

우선순위제어기능을 가진 ATM스위치의 성능분석

正會員 張 在 申* 正會員 申 炳 喆* 正會員 朴 權 喆**

Performance Analysis of ATM Switch with
Priority Control MechanismsJae Shin Jang*, Byung Cheol Shin*, Kwon Chul Park** *Regular Members*

要 約

본 연구에서는 입력이 시간 지연에 민감한 트래픽과 손실에 민감한 트래픽으로 이루어진 경우의 우선순위 제어기능을 가진 출력버퍼형의 ATM 스위치의 성능을 분석하였다. 손실우선순위제어 메카니즘으로는 부분 버퍼공유 방법을 택했고, 시간우선순위제어 메카니즘으로는 HOL(Head Of Line)우선순위 제어 메카니즘을 택했다. 손실에 민감한 트래픽은 포아송 프로세스로 모델링하였고 시간 지연에 민감한 트래픽은 MMPP로 모델링하였다. 손실우선순위제어에서는 손실우선순위제어 메카니즘을 사용할때 손실에 민감한 트래픽의 셀손실률이 감소함을 확인하였고, 시간우선순위제어에서는 HOL 우선순위제어 메카니즘을 사용할때 시간 지연에 민감한 트래픽의 평균 시간지연이 감소함도 확인하였다. 이 결과로부터 시간우선순위제어와 손실우선순위제어를 동시에 사용하면 시간 지연에 민감한 트래픽과 손실에 민감한 트래픽의 서비스질을 동시에 만족하게 할 수가 있다는 것을 확인하였다.

Abstract

In this work, the performance of both delay sensitive traffics and loss sensitive traffics of the output buffered ATM switch with priority control mechanisms has been evaluated. We choose the partial buffer sharing mechanism as the loss priority control mechanism and the HOL(Head Of Line) priority control mechanism as the time priority control mechanism. We model loss sensitive traffics with Poisson process and delay sensitive traffics with MMPP. With loss priority control, it is confirmed that loss probability of loss sensitive traffice decreases when the loss priority control mechanism is chosen. With time priority control, it has also been confirmed that mean cell delay of delay sensitive traffics decreases when the HOL priority control mechanism is used. From this analysis, it has been confirmed that the requirements of QOS for both loss sensitive and delay sensitive traffics can be satisfied in the ATM switch by combining both the loss priority control mechanism and the HOL priority control mechanism.

*韓國科學技術院 電氣 및 電子工學科
Department of Electrical and Electronic Engineering,
KAIST

**韓國電子通信研究所
ETRI
論文番號 : 93 - 122

I. 서 론

기존의 다양한 정보통신서비스를 단일화된 통신망으로 통합하기 위해서 협대역 ISDN(Integrated Services Digital Network)이 출현하였지만 이것은 종래의 데이터와 음성을 위하여 64kbps의 협대역 서비스에 국한되어 있어서 64kbps이상의 동화상 서비스는 충족시키질 못했고, 또 멀티미디어에 대한 요구가 증대됨에 따라 BISDN(Broadband ISDN)에 관심이 집중되기 시작했다. BISDN의 핵심기술인 ATM(Asynchronous Transfer Mode)방식에서는 어떠한 종류의 트래픽이라도 셀(cell)이라고 하는 53바이트 길이의 패킷에 실어 전송된다. 예상되는 BISDN 트래픽으로는 상호교신성 데이터, 비디오 신호와 같이 특성이 bursty한 트래픽과 큰 화일정보와 같이 트래픽 특성이 연속적인 특성을 가진 트래픽이 있고 시간지연에 민감하면서 손실에는 덜 민감한 음성과 같은 신호가 있는 반면에 실시간 전송은 덜 요구되지만 손실에는 매우 민감한 데이터와 같은 신호도 있다. 이러한 다양한 트래픽들을 쉽게 수용하면서 채널사용효율을 증대시키기 위해 통계적 다중화방법을 사용하는데 이것 때문에 망에서는 폭주현상이 발생하게되며 버퍼의 오버플로우에 의한 셀손실이 발생한다. 그러므로 ATM에서는 망의 폭주에 의한 셀손의 최소화과 망자원 사용의 극대화란 서로 상충되는 목표를 달성하는게 주요 현안중의 하나이다.

CCITT에서는 5바이트의 헤더필드중에 1비트를 CLP 셀포기우선순위 비트로 할당되어 있는데 사용자에 의해서든지 혹은 망에 의해서 각 셀의 포기우선순위가 결정되며 낮은 우선순위로 할당된 셀은 CLP 비트가 1이므로 트래픽제어에 이용할 수가 있게된다.

본 연구에서는 음성 입력트래픽에서 실제의 특성을 더욱더 잘 반영하게 하기위해 MMPP(Markov Modulated Poisson Process)로 모델링하였고, 그때의 우선순위제어방법의 성능을 분석하였다. II 장에서는 ATM 스위치의 모델링에 대해 그리고 III 장에서는 우선순위기법에 대해 알아보며 IV 장에서는 우선순위기능을 가진 ATM 스위치의 성능을 분석하고 V 장에서는 성능분석된 결과를 제시하며 VI 장에서 결론으로 끝맺는다.

II. ATM 스위치의 모델링

ATM 스위치로는 그림2-1처럼 충돌이 발생하지

않는 공간분할스위치를 사용했다고 가정한다. 여기서 S는 스위치의 입력단자의 수를 나타낸다. 스위치 내부에서는 충돌이 발생하지 않을지라도 스위치가 통계적 다중화와 비슷하게 동작하기 때문에 한 슬롯타임동안에 동일한 출력주소를 가진 서로 다른 셀들이 각각 다른 입력을 통해 스위치의 어떤 출력단자에도 도착할 수가 있어 대기상태 또는 큐잉이 발생한다. 그래서 스위치내에는 버퍼를 설치하는 것이 필요하며 여기서 가정한 스위치는 출력버퍼형의 공간분할스위치이다. 스위치는 슬롯단위로 동기되어 동작하며 내부처리속도가 임출력 링크의 속도보다 S배 정도 빨라서 슬롯타임에 원하는 출력으로 모두 빠져나간다고 가정한다. 각 입력을 통해 들어온 셀들은 상호 독립이고 동일한 분포를 갖고 일단 스위치입력으로 들어온 셀들이 S개중 한개를 택해 나가는 것이 균일한 분포를 따른다면 성능분석을 위해 한 출력에만 초점을 맞출 수가 있다. 그러면 ATM 스위치는 그림2-2처럼 입력의 갯수는 S개이고 버퍼의 크기는 N으로 유한하며 서비스시간이 일정한 큐잉모델이 된다.

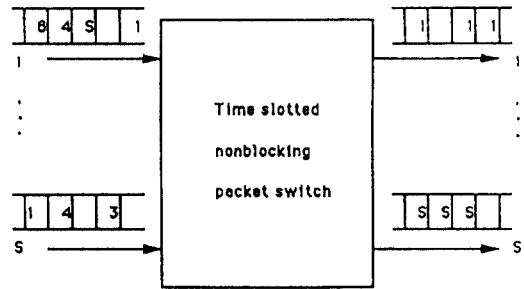


그림 2-1. 비충돌이 비블통 공간분할스위치
Fig 2-1. Nonblocking space division switch

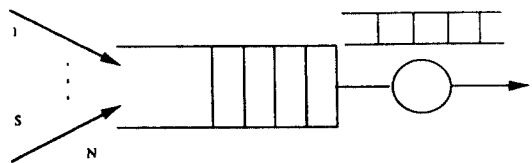


그림 2-2. 스위치 모델링
Fig 2-2. Modeling of the switch

Ⅲ. 우선순위제어방법

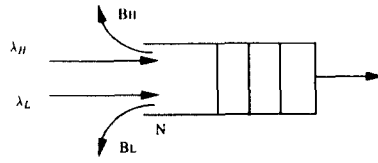
BISDN의 트래픽의 경우에 트래픽의 요구조건들이 서로 다르므로 요구되는 평균지연과 포기우선순위제어 값을 만족시켜주기 위해서 망에서는 우선순위제어를 하게된다. ATM 스위치에 우선순위제어기능을 구현하는 방법에는 시간우선순위제어방법과 손실우선순위제어방법이 있다.

1. 손실우선순위제어

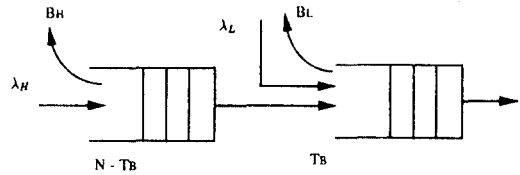
손실우선순위제어는 우선순위가 높은 셀들의 포기우선순위제어로 할당될 수 있다. 이 방법에는 푸쉬아웃 메카니즘과 부분 버퍼공유 메카니즘이 있다(그림 3-1 참조). 푸쉬아웃 메카니즘은 그림3-1(a)와 같이 버퍼가 찰때까지 모두 받아들이다가 버퍼가 차고난 뒤에는, 도착하는 셀중에 우선순위가 낮은 셀은 버리고 우선순위가 높은 셀들은 버퍼내에 우선순위가 낮은 셀이 있을 경우에는 이것을 밀어내고 들어가는 방법이다. 이 방법은 버퍼전체를 공유하기 때문에 뒤에 설명하는 부분 버퍼공유 방법에 비해 효율은 좋지만 버퍼관리가 상대적으로 복잡해서 ATM과 같은 고속전송에는 적합하지가 않는 것 같다. 부분적인 버퍼공유방법은 그림3-1 (b)와 같이 문턱값 T_B 까지는 버퍼를 공유하다가 큐의 길이가 T_B 를 넘어서면 우선순위가 높은 셀만을 받아들이는 바법이다. 이 방법은 푸쉬아웃방법에 비해 효율은 조금 떨어지지만 구현하기가 쉽고 하드웨어적으로 처리할 수가 있어 ATM과 같은 고속전송에 적합할 것으로 생각된다. 그런데 이 방법은 문턱값 근처에서 큐의 길이가 진동할 수가 있으므로 그림3-2와 같이 문턱값을 2개 두어 hysteresis 현상을 갖게한 변형된 부분적인 버퍼 공유 방법을 생각하게 되었다. 이 경우는 큐가 증가할 경우에는 L_2 이 문턱값으로 작용하고 일단 넘어섰다가 감소하기 시작하면 L_1 가 문턱값으로 작용하는 메카니즘이다.

2. 시간우선순위제어

시간우선순위제어방법은 버퍼내에서 우선순위가 높은 셀을 먼저 전송해주어 전체적으로 우선순위가 높은 셀의 큐잉지연시간을 줄이는 방법이다. 지연측에서는 우선순위가 높은 트래픽은 실시간 전송을 요구하는 음성이나 비디오같은 트래픽이며 호단위로 할당된다. 우선 여기서 가정하는 것은 호단위로 시간우선순위를 할당할 수 있고, 할당된 우선순위를 A-



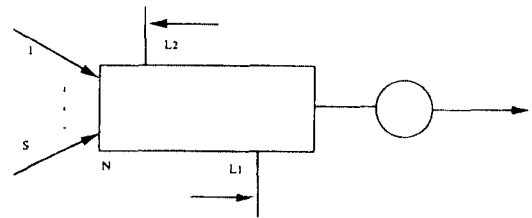
(a) Push-out 방법



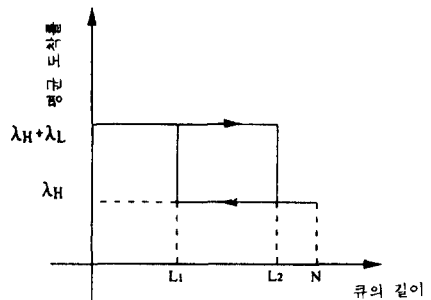
(b) Partial buffer sharing 방법

그림 3-1. 손실우선순위제어 메카니즘

Fig 3-1. Loss priority control mechanism



(a) 스위치 모델링



(b) 큐의 상태에 따른 평균 도착률의 변화

그림 3-2. 두개의 threshold를 가진 메카니즘

Fig 3-2. Mechanism with two thresholds

TM 스위치에서 인식할 수가 있으며 우선순위에 따라 처리도 가능하다는 것이다. 대표적인 우선순위 제어방법으로는 HOL 우선순위가 있으며 이 방법은 그림3-3와 같이 높은 우선순위의 셀들을 항상 먼저 서비스를 해주며 같은 우선순위의 셀들은 FIFO형태로 서비스한다. 새로이 도착하는 셀은 항상 같은 우선순위의 셀중에 제일 뒤에 위치하며 앞에 있는 셀들이 모두 서비스를 받고난 뒤에 서버에 들어간다. 본 연구에서는 nonpreemptive인 경우에 대해 성능분석을 했다.

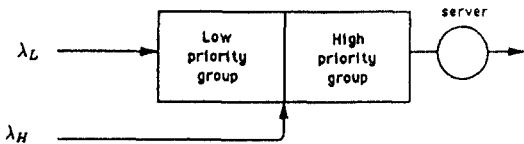


그림 3-3. HOL(Head Of Line) priority 방법의 모델링
Fig 3-3. Modeling of HOL priority mechanism

IV. 성능분석

1. 입력트래픽 모델링

손실우선순위제어에서 높은 우선순위 트래픽은 데이터와 같은 트래픽이고 낮은 우선순위 트래픽은 음성과 같은 트래픽이거나 변수제어기능에 의해 CLP 1로 주어진 트래픽들인데 여기서는 대부분 음성으로 이루어져 있다고 보고 그림4-1과 같은 2-상태 MMPP로 모델링했다. 그리고 높은 우선순위가 트래픽은 대부분 데이터와 같은 트래픽으로 이루어졌다고 보고 포아송 프로세스로 모델링했다. 그리고 MMPP와 포아송 프로세스의 중첩은 도착율의 합으로 구성된 새로운 MMPP로 볼 수 있다[2].

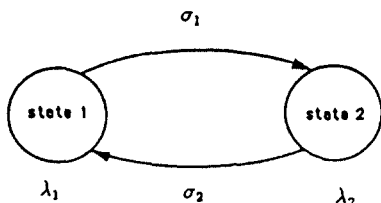


그림 4-1. 2-상태 MMPP(Markov Modulated Poisson Process)

Fig 4-1. two-state MMPP

시간우선순위제어에서의 높은 우선순위 트래픽은 음성과 같은 트래픽이고 낮은 우선순위 트래픽은 데이터와 같은 트래픽이 된다. 이 경우에 있어 음성 트래픽은 손실우선순위제어에서 MMPP로 모델링하는 것과는 달리 간단히 하기 위해서 포아송 프로세스로 가정했다.

2. 파라미터를 결정하는 asymptotic matching 방법

2-상태 MMPP는 Q와 Λ 로 나타낼 수가 있는데 이를 구성하는데는 4가지 파라미터가 있다. 이 값들은 실제 트래픽의 특성을 잘 반영하는 값이 되도록 파라미터 matching을 해야하는데 여기에는 몇가지 방법이 있다. Heffes[2]는 모멘트를 일치시키는 방법을 사용했는데 이 경우는 셀레벨의 특성은 잘 나타내지만 burst레벨의 특성은 잘 반영하지 못하는 단점이 있다[6]. 그래서 여기서는 Biocchi[6]가 사용한 근사적 정합맞춤방법을 사용하여 4가지 파라미터를 구하고자 한다. 최대 N개의 상호 독립적인 균질한 신호원 트래픽을 발생시키고 있다고 가정하고 각 신호원은 다음과 같이 세가지 파라미터로 나타낼 수 있다.

- F_p : 최대 비트율
- p: activity factor
- LF_B : 평균 버스트 길이

다음과 같은 표기를 정의하면 각 신호원을 그림4-2와 같은 ON-OFF 방법으로 나타낼 수 있다.

α : 출생율 ($= p \cdot F_p / (1-p) \cdot L_B$)

β : 사망율 ($= F_p / L_B$)

Λ : ON 상태에서 셀의 발생율 ($= F_p / L_I$)

L_I : ATM 셀의 유효부하의 길이

C: 출력용량

M: 최대 비트율로 할당할 수 있는 신호원의 최대 수

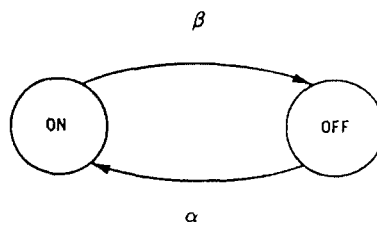


그림 4-2. 개폐프로세스 모델

Fig 4-2. Modeling of on-off process

이때 $J(t) = \{0, 1, \dots, N\}$ 을 시간 t 에서 활성상태에 있는 신호원의 수라고 하면 이것은 출생-사망 프로세스가 되며 그림4-3과 같이 나타낼 수 있다. $J(t) = j$ 일때 셀방사율은 $j\Lambda$ 가 된다. 여기서는 입력프로세서를 2-상태 MMPP로 모델링하는데 상태가 $J(t) = \{0, \dots, M\}$ 에 있는 저부한구간을 UL 상태, $J(t) = \{M+1, \dots, N\}$ 에 있는 과부한 구간을 OL 상태라 하여 다음과 같은 4가지 파라미터를 구하는 것이 목적이다.

$\sigma_{OL}(\sigma_{OL})$: OL(UL)상태를 벗어나는 평균 천이율.
 $\lambda_{OL}(\lambda_{OL})$: OL(UL)상태에서 포아송 프로세스의 평균 도착률.

먼저 OL state에서 살펴보면 M이 흡수상태가 된다. 전체 시스템의 state가 $[M, \dots, N]$ 으로 이루어진 과도마르코프 프로세스 Ω 를 정의하고 흡수가 일어날 때까지 걸리는 시간을 τ 라고 하면 τ 의 잔존함수는 다음과 같이 근사적으로 표현된다[12].

$$G_{\tau}(x) = De^{-ix} + O(e^{-ix}), X \rightarrow \infty \quad (4.1)$$

여기서 v 은 Q-행렬에서 공유치중에 실수부값이 가장 큰 값이다. 위식은 평균적으로 $1/v$ 후에는 OL 상태를 벗어나는 것을 나타내므로 $\sigma_{OL} = v$ 을 선택한다. $\lambda_{OL}(\lambda_{UL})$ 은 OL(UL)상태에서 평균 방사율이 되므로 다음 식으로 된다.

$$\lambda_{OL} = \Lambda \cdot \sum_{i=M+1}^N i \cdot \frac{\pi}{\pi_{OL}}$$

$$\lambda_{UL} = \Lambda \cdot \sum_{i=0}^M i \cdot \frac{\pi}{\pi_{UL}} \quad (4.2)$$

여기서 $\pi_{OL} = \sum_{i=M+1}^N \pi_i$, $\pi_{UL} = \sum_{i=0}^M \pi_i$ 이다. σ_{UL} 은 유효도착율의 관계식에 의해 다음 식과 같이 구할 수 있다.

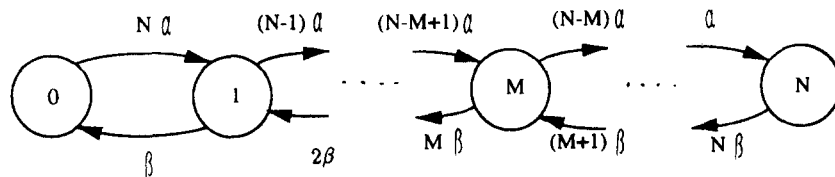


그림 4-3. Talker activity model
 Fig 4-3. Talker activity model

$$\sigma_{UL} = \sigma_{OL} \cdot \frac{N \cdot p \cdot \Lambda - \lambda_{UL}}{\lambda_{OL} - N \cdot p \cdot \Lambda} \quad (4.3)$$

이상으로 MMPP의 4가지 파라미터를 모두 구했다.

3. 성능분석

3.1 부분버퍼공유 방법

MMPP는 다음과 같이 Q와 Λ 로 표시할 수 있다.

1). $L < T_B$

$$Q = \begin{bmatrix} -\sigma_1 & \sigma_1 \\ \sigma_2 & -\sigma_2 \end{bmatrix}, \Lambda = \begin{bmatrix} \lambda_1 + \lambda_d & 0 \\ 0 & -\lambda_2 + \lambda_d \end{bmatrix} \quad (4.4)$$

2). $L \geq T_B$

$$Q = \begin{bmatrix} -\sigma_1 & \sigma_1 \\ \sigma_2 & -\sigma_2 \end{bmatrix}, \Lambda = \begin{bmatrix} \lambda_d & 0 \\ 0 & \lambda_d \end{bmatrix} \quad (4.5)$$

여기서

L : 현재 큐의 길이, T_B : 문턱값

$\lambda_1(\lambda_2)$: 2-상태 MMPP의 경우일때 상태1(2)에서 평균 도착률

$\sigma_1^{-1}(\sigma_2^{-1})$: 상태 1(2)에 머무는 평균시간

λ_d : 높은 우선순위트래픽의 평균 도착률

$N(t)$ 를 $(0, t)$ 사이에서 도착하는 셀의 수라고 하고 $J(t)$ 를 시간 t 에서 Markov process Q의 상태라고 할 때 다음 조건부확률을 정의한다.

$$P_{i,j}(n, t) = \text{prob.}(N(t) = n, J(t) = j \mid N(0) = 0, J(0) = i)$$

$$P(n, t) = [P_{i,j}(n, t)], \quad 1 \leq n, 1 \leq j \leq 2 \quad (4.6)$$

그리고 Z변환을 구하면

$$P(z, t) = \sum_{n=0}^{\infty} P(n, t) \cdot z^n, \quad |z| \leq 1$$

$$= e^{R(z) \cdot t}, \quad R(z) = Q + (z-1)\Lambda \quad (4.7)$$

이 된다. T를 슬롯타임이라할 때 0, T, 2T,...와 같은 이산시간에서 버퍼의상태를 살펴보자. 먼저 L_n 을 nT에서 버퍼에 들어있는 셀의 수라고 하고, J_n 을 nT에서 입력프로세스의 상태를 나타낸다고 하면 $\{(L_n, J_n), n \geq 0\}$ 은 상태공간이 $\{0, 1, 2, \dots, N\} \times \{1, 2\}$ 인 유한상태 Markov chain이 되며 정상상태 확률벡터 $\Pi = (\pi_0, \dots, \pi_N)$ 는 다음식에 의해 구할 수 있다.

$$\Pi \cdot P = \Pi$$

$$\Pi \cdot e = 1 \quad (4.8)$$

여기서 $\pi_k = (\pi_{k1}, \pi_{k2})$ 인 벡터이고 $\pi_{ki} = \lim_{n \rightarrow \infty} \text{prob.}(L_n = k, J_n = i)$ 이다. e는 모든 원소가 1인 $(2N+2) \times 1$ 벡터이어서 식(4.8)의 두번째 식은 확률벡터 Π 의 각 원소의 합이 1이 됨을 의미한다. 버퍼의 크기를 상태 공간으로 한 경우의 상태 천이도를 그리면 그림 4-4와 같다.

그리고 천이확률메트릭스 P는 다음과 같이 구해진다.

$$P = \begin{bmatrix} A_0 + A_1 & A_2 & \dots & A_{k+1} & B_{k+1,1} & \dots & B_{k+1,N-1-k} & \sum_{i=N-k}^{S-1} B_{k+1,i} \\ A_0 & A_1 & \dots & A_k & B_{k,1} & \dots & B_{k,N-1-k} & \sum_{i=N-k}^{S-k} B_{k,i} \\ 0 & A_0 & \dots & A_{k-1} & B_{k-1,1} & \dots & B_{k-1,N-1-k} & \sum_{i=N-k}^{S-k+1} B_{k-1,i} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & A_2 & B_{2,1} & \dots & B_{2,N-1-k} & \sum_{i=N-k}^{S-2} B_{2,i} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & C_1 & C_2 & \dots & C_{N-k} & \sum_{i=N-k+1}^S C_i \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 & \dots & C_0 & \sum_{i=1}^S C_i \end{bmatrix} \quad (4.9)$$

여기서 문턱값 T_B 는 k이고 A_k 는 현재 큐의 길이가 문턱값 값보다 작을때에 슬롯 타임동안에 k개의 셀이 도착할 확률이다. $B_{k,j}$ 는 셀이 도착하기전에는 threshold보다 큐의 길이가 작았다가 슬롯타임동안에 셀이 k개 도착하여 큐의 길이가 문턱값이 같아지면 이후에는 도착하는 셀들중에서 우선순위가 높은 셀만을 받아들인데 j개의 높은 우선순위 셀을 받아들일 확률이다. C_k 는 현재 큐의 길이가 문턱값보다 클 경우

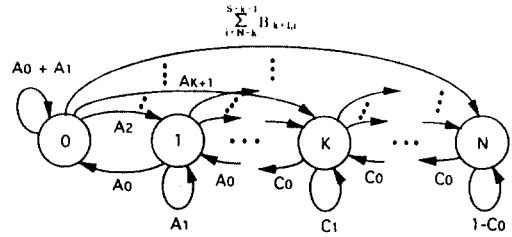


그림 4-4. 상태 천이도(threshold가 한개인 경우)
Fig 4-4. State transition diagram

에서 슬롯타임동안에 k개의 셀이 도착할 확률이고, $A_k, B_{k,j}, C_k$ 는 모두 2×2 매트릭스로 표시되고 다음과 같이 구해진다.

$$A_k = P_1(k, T), \quad k = 0, 1, 2, \dots, S, \quad \text{단, } L_n < T_B$$

$$B_{n_1, n_2} = \sum_{n=n_1+n_2}^S A_n \binom{n-n_1}{n_2} \left(\frac{\lambda_d}{\lambda_d + \lambda^*} \right)^{n_2} \cdot \left(\frac{\lambda^*}{\lambda_d + \lambda^*} \right)^{n-n_1-n_2}$$

$$C_k = P_u(k, T), \quad k = 0, 1, 2, \dots, S, \quad \text{단, } L_n \geq T_B \quad (4.10)$$

윗 식에서 $P_1(k, T), P_u(k, T)$ 는 식(4.6)에서 정의된 값으로 시간 T동안에 k개의 셀이 도착할 확률이며 여기에서 사용된 ATM 스위치는 $S \times S$ ATM 스위치라고 가정하였기 때문에 임의의 출력에 동시에 도착할 수 있는 셀의 최대 수는 S개로 제한된다($k \leq S$). $P_1(k, T), P_u(k, T)$ 는 식(4.7)로부터 역변환에 근거하여 계산되며[13], $P_1(k, T)$ 는 버퍼크기 L이 문턱값 보다 작은 경우로($L < T_B$) 식(4.4)의 Λ 를 사용하여 구할 수 있고, $P_u(k, T)$ 는 버퍼크기 L이 문턱값 보다 큰 경우로($L > T_B$) 식(4.5)의 Λ 를 사용하여 구할 수 있다.

우선순위가 높은 트래픽의 포기우선순위를 P_{B1} 이라고 하고 우선순위가 낮은 트래픽의 포기우선순위를 P_{B2} 라고 하면 다음과 같이 구해진다.

$$P_{B1} = \sum_{i=1}^2 \pi_{N,i} \quad (4.11)$$

$$P_{B2} = \sum_{i=1}^2 \sum_{j=1}^N \pi_{j,i}$$

3.2 HOL 우선순위 방법

시간 지연측면에서 높은 우선순위 트래픽은 실시

간 전송을 요구하는 음성이나 비디오같은 트래픽이며 호단위로 우선순위가 할당된다. 우선 여기서 가정하는 것은 호단위로 할당된 시간 우선순위를 ATM 스위치에서 인식할 수 있으며 우선순위에 따라 프로세싱도 가능하다는 것이다. 그리고 앞부분에서는 높은 우선순위 트래픽의 경우는 실제로 MMPP로 모델링되었는데 복잡하기 때문에 여기서는 근사적으로 MMPP에서의 effective arrival rate를 mean rate로 갖는 포사송 프로세스로 가정한다.

셀의 지연을 크게 세부분으로 나눌 수가 있다. 첫째는 도착했을때에 서버가 다른 셀을 서비스중에 있어 생기는 지연이고 둘째는 이미 도착해 있는 셀중에 우선순위가 자기와 같거나 높은 셀들에 의한 지연이다. 셋째는 나중에 도착했지만 우선순위가 자기보다 높아 먼저 서비스를 받는 셀들에 의한 지연이다. 여기서는 간단하게 하기위해서 MMPP로 모델링된 높은 우선순위의 트래픽을 effective arrival rate를 mean rate로 갖는 포사송 프로세스로 가정했다. 그러므로 참고문헌[7]을 이용하면 우선순위별 평균지연시간은 다음과 같다.

$$\begin{aligned}
 W_0 &= \lambda_H \cdot (1 - P_{B1}) / 2 + \lambda_L \cdot (1 - P_{B2}) / 2 \\
 W_H &= W_0 + \lambda_H \cdot (1 - P_{B1}) \cdot W_H \\
 W_L &= W_0 + \lambda_H \cdot (1 - P_{B1}) \cdot W_H + \lambda_L \cdot (1 - P_{B2}) \cdot W_L + \lambda_H (1 - P_{B1}) \cdot W_L
 \end{aligned}
 \tag{4.12}$$

여기서

W_0 : 남아있는 평균 서비스시간

λ_H : 높은 우선순위셀의 평균 도착률

λ_L : 낮은 우선순위셀의 평균 도착률

W_H : 높은 우선순위셀이 시스템에서 기다리는 평균시간

W_L : 낮은 우선순위셀이 시스템에서 기다리는 평균시간

P_{B1} : 높은 우선순위셀의 셀 손실률

P_{B2} : 낮은 우선순위셀의 셀 손실률

3.3 두개의 문턱값을 가진 손실우선순위 방법

문턱값이 한개인 경우에 대해서는 제III장 1절에서 모델을 설명하고, 이에 관한 분석은 제IV장 3.1에서 논의되었으며 이와 관련한 상태천이도는 그림 4-4에 있다. 문턱값이 두개인 경우의 모델에 대한 설명은 제III장 1절에서 하였고 성능분석을 하는 절차는 제IV장 3.1에서 논의된 것과 유사하기 때문에 여기서는 생략하겠다. 이와 관련하여 버퍼의 상태공간으로 하여 천이도를 그리면 그림4-5와 같다.

$\{(L_n, J_n), n \geq 0\}$ 은 $\{0, 1, 2, \dots, L_1, L_1 + 1, \dots, L_2 - 1, L_1 + 1, \dots, L_2 - 1, L_2, L_2 + 1, \dots, N\} \times \{1, 2\}$ 의 상태공간을 가지는 finite state Markov chain이 되며 정상상태 확률벡터 $\Pi = (\pi_0, \dots, \pi_N)$ 는 다음 식에 의해 구할 수 있다.

$$\begin{aligned}
 \Pi \cdot P &= \Pi \\
 \Pi \cdot e &= 1
 \end{aligned}
 \tag{4.13}$$

여기서 e 는 모든 원소가 1이고 크기가 $(2N + 2L_2 - 2L_1) \times 1$ 인 벡터이며 천이확률메트릭스 P 는 다음과 같이 표시된다.

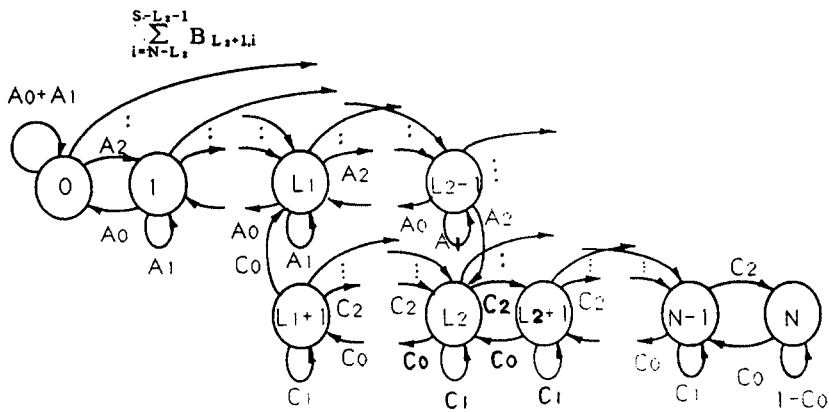


그림 4-5. 상태 천이도(threshold가 두개인 경우)
Fig 4-5. State transition diagram

P =

$$\begin{bmatrix} A_0 + A_1 & A_2 & \cdots & A_{L_1+1} & \cdots & A_{L_2} & 0 & \cdots & 0 & A_{L_2+1} \\ A_0 & A_1 & \cdots & A_{L_1} & \cdots & A_{L_2-1} & 0 & \cdots & 0 & A_{L_2} \\ & \vdots & & & & & & & & \vdots \\ 0 & 0 & \cdots & A_1 & \cdots & A_{L_2-L_1} & 0 & \cdots & 0 & A_{L_2-L_1+1} \\ & \vdots & & & & & & & & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & A_1 & 0 & \cdots & 0 & A_2 \\ 0 & 0 & \cdots & C_0 & \cdots & 0 & C_1 & \cdots & C_{L_2-L_1-1} & C_{L_2-L_1+1} \\ & \vdots & & & & & & & & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 0 & 0 & \cdots & C_0 & C_1 \\ & \vdots & & & & & & & & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix} B_{L_2+1,1} \cdots B_{L_2+1,N-L_2-1} & \sum_{i=N-L_2}^{S-L_2-1} B_{L_2+1,i} \\ B_{L_2,1} \cdots B_{L_2,N-L_2-1} & \sum_{i=N-L_2}^{S-L_2} B_{L_2,i} \\ & \vdots \\ B_{L_2-L_1+1,1} \cdots B_{L_2-L_1+1,N-L_2-1} & \sum_{i=N-L_2}^{S-L_2+L_1-1} B_{L_2-L_1+1,i} \\ & \vdots \\ B_{2,1} \cdots B_{2,N-L_2-1} & \sum_{i=N-L_2}^{S-2} B_{2,i} \\ C_{L_2-L_1+1} \cdots C_{N-L_1-1} & \sum_{i=N-L_1} C_i \\ & \vdots \\ C_2 \cdots C_{N-L_2} & \sum_{i=N-L_2+1} C_i \\ & \vdots \\ 0 \cdots C_0 & \sum_{i=1}^S C_i \end{bmatrix}$$

(4.14)

그리고 정상상태 확률벡터 $\Pi = (\pi_0, \dots, \pi_N)$ 와 버퍼 점유도 $\{0, 1, \dots, L_1, \dots, L_2 - 1\}$ 으로 이루어진 Markov chain Ω^* 과 $\{L_1 + 1, \dots, L_2, \dots, N\}$ 으로 이루어진 Markov chain \bar{Q} 에 의한 큐의 확률분포와의 관계는 다음과 같다.

$$\pi_i = \begin{cases} \pi^* i, & i \leq L_1 \\ \pi^* i + \bar{\pi}_i, & L_1 < i < L_2 \\ \bar{\pi}_i, & i \geq L_2 \end{cases} \quad (4.15)$$

그러므로 우선순위별 포기우선순위를 구해보면 다음과 같다.

$$P_{B1} = \sum_{i=1}^2 \bar{\pi}_{N,i}$$

$$P_{B2} = \sum_{i=1}^2 \sum_{L=1}^N \bar{\pi}_{N,i} \quad (4.16)$$

V. 수치계산결과

그림5-1은 손실우선순위제어의 경우, 3가지 트래픽 형태에 대해 문턱값 T_B 를 변화시킬때 각 트래픽에 대한 셀손실률을 구한 그림이다. T_B 의 감소에 따라 낮은 우선순위 트래픽의 셀손실률은 증가하고 높은 우선순위 트래픽의 셀손실률은 감소하는데 이는 예측된 결과이다. 그림5-2는 effective arrival rate λ^* 의 변화에 따른 셀손실률의 변화를 나타낸 것으로 λ^* 의 증가에 따라 셀손실률이 증가하며 3가지 모두 포사송 프로세스보다 큼을 알 수 있다. 그림5-3은 load ratio의 변화에 따른 허용부하량의 변화를 살펴본 것으로 load ratio의 증가에 따라 증가한다. 그림5-4는 asymptotic matching 방법에 의해 구한 파라미터를 근거로 λ_i 를 변화시킬때 셀손실률의 변화를 살펴본 것이다. 트래픽환경으로는 155.520Mbps중에서 DS-3급의 2채널에 비디오트래픽을 할당하고(89.472Mbps), 데이터트래픽에 2.048Mbps급의 6채널(12.288Mbps)을 할당했을때 음성인 경우는 64Kbps채널을 400개 정도 할당할 수 있는데 이런 경우에 최대 음성원의 수가 450개 정도 된다고 가정했을 경우이다. 그림5-5는 HOL 우선순위제어를 했을 경우에 평균지연시간을 구한 것으로 예상했던 것처럼 높은우선순위로 할당된 트래픽의 큐잉지연시간이 가장 적음을 알 수 있다. 그림5-6은 2개의 문턱값을 가진 손실우선순위 경우에 대한 그림으로 한개의 문턱값을 가진 경우보다 오히려 성능은 뒤떨어짐을 알 수 있다.

VI. 결론

이상에서 손실우선순위제어 방법으로 부분버퍼공유 방법을 사용하여 우선순위제어를 할 경우에 각 제어방법이 높은 우선순위트래픽과 낮은 우선순위트래픽의 셀손실률에 미치는 영향등에 대해 살펴보았다. 시간우선순위 방법으로는 HOL 우선순위 방법을 사용하여 우선순위제어를 함으로써 각 트래픽의 평균 큐잉지연시간에 미치는 영향에 대해 살펴보았다. 그리고 손실우선순위제어 방법의 경우에 문턱값 근처에서 큐의 길이가 불안정하게될 경우를 개선하기 위해 문턱값을 2개 두어 hysteresis 특성을 갖게 했을

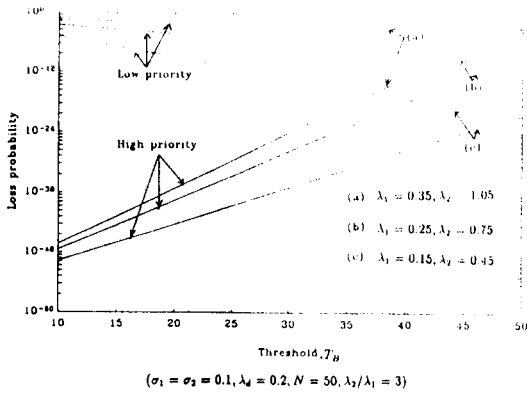


그림 5-1. T_B 의 변화에 따른 셀손실률의 변화
 Fig 5-1. Variation of cell loss probability as a function of T_B

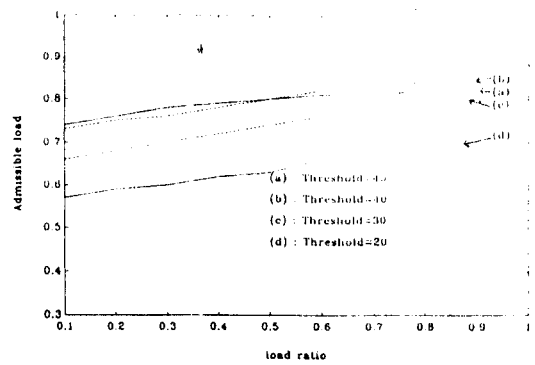
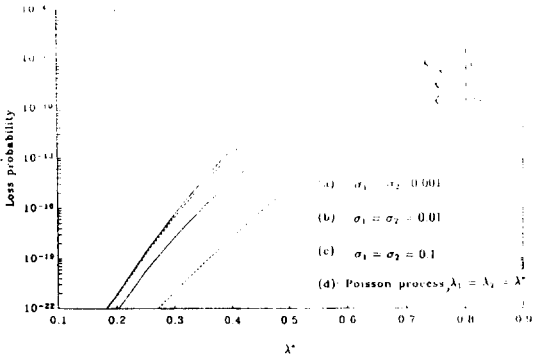
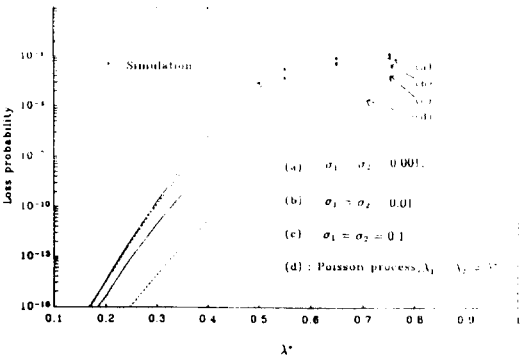


그림 5-3. 부하율($\lambda d / (\lambda d + \lambda^*)$)의 변화에 따른 허용부하량의 변화($N=50, \sigma_1 = \sigma_2 = 0.1$).
 Fig 5-3. Variation of admissible load as a function of load ratio

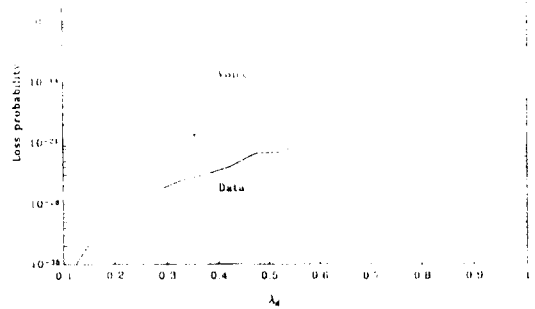


(a) 높은 우선순위트래픽
 (a) high priority traffic



(b) 낮은 우선순위트래픽
 (b) low priority traffic

그림 5-2. λ^* 의 변화에 따른 셀손실률의 변화($N=32, T_B=25, \lambda_1/\lambda_2=3, \lambda_d=0.2$)
 Fig 5-2. Variation of cell loss probability as a function of λ^*



($N=50, T_B=40, \lambda_{ul}=7.3 \times 10^{-2}, \lambda_{dl}=1.97 \times 10^{-1}, \sigma_{ul}=1.36 \times 10^{-18}, \sigma_{dl}=1.74 \times 10^{-3}$)

그림 5-4. 추출된 파라미터에 대해 λd 의 변화에 따른 셀손실률의 변화
 Fig 5-4. Variation of cell loss probability with extracted parameter as a function of λd

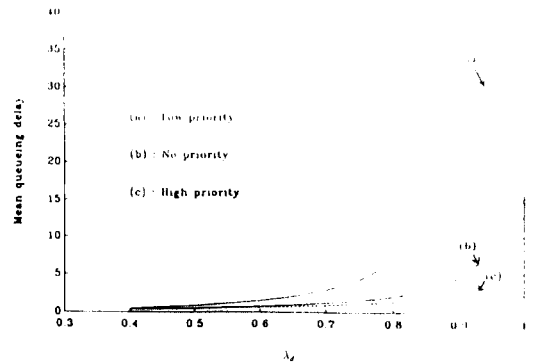


그림 5-5. HOL 우선순위방법을 사용했을때 λd 의 변화에 따른 셀의 평균지연시간의 변화($N=32, T_B=25, \sigma_1 = \sigma_2 = 1, \lambda_2 = 0.15, \lambda_1 = 0.05$)

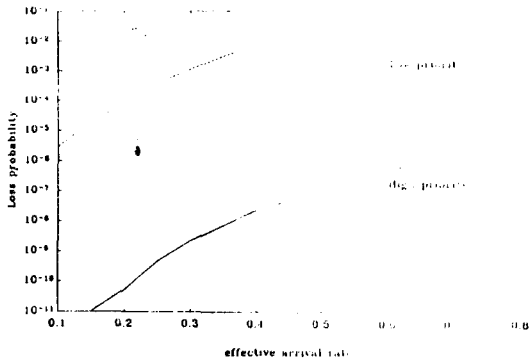


그림 5-6. 두개의 문턱값을 가진 우선순위제어시스템에서 effective arrival rate와 셀 손실률과의 관계($N = 32, L_1 = 23, L_2 = 27, \lambda_2/\lambda_1 = 3, \lambda_1 = 0.2, \sigma_1 = \sigma_2 = 0.1$)

Fig 5-6. Figure of loss priority control mechanism two threshold

경우에 우선순위제어방법의 성능도 살펴보았는데 성능은 문턱값이 한개 있을 경우에 비해 좋지 않음을 알 수 있었다. 문턱값을 두개 두어 제어를 수행하기 때문에 이 방법의 경우는 문턱값이 한개인 경우보다 구현은 복잡하지만 문턱값을 잘 선택하면 비슷한 성능을 가지게 할 수도 있고, 또 앞에서 언급한 것처럼 문턱값이 한개인 경우에는 문턱값 근처에서 큐의 길이가 불안정할 경우가 발생하지 않기 때문에 괜찮은 방법이라고 생각된다. 이 경우와 시간우선순위제어에서 입력 트래픽을 MMPP로 모델링하여 성능분석하는 것이 아직 과제로 남아있다.

참 고 문 헌

1. M. F. Neuts, Structured stochastic matrices of M/G/1 type and their applications, 1989.

2. H. Heffes, "A Markov modulated characterization of packetized voice and data traffic and related statistical multiplexer performance," IEEE JSAC, pp.856-868, Sep.1986.

3. K. Rothermel, "Priority mechanisms in ATM networks," IEEE Globecom '90, pp.847-851, 1990.

4. H. Kroner, "Priority management in ATM switching node," IEEE JSAC, pp.418-427, April 1991.

5. K. Sriram, "Traffic smoothing effects of bit dropping in a packet voice multiplexer," IEEE INFOCOM '88, pp.759-1988.

6. A. Baiocchi, "Loss performance analysis of an ATM multiplexer loaded with high-speed ON-OFF sources," IEEE JSAC, pp.388-393, April 1991.

7. L. Kleinrock, Queueing systems II, 1976.

8. 김영한, Performance of bandwidth allocation strategies in fast packet-switched BISDNs, KAIST 박사학위논문, 1990.

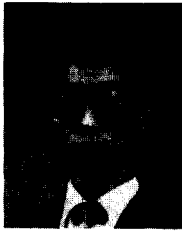
9. V. Ramaswami, "The N/G/1 queue and its detailed analysis," Adv/ Appl. Prob., pp. 222-261, 1980.

10. M.F. Neuts, "A versatile Markovian point process," J. Appl. Prob., 1980.

11. 장재신, 신병철, "우선순위기능을 가진 ATM스위치의 성능분석," JCCI, 4, 1992.

12. M.F. Neuts, Matrix-geometric solutions in stochastic models-an algorithmic approach, Johns Hopkins university press, 1981.

13. D.M. Lucantoni, "Efficient algorithm for solving the nonlinear matrix equations arising in phase type queues," commu. statist.-stochastic models, pp.29-51, 1985.



張 在 申(Jae Shin Jang) 정회원
 1965년 3월 25일생
 1990년 2월 : 동아대학교 전자공학과 졸업(공학사)
 1992년 2월 : 한국과학기술원 전기 및 전자공학과 졸업(공학석사)
 1992년 3월~현재 : 한국과학기술원 전기및 전자공학과 박사과정

※주관심분야 : BISDN, 이동통신분야



申 炳 喆(Byung Cheol Shin) 정회원
 1952년 11월 2일생
 1975년 2월 : 서울대학교 전기공학과 졸업
 1977년 2월 : 한국과학기술원 전기및 전자공학과 석사졸업
 1984년 8월 : 한국과학기술원 전기및 전자공학과 박사졸업
 1977년 2월~1980년 2월 : 한국전자통신연구소 연구원
 1987년 3월~1988년 2월 : SRI International에서 Post Doc. 연수
 1984년 9월~현재 : 한국과학기술원 전기및 전자공학과 부교수

※주관심분야 : Multimedia 통신, High Speed Network, Wireless Communication, ATM Network분야

朴 權 喆(Kwon Chul Park) 정회원
 1953년 10월 8일

1977년 : 고려대학교 전자공학과(학사)
 1979년 : 고려대학교 전자공학과(석사)
 1988년 : 고려대학교 통신공학과(박사)
 1982년 7월 1일~현재 : 한국전자통신연구소 책임연구원