

# ATM 교환기의 프로세서간통신을 위한 메시지 라우팅 방법

박혜숙, 문승진, 박만식, 송광석, 이형호

한국전자통신연구원

## Message Routing Method for Inter-Processor Communication of the ATM Switching System

HeaSook Park, SungJin Moon, Mansik Park, KwangSuk Song and Hyeong Ho Lee

Electronics and Telecommunications Research Institute (ETRI)

E-mail : parkhs@etri.re.kr

### Abstract

This paper describes an interconnection network structure which transports information among processors through a high speed ATM switch. To efficiently use the high speed ATM switch for the message-based multiprocessor, we implemented the cell router that performs multiplexing and demultiplexing of cells from/to processors. In this system, we use the expanded internal cell format including 3bytes for switch routing information. This interconnection network has 3 stage routing strategies: ATM switch routing using switch routing information, cell router routing using a virtual path identifier (VPI) and cell reassembly routing using a virtual channel identifier (VCI). The interconnection network consists of the NxN folded switch and N cell routers with the M processor interface. Therefore, the maximum number of NxM processors can be interconnected for message communication. This interconnection network using the ATM switch makes a significant improvement in terms of message passing latency and scalability. Additionally, we evaluated the transmission overhead in this interconnection network using ATM switch.

### 1. 서 론

분산된 다중프로세서 구조에서는 프로세서간의 빠른 정보 전달이 시스템의 전체 성능을 높이는 중요한 인자이다. ATM 교환기시스템의 주요 기능인 호처리, 과금, 및 유지보수 기능들은 IPC 메시지를 통해 이루어진다. 따라서 IPC 메시지는 고속의 ATM망을 통해 전달되는 경향

이 있다[1].

고성능, 낮은 지연이라는 특성을 가지는 ATM은 B-ISDN(Broadband Integrated Service Digital Network)을 위한 전달/교환 방법으로 채택되고 있다. ATM 망에서는 셀이라는 고정된 길이의 패킷이 전달단위이다. 한 셀은 5바이트의 헤드 정보와

48바이트의 데이터로 구성된다. 셀 헤드 정보는 셀을 분해/조립하는데 사용되며, 셀의 구별인자인 VPI(Virtual Path Identifier)와 VCI(Virtual Channel Identifier) 등의 정보를 포함한다[2,3].

본 논문의 ATM 교환기에서는 ATM 스위치를 통해 프로세서간 통신이 이루어진다. 따라서 정보 전달의 단위인 메시지가 ATM 스위치를 통해서 전달되기 위해서는 셀로 분해되어야 하며, 전달된 셀들은 메시지로 조립되어 응용프로그램으로 전달된다. 즉, 메시지를 송신/수신하는 프로세서는 셀 분해/조립 (SAR:Segmentation And Reassembly) 기능을 포함하고 있어야 한다. 또한, 스위치 라우팅을 위해 표준 ATM 셀에 라우팅 정보를 추가함으로써 VPI와 VCI는 셀을 구별하기 위해 프로세서를 구별하는 목적으로 사용할 수 있다. 따라서 본 논문의 3단계 라우팅은 ATM 스위치 라우팅 정보, VPI와 VCI에 의해 이루어진다.

셀 다중화/역다중화 장치를 통하여 다중 프로세서를 하나의 스위치 링크에 연결함으로써 프로세서 규모의 확장성 및 가변성이 용이한 계층적 구조를 가진다. 이러한 구조에서 프로세서간 통신 메시지는 3단계의 라우팅을 통해 송신측 프로세서에서 착신측 프로세서로 전달된다. 송신측에서 셀로 분해되어 스위치로 전달된 메시지는 스위치 라우팅 태그로 라우팅 되어 착신측 서브시스템의 셀 다중화/역다중화 장치로 전달된다. 셀 다중화/역다중화 장치는 셀의 VPI 값을 VPI 테이블에서 참조하여 셀을 전송할 프로세서의 포트를 결정한다. 프로세서로 수신된 셀은 VCI 값에 의해 메시지로 조립되어 응용

프로그램으로 전달된다. 위의 3단계 라우팅 중에서 스위치 라우팅을 위해 본 논문에서는 56바이트로 확장된 내부 셀 포맷을 사용한다. 결론적으로 NxN ATM 스위치에 N개의 셀 다중화/역다중화 장치를 연결하고 각각의 셀 다중화/역다중화 장치에 M개의 프로세서를 연결하면 최대 NxM개 프로세서간의 메시지 통신이 가능하다. 즉, 시스템의 구조에 따라 셀 다중화/역다중화 장치와 프로세서의 수를 결정할 수 있다.

ATM 스위치를 메시지 전달망으로 사용함으로써 고속의 정보 전달이 가능하지만, 메시지 오버헤드 문제가 있다. 본 논문에서는 IPC 포맷으로 AAL (ATM Adaptation Layer) 타입5 포맷을 사용하는데 메시지를 셀로 분해/조립하기 위한 ATM 헤드정보, AAL 타입5 트레일러와 셀 단위로 전송을 위한 패딩 데이터는 ATM을 사용함으로써 생기는 오버헤드이다.

본 논문의 구성은 다음과 같다. II장에서는 본 논문의 메시지 라우팅 방법이 적용되는 ATM 교환 시스템의 구조에 대하여 설명한다. III장에서는 ATM 스위치, 셀 다중화/역다중화 장치, 그리고 프로세서의 셀 분해/조립 장치를 통한 메시지의 3단계 라우팅 방법에 대해 설명한다. IV장에서는 3단계 라우팅 방법의 분석과 IPC 메시지의 오버헤드에 관해서 분석한다. 그리고, V장에서는 결론 및 본 논문의 발전 방향을 기술한다.

## II. ATM 교환시스템의 구조

그림 1은 ATM 교환시스템의 구조를 보여준다. CDM<sub>i</sub>는 i번째 셀 다중화/역다중화 장치를 나타내고 P<sub>ij</sub>는 i번째 셀 다중화/역다중화 장치에 연결된 j번째 프로세서를 표시한다. 프로세서 NxN의 이중화된 ATM 스위치는 셀의 스위치 라우팅 정보를 분석하여 해당 ATM 링크로 셀을 전달한다.

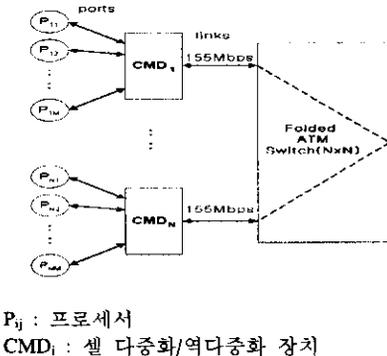


그림 1. ATM 교환시스템의 프로세서 연결도

스위치에 155Mbps 링크로 연결된 셀 다중화/역다중화 장치는 고속의 ATM 링크와 저속의 다중프로세서를 연결하는 인터페이스를 제공하는 셀 처리장치이다. 그림 2는 셀 다중화/역다중화 장치의 간략한 기능구조도이다.

분산 시스템의 규모와 처리 용량에 따라 1-M개까지의 프로세서가 정합될 수 있는 확장성을 가진다. 셀 다중화/역다중화 장치는 다중프로세서에서 수신된 셀을 ATM 스위치로 전달하는 다중화 기능과 ATM 스위치에서 수신된 셀을 다중프로세서로 전달하는 역다중화 기능을 가진다. 프로세서에서 ATM 스위치로 셀을 전달할 경우는 셀을 바이패스하고, ATM 스위치에서 수신된 셀을 다중프로세서로 전달할 때는 셀의 헤드 정보인 VPI를 참조하여 셀을 해당 프로세서의 포터로 전달한다. 즉, 셀 다중화/역다중화 장치는 VPI 값에 의한 셀 라우팅을 수행한다. ATM 스위치의 링크로부터 수신된 셀들이 프로세서의 집합  $\{P_0, P_1, \dots, P_{M-1}\}$ 으로 전달되기 위해서 프로세서 포트를 선택해야 한다. VPI 테이블에 설정된 값을 참조하여 프로세서의 포트를 선택하며, 다중 포트 또는 모든 포트로 셀을 전달할 수도 있다. 따라서 셀 다중화/역다중화 장치는 점대점, 멀티캐스팅과 방송 셀을 프로세서로 전달할 수 있다. 셀이 전달되는 포트의 선택은 식 1로 결정된다. 아래 식에서  $MAXVPI$ 의 값은 256이다.

셀 전송 포트  

$$= \{ P_i \mid VPITable[x].i = 1, \quad (식 1)$$

$$0 \leq x < MAXVPI, 0 \leq i < M-1 \}$$

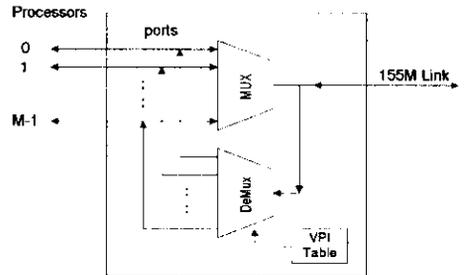


그림 2. 셀 다중화/역다중화 장치의 구성도

셀 다중화/역다중화 장치에 연결된 프로세서들은 메시지를 셀로 분해하여 전달하고, 수신된 셀을 메시지로 조립하는 SAR 기능을 가진다. 분산되어 있는 다중프로세서로부터 수신되는 셀들을 조립하기 위해서는 각각의 프로세서에서 전달된 셀들을 구별해야 한다. ATM에서 셀 구별 인자는 VPI와 VCI이다. 따라서 본 논문에서는 메시지 송신측 프로세서는 VCI로, 수신측 프로세서는 VPI를 가지고 구별한다.

## III. 메시지의 3단계 라우팅 방법

그림 3은 5바이트의 ATM 셀 헤드정보의 내용과 본 논문에서의 확장된 내부 셀 포맷을 보여주고 있다. 표준 셀 포맷에 3바이트의 스위치 라우팅 정보가 부가된 확

장된 포맷을 가진다.

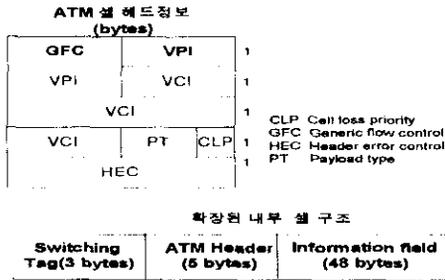


그림 3. 확장된 내부 셀 포맷

그림 4는 본 논문에서 사용하는 메시지의 포맷인 AAL 타입5의 CPCS-PDU (Common Part Convergence Sublayer-Protocol Data Unit) 포맷을 보여준다. 페이로드 부분은 64K바이트까지 확장가능하며, 8바이트의 트레일러는 메시지의 분해/조립을 위한 제어 정보들을 포함한다. PAD 필드는 CPCS-PDU를 48바이트로 정렬하기 위해 사용된다. 페이로드 부분은 16바이트의 메시지 헤드 정보와 순수 사용자 데이터로 구성된다. 16바이트의 메시지 헤드는 송신/수신 프로세서 식별자, 송신/수신 프로세스 식별자, 시그널 인자, 메시지 길이와 메시지의 우선순위 등의 정보를 포함한다.

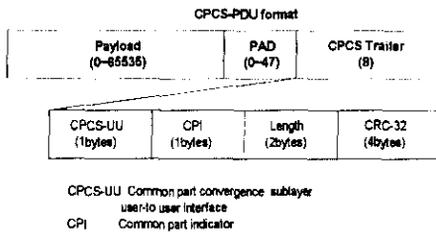


그림 4. 메시지 포맷(AAL 타입5 포맷)

본 논문에서 프로세서간의 메시지 라우팅은 3단계로 이루어진다. 그림 5는 메시지 라우팅의 구성도를 나타낸 것으로 ATM 스위치, 셀 다중화/역다중화 장치와 프로세서의 3계층에서의 라우팅을 위해 스위치 라우팅 정보, VPI와 VCI가 각각 맵핑되어 있음을 보여주고 있다. 그림 5에서는 N개의 셀 다중화/역다중화 장치에 각각 4개의 프로세서가 연결되어 있는 예이다. 첫 단계는 ATM 스위치에서의 셀 라우팅이다. 분산되어 있는 서브 시스템간의 경로를 설정하는 것으로  $SR_{1,N}$ 은 서브시스템 1에서 N으로의 스위치 경로이고,  $SR_{N,1}$ 은 서브시스템 N에서 1로의 경로이다.

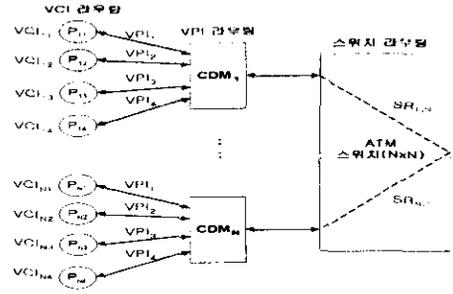


그림 5. 메시지 라우팅의 예

#### IV. 분석

본 논문에서의 3단계 라우팅은 많은 오버헤드를 가진다. 3바이트의 스위치 라우팅 태그, 1바이트의 VPI와 2바이트의 VCI를 고려하면 최대  $2^{48}$ 개 종류의 메시지를 라우팅할 수 있다. 그러나 본 논문의 프로세서간 라우팅 방법은 표준 ATM 셀포맷을 그대로 두고 스위치 라우팅 정보를 추가함으로써 시스템의 확장성과 유연성을 가질 수 있으며, VPI와 VCI는 한 서브시스템 내에서의 셀 라우팅과 메시지 조립을 위한 정보로 사용하였다. 따라서 VPI와 VCI는 스위치 라우팅에 사용되지 않는다.

한 메시지가 ATM 스위치를 통해 다른 프로세서로 전달되기 위해서는 셀로 분해되어야 하므로, ATM을 사용함으로써 발생하는 오버헤드 데이터는 ATM 헤드, AAL 타입5 트레일러, 패딩데이터 등이다. 한 메시지를 전달하기 위해 전송한 전체 데이터 크기에 대한 전송하고자 하는 순수 사용자 메시지의 크기의 비를 계산함으로써 오버헤드의 정도를 파악할 수 있다.

ATM 셀 전송을 통한 메시지 전송 오버헤드를 계산하기 위해 아래의 변수들을 정의한다.

- $L_{mh}$  메시지 헤드정보의 크기 (16바이트)
- $L_{tal}$  AAL 타입5의 트레일러 크기 (8바이트)
- $L_{cpyl}$  한 셀의 페이로드의 크기 (48바이트)
- $N$  확장된 내부 셀의 크기 (56바이트)
- $X$  한 메시지의 크기

크기  $X$ 의 한 메시지를 전달하기 위해 전송할 셀의 수  $C_X$ 는 식 2와 같다.

$$C_X = \left\lceil \frac{L_{mh} + X + L_{tal}}{L_{cpyl}} \right\rceil \quad (\text{식 } 2)$$

단,  $\lceil a \rceil$ 는  $a$  보다 크거나 같은 정수

식 2에 의해 크기  $X$ 의 메시지를 전달하기 위해 전

송할 전체 메시지의 크기  $T_X$ 는 식 3과 같다.

$$T_X = N \times C_X \quad (\text{식 3})$$

식 3에 의해 전송한 전체 메시지에 대한 순수 사용자 메시지의 비  $R_X$ 는 식 4와 같다.

$$R_X = \frac{X}{T_X} \quad (\text{식 4})$$

위의 식에서 다른 변수들의 크기는 상수이기 때문에 48바이트의 정렬을 위한 패딩 정보만이 0~47바이트의 가변 크기를 가진다. 따라서 사용자 메시지의 크기가 전송 오버헤드의 중요한 인자이다. 즉, 패딩 데이터의 크기를 줄이는 사용자 메시지의 크기로 데이터를 전달함으로써 전송 오버헤드를 줄일수 있다. 그림 6은 식 4에 의한 사용자 메시지 크기  $X$ 와  $R_X$ 의 관계를 그래프로 나타낸 것이다.  $(1 - R_X)$ 는 메시지 전송 오버헤드의 비로 사용자 메시지 크기가 300 이상일 경우 0.15 정도임을 그림 6에서 알 수 있다. 즉, 본 논문의 프로세서간 통신을 위한 메시지의 크기는 300 바이트 이상이어야 어느 정도 오버헤드를 줄일 수 있다.

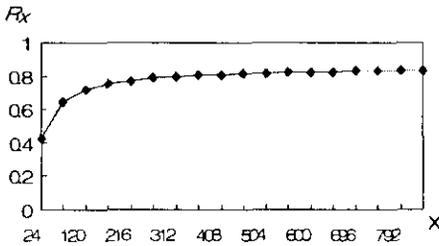


그림 6.  $X$ 와  $R_X$ 의 관계 그래프

### V. 결론

본 논문에서는 ATM 교환시스템의 분산구조에서의 ATM 스위치를 통한 프로세서간 메시지 통신에 관하여 설명하였다. 셀 다중화/역다중화 장치를 통하여 다중 프로세서를 하나의 스위치 링크에 연결함으로써 프로세서 규모의 확장성 및 가변성이 용이한 계층적 구조를 가진다. 이러한 구조에서 프로세서간 통신 메시지는 3단계의 라우팅을 통해 송신측 프로세서에서 착신측 프로세서로 전달된다. 위의 3단계 라우팅 중에서 스위치 라우팅을 위해 본 논문에서는 56바이트로 확장된 내부 셀 포맷을 사용한다. 결론적으로  $N \times N$  ATM 스위치에  $N$ 개의 셀 다중화/역다중화 장치를 연결하고 각각의 셀 다중화/역다중화 장치에  $M$ 개의 프로세서를 연결하면 최대  $N \times M$ 개

프로세서간의 메시지 통신이 가능하다. 즉, 시스템의 구조에 따라 셀 다중화/역다중화 장치와 프로세서의 수를 결정할 수 있다. 또한 메시지를 ATM 셀로 분해하여 전송함으로써 발생하는 메시지 전송 오버헤드의 비는 사용자 메시지의 크기가 300바이트 이상일 경우 약 15~20% 정도이다.

### VI. 참고문헌

[1] R. Sharma and S. Keshav, Signalling and Operating System Support for Native-Mode ATM Applications, *SIGCOMM 94*, London, England, pp. 149~157, 1994.  
 [2] Daniel Winkelstein, Daniel Stevenson, Supercomputer Communications as an Application for ATM Local Area Networks, *High-Speed Communication Networks*, pp. 15~24, 1992.  
 [3] Rainer Handel, Manfred N. Huber and Stefan Schroder, *ATM Networks: Concepts, Protocols, Applications*, Addison Wesley, 1994.