a general-purpose computer. The ATM traffic and PS traffic are handled by making use of a specific hardware mechanism inside the high-speed integrated switching storage (ISS). Switching program execution is performed so that instructions stored in the high-speed

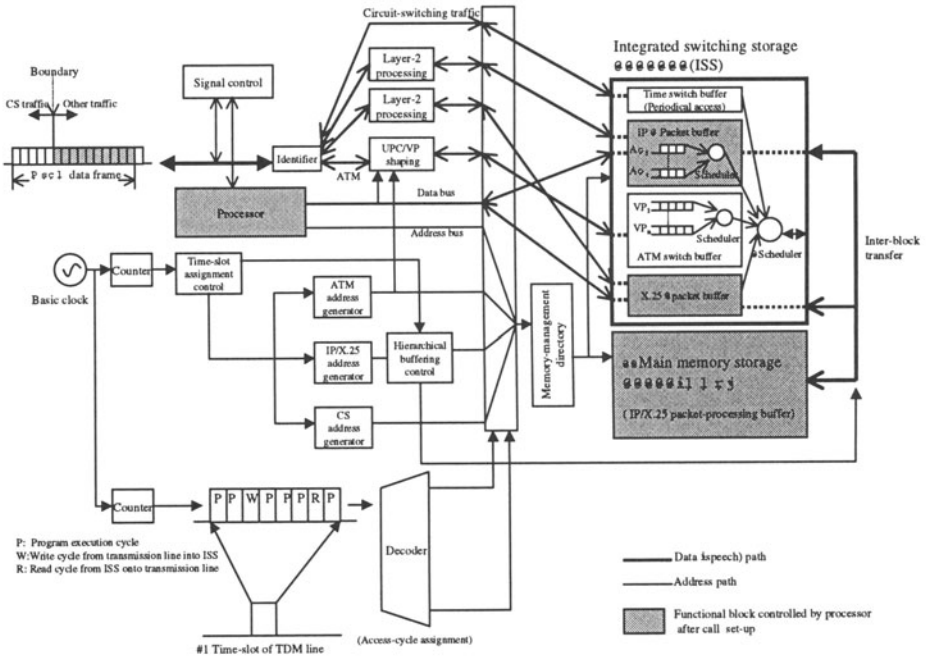

Fig. 1 Proposed integrated switching-system architecture.

ISS are read out and processed. When the required information (instructions or data) is not stored in the ISS, the appropriate data block is fetched from the large-capacity main memory storage, processed, and stored in the ISS so as to be available for subsequent processing. This process is the same as cache processing in a general-purpose computer.

(3) The communication channels (with multiple time-slots) assigned for PS, ATM, and CS traffic are multiplexed, and each time slot is identified during call set-up as containing a PS, CS, or ATM call. Doing this requires that the calling terminal send the call-identification signaling information in advance to the switching node. This information is sent through a control-signal channel, as in the ISDN user-network interface.

(4) The IP/X.25 traffic-processing function processes the appropriate data stored in the ISS. When the IP-packet /ATM traffic density is low and the data buffer areas for these traffic types in the ISS have vacancies, the processor handles these types of traffic on a first-in first-out (FIFO) basis for each service-type queue, sending the data into the outgoing TDM transmission line.

(5) When IP/X.25 packets arrive with a long service time, such as for a mail-box service, or when there is a shortage of buffer space for these packets in the ISS due to a large number of packets arriving simultaneously, the data packets are transferred to the MMS on a FIFO basis. In contrast, when there is a shortage of ATM-traffic buffer space, ATM cells are lost in the ISS,

which leads to degradation in the ATM QoS.
(6) A CS call periodically accesses the same area of the CS buffer space (which is equivalent to a time switch) in the ISS until a disconnection signal is received by the signal control block and processed by the processor. The CS function is thus achieved by this time-slot interchanging action. Therefore, it can easily guarantee a short and constant delay, with real-time characteristics. In contrast, PS traffic is processed continuously by the processor according to the corresponding communication protocol, such as error recovery/re-transmission and virtual-to-real address conversion for the corresponding program execution.
(7) ATM traffic can be handled in a manner similar to that for CS traffic, except it also needs usage parameter control (UPC), VP shaping control, and required-QoS control for each virtual path (VP). Also, the time-slot assignment for ATM traffic in a TDM frame for each specific call is not fixed. Only ATM traffic needs these mechanisms ; other types of traffic do not need them. ATM traffic also follows the Q.2931 protocol.

In the proposed switching architecture, IP packets can be handled in both the IP layer and the TCP/UDP layer by making use of the common processor. A conventional router can normally handle only IP addresses (by referring to a routing table), but in this proposed architecture, it can also handle the port numbers included in the TCP/UDP header. For each IP packet call, after identifying the port number (which normally determines a specific application server; WWW, FTP, etc.), the processor stores the IP packet in the corresponding service queue buffer (A1-Am). Because this priority-control mechanism can be further enhanced by combining the IP address, the type of service (which is included in the IP header), and the port number, it is effective for handling application- or protocol-specific priority control. This differs substantially from the end-to-end user connection-oriented QoS control normally used in ATM switch networks, in which several priority-specific VP buffers (VP_1-VP_n) are used.

This proposed design concept is also applicable to a single-stage memory system if the ISS buffer is large enough to handle the packet traffic and the required packet-switching speed is comparatively low. Figure 2 shows the ISS access cycle flow based on the above discussion.

Integrated PS/CS/ATM communications can be made available by selectively offering packet- or circuit-switching services that correspond to the user's rapidly changing requests. To adequately respond to these requirements, the PS/CS/ATM integrated switching concept needs to be carefully examined from the perspectives of economy and flexibility.

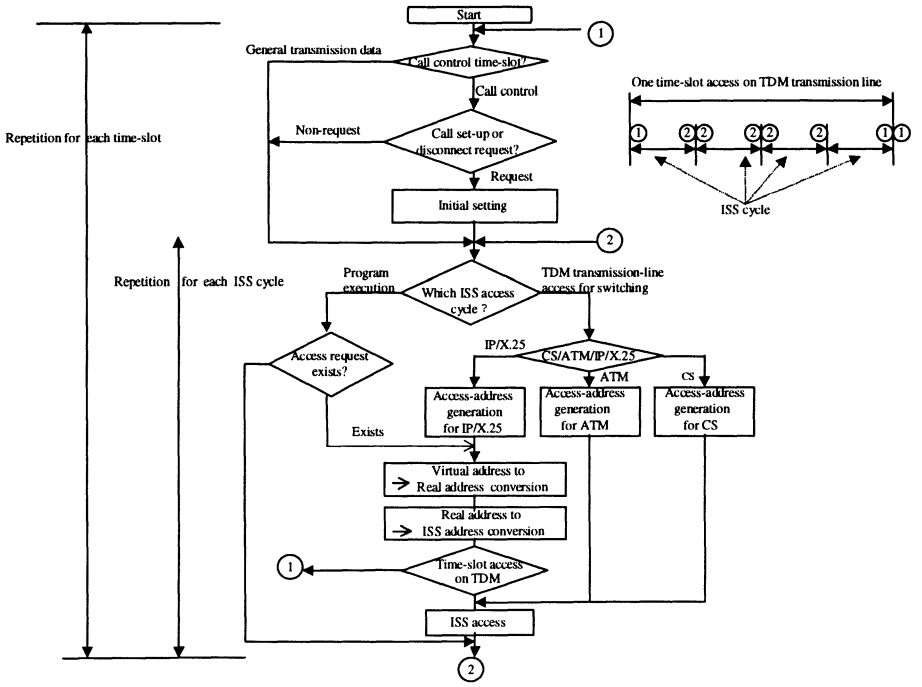

Fig.2 Proposed ISS access flow.

3. DETERMINING THE OPTIMUM BOUNDARY CONDITION

To determine the optimum boundary condition, we define the costs of transferring B bits of information per day by circuit switching and by packet switching as Cc and Cp, respectively. Assuming the parameters defined in Table 2, we can derive the following equations.

$$C_c = A_1 \bullet B / D + A_2 \bullet B/ (R_1\rho) \quad (1)$$

$$C_p = A_3 \bullet B/P + A_4 \bullet B/D + A_5 \bullet B/\lambda \quad (2)$$

By using these formulae, we can minimize newly defined relative cost Zrel. Let Ncs and Nps be the number of channels assigned to circuit and packet switching, respectively, and let q be the basic channel bearer rate (e.g., 64 kb/s). If a primary-rate E1 transmission line is assumed, we get Ncs + Nps = m (constant=30).

The parameters R_1, R_2, and λ in Table 1 are defined as follows.

$$R_1 = q \bullet N_{cs} \quad (3)$$

$$R_2 = q \bullet N_{ps} \quad (4)$$

$$\lambda = q \bullet N_{ps}/P \quad (5)$$

Thus, the relative cost Z_{rel} is expressed as follows.

$$Z_{rel} = (C_c + C_p)/B = A_1/D + A_2/(R_1\rho) + A_3/P + A_4/D + A_5/\lambda$$

$$= [(A_1+A_4)/D + A_3/P] + A_2/(q \bullet N_{cs} \bullet \rho) + (A_5 \bullet P)/q \bullet (m - N_{cs}) \quad (6)$$

Although Ncs is actually an integer (≥0), we assume it is a real number and that Zrel is a continuous function of Ncs. The minimum value of Zrel is found by setting the derivative of Zrel equal to zero (i.e., Zrel'=0) and calculating (Ncs)opt, the corresponding value of Ncs for $0 \leq Ncs \leq m$.

The (Ncs)opt that gives the minimum Zrel under these conditions is as follows.

$$(N_{cs})_{opt}=(m/q)\bullet[(A_2\bullet A_5\bullet P/\rho)^{1/2}-A_2/\rho]/[A_5\bullet P/q-A_2/(q\bullet\rho) \quad (7)$$

Table 2 Parameters for determining for optimum boundary condition.

A_1	Circuit-switching processing cost per call (unit/call)
A_2	Circuit-switching speech-path channel cost per sec (unit/s)
A_3	Packet-switching processing cost per packet (unit/packet)
A_4	Packet-switching processing cost per virtual call (unit/call)
A_5	Packet data buffer occupation cost (unit/ bit ·s)
B	Total information amount per day (bit/day)
C_c	Relative cost of circuit switching for transmission of B bits per day
C_p	Relative cost of packet switching for transmission of B bits per day
D	Total information amount per call (bit/call)
P	Packet size (bit/packet)
R1	Line speed for circuit switching (bit/s)
R_2	Line speed for packet switching (bit/s)
λ	Reciprocal of holding time for buffer occupation for packet switching
ρ	Line utilization rate for circuit switching (time required for information transmission divided by line holding time)

4. EVALUATION

To evaluate the effectiveness of the method proposed for determining the optimum boundary condition, let us consider practical examples in which we assume appropriate traffic parameters for Eq. (6). We also assume that either PS or CS traffic is dominant. Table 3 shows seven typical traffic-model cases in which PS and CS traffic are appropriately combined, taking current communication-service trends into account.

Based on the results for these cases, in which either the CS or PS traffic is dominant depending on certain conditions, the optimum boundary conditions can be determined.

Table 3 Models for evaluating method for determining optimum boundary condition

Case	Model	PS characteristics	CS characteristics
1	X.25 + CS	Short packet (128 B)	Low traffic density (0.1) Medium speed (64 Kb/s)
2	X.25 + CS	Long packet (4096 B)	Low traffic density (0.1) Medium speed (64 Kb/s)
3	IP + CS	IP packet (576 B)	Low traffic density (0.1) Medium speed (64 Kb/s)
4	IP + CS	IP packet (576 B)	High traffic density (0.8) Low speed (8 Kb/s)
5	IP + CS	IP packet (576 B)	High traffic density (0.8) High speed (128 Kb/s)
6	ATM + CS	ATM cell (53 B)	Low traffic density (0.1) Medium speed (64 Kb/s)
7	ATM + CS	ATM cell (53 B)	High traffic density (0.9) Medium speed (64 Kb/s)

Accurate values for parameters A_1 to A_5 are not needed to evaluate the relative costs or to obtain the conditions for minimizing transmission costs. We can set the same approximate value for parameters A_1 and A_4 based on the author's experience with many types of packet, circuit, and ATM switching-system implementations. Generally, A_2 is of the same order as A1. The important thing is setting the values for A_3 and A_5 compared to those for A_1 and A_2. The absolute value of A3 is normally one order of magnitude lower than that for A_4, which depends on the application service and traffic characteristics. The absolute value of A_5 is probably one order of magnitude lower given recent trends in the development of economical high-speed random access memory. Therefore, the values of parameters A_1 to A_5 were set based on the above reasoning and were the same for the seven examined cases, enabling extraction of the essential requirements.

To evaluate cost function Zrel, a utilization of 2 Mb/s was assumed for an international dedicated line, whose running cost is far higher than that of other communication facilities. Assuming a frame structure of 2.048 Mb/s, 30-B (64 Kb/s) channels were assumed because one frame (= 125μs) is composed of 32 64-Kb/s time-slots (time-slot #0 - #31); #0 is used for synchronization and #16 is used for signaling (D-channel).

Relative cost Zrel is approximately expressed as follows for the seven traffic models, assuming the parameter settings shown in Table 3. In cases 1 to 3 and 6 for normal CS-application services, real-time voice and interactive large-file transmission applications are assumed. The corresponding required average transmission speed is 64 Kb/s and the traffic density is 0.1.

Case 1: X.25 short packet + medium-speed CS

A_1=100 (unit), A_2=100 (unit), A_3=10 (unit), A_4=100 (unit), A_5=0.1 (unit),

D=1 (Mb), P=128 (B), q=64 (Kb/s),ρ = 0.1, Ncs+Nps=30

Zrel = 2 + 200/(64Ncs) + 20/{64•(30-Ncs)} (8)

Case 2: X.25 long packet + medium-speed CS

A_1=100 (unit), A_2=100 (unit), A_3=10 (unit), A_4=100 (unit), A_5=0.5 (unit),

D=1 (Mb), P=4096 (B), q=64 (Kb/s), ρ = 0.1, Ncs+Nps=30

Zrel = 0.1 + 200/(64Ncs) + 51/(30-Ncs) (9)

Here, 576 Byte size of IP packet is assumed according to the specification on RFC879 on TCP/IP case.

Case 3: IP packet + medium-speed CS

A_1=100 (unit), A_2=100 (unit), A_3=10 (unit), A_4=100 (unit), A_5=0.5 (unit),

D=1 (Mb), P=576 (B), q=64 (Kb/s), ρ = 0.1, Ncs+Nps=30

Zrel = 0.44 + 200/(64Ncs) + 92/64(30-Ncs) (10)

Case 4: IP packet + low-speed CS

A_1=100 (unit), A_2=100 (unit), A_3=10 (unit), A_4=100 (unit), A_5=0.5 (unit),

D=1 (Mb), P= 576 (B), q=8 (Kb/s), ρ = 0.8, Ncs+Nps=240

Zrel = 0.44 + 200/(64Ncs) + 740/64(240-Ncs) (11)

For Case 4, because the basic time-slot speed in CS is 8 Kb/s, the total number of equally available time-slots is assumed to be 240 (= 30 x 8). Because an 8-Kb/s-bandwidth efficiently compressed-voice technology[6] (conforming to ITU-T G.729, Conjugate Structure Algebraic CELP) is assumed, the corresponding traffic-density parameter should be set to a large value (such as 0.8). To make this case comparable with the other cases, the optimum boundary range of 0 to 240 is converted to 0 to 30.

Case 5: IP packet + high-speed CS

A_1=100 (unit), A_2=100 (unit), A_3=10 (unit), A_4=100 (unit), A_5=0.5 (unit),

D=1 (Mb), P= 576 (B), q=64 (Kb/s), ρ = 0.8, Ncs+Nps=15

Zrel = 0.44 + 200/(1024Ncs) + 46/84(15-Ncs) (12)

For Case 5, since the basic time-slot speed in CS is 128 Kb/s, the total equivalently available number of time-slots is assumed to be 15(=30 / 2). Because a 128-Kb/s time-slot is assumed to be used for high-quality stereophonic sound, the corresponding traffic-density parameter should be set to a large value (such as 0.8). To make this case comparable with the other cases, the optimum boundary range of 0 to 15 is converted to 0 to 30.

Case 6: ATM cell + medium-speed CS

A_1=100 (unit), A_2=100 (unit), A_3=10 (unit), A_4=100 (unit), A_5=0.5 (unit),

D=1 (Mb), P= 53 (B), q=64 (Kb/s), ρ =0.1, Ncs+Nps=30

Zrel = 4.7 + 200/(64Ncs) + 0.66/(30-Ncs) (13)

For Case 6, the ATM traffic was assumed to have a fixed packet size for evaluation simplicity.

Case 7: ATM cell + medium-speed CS

A_1=100 (unit), A_2=100 (unit), A_3=10 (unit), A_4=100 (unit), A_5=0.5 (unit),

D=1 (Mb), P= 53 (B), q=64 (Kb/s), $\rho = 0.9$, Ncs+Nps=30

$$Zrel = 4.7 + 200/(576Ncs) + 0.66/(30-Ncs) \quad (14)$$

The equations for the seven cases can be illustrated in Figs. 3 (a) to (g) for evaluating the optimum boundary. Several points should be noted.

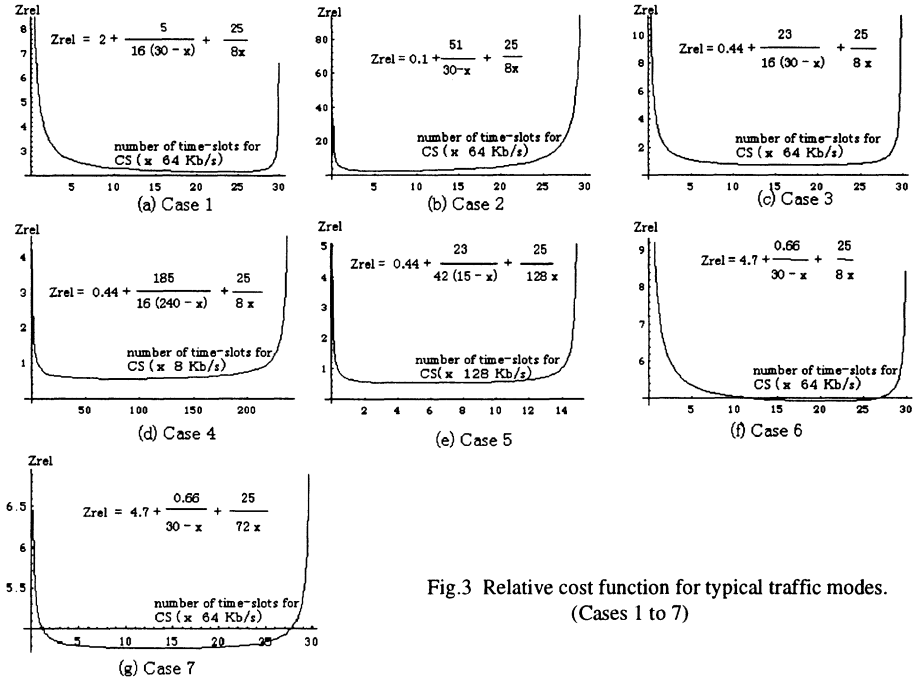

Fig.3 Relative cost function for typical traffic modes. (Cases 1 to 7)

(1) When the CS processing cost exceeds that of PS, i.e., in the X.25 short-packet case, the initially assigned boundary range for CS is comparatively wide, and the required number of time-slots for CS is large (17 to 27).

(2) When the PS processing cost exceeds that of CS, i.e., in the X.25 long-packet case, the initially assigned boundary range for CS is comparatively narrow (3 to 10), and the required number of time-slots for CS is small.

(3) When the required speed for CS is high, the initially assigned boundary range is comparatively wide; however, the optimum value assigned for CS traffic is approximately the same as in the low-speed (8-Kb/s) case if the corresponding traffic density is the same.

(4) When the traffic density of CS increases, the assigned number of time-slots for CS can be increased to achieve minimum cost on the condition that ATM traffic is considered to be a fixed size short packet.

After the initially assigned channel allocation between the switching functions has been determined based on the proposed method, it can be adjusted to match the variations in traffic types by using a movable boundary

scheme [4].

The proposed concept can also be applied to a mixture of frame-relay traffic. In this paper, ATM traffic of UBR class (best-effort type) was assumed for analytical simplicity because ATM traffic is considered to be equivalent to fixed-length-packet traffic.

5. CONCLUSION

A new concept was proposed for a switching system architecture in which packet-switching, ATM, and circuit-switching traffic with QoS control are integrated based on a hierarchical memory system. The optimum boundary condition for frames sent over an expensive TDM transmission line on which PS, ATM, and CS traffic are integrated was evaluated by setting several appropriate cost-related parameters. The PS processing cost parameter was found to be more sensitive than the CS one. This proposed method can also be applied to a mixture of CS, PS, ATM (UBR class), and frame-relay traffic.

Future studies should consider the allowable call-blocking ratio in the CS case and the allowable maximum buffering delay in the PS case in addition to the ATM traffic characteristics during burst mode in order to precisely define the optimum boundary condition. In addition, the effective ATM cell handling technology for versatile cell transfer mode such as ABR, VBR, and GFR should also be investigated.

References

[1] Ishikawa, H., Kosuge, Y., and Miyaho, N., "A Study on Hybrid Switch Architecture Using Hierarchical Memory System", IEICE Trans.Commun., Vol. 68-B, No. 1, pp. 30-37, Jan., 1985.

[2] Lee, W. C., Hluchy, M.G., and Humblet, P.A., "Routing Subject to Quality of Service Constraints in Integrated Communication Networks", IEEE Network, July/Aug., pp. 46-55, 1995.

[3] Mase, K. and Kimura, "Quality of Service Issues in Communication Networks," IEICE B-1, Vol. J80-B-I, pp. 283-295, June 1997.

[4] Niitsu, Y., "Evaluation of the Movable Boundary Scheme in Circuit and Packet Switched Networks", Transactions of IEICE, Vol. J68-B, No. 10, 1985.

[5] Miyaho, N. and Miura, A., "Integrated Switching Architecture and its Traffic Handling Capacity in Data Communication Networks", IEICE Trans. Commun., Vol. E-79-B, No. 12, pp.1887-1899, Dec., 1996.

[6] R. Salami, et.al., "Design and description of CS-ACELP : Toll quality 8Kb/s speech coder," IEEE Trans. Speech on Audio Processing, Vol.6, No.2, pp116-130,1998.