

# 지연 서비스품질 요구사항을 고려한 효과적인 라운드 로빈 패킷 전송 스케줄링 기법

正會員 유 상 조\*, 박 수 열\*, 김 휘 용\*, 김 성 대\*

## An Effective Round-Robin Packet Transmit Scheduling Scheme Based on Quality of Service Delay Requirements

Sang Jo Yoo\*, Su Yeal Park\*, Hui Yong Kim\*, Seong-Dae Kim\* *Regular Members*

### 요 약

효과적인 패킷전송 스케줄링 알고리즘은 호 수락제어를 통하여 할당된 대역폭 및 기타 서비스 품질 요구사항등을 고려하여 각 연결에게 자원을 공정하게 배분하고 잘못된 사용자로부터 각 연결을 보호할 수 있어야 한다. 본 논문에서는 지연 서비스 품질 요구사항을 고려한 지연허용패킷 저장기법(DTPRS)을 이용하여 효과적인 라운드 로빈 방식의 패킷 전송 스케줄링을 제안한다. 제안된 방법은 각 연결에 대하여 공정한 대역폭 할당을 제공할 뿐만 아니라, 각 연결의 서로 다른 지연특성에 기반하여 지연이 허락되는 패킷의 출력링크 시간 슬롯을 저장하고 이를 지연에 급박한 패킷에 할당함으로써 지연, 지연변이 및 셀손실을 감소시킬 수 있다. 또한 본 논문에서 제안한 방법은 O(1)의 복잡도로 패킷을 처리하기 때문에 ATM과 같은 실제 통신망에 적용이 가능하다.

### ABSTRACT

An efficient packet transmit scheduling algorithm should be able to allocate the resource to each connection fairly based on the bandwidth and quality of service requirements negotiated during the call admission procedure and it should be able to isolate the effects of users that are behaving badly. In this paper, we propose an effective round-robin packet transmit scheduling mechanism, which we call the delay tolerant packet reserving scheme (DTPRS) based on delay QoS requirements. The proposed scheme can not only provide fairness and but also reduce delay, delay variation, and packet loss rate by reserving output link time slots of delay tolerant packets and assigning the reserved slots to delay urgent packets. Our scheme is applicable to high speed networks including ATM network because it only requires O(1) work to process a packet, and is simple enough to implement.

\*한국과학기술원 전기 및 전자공학과  
論文番號:97218-0627  
接受日字:1997年 6月 27日

## I. 서론

최근 멀티미디어 서비스에 대한 이용자의 요구가 높아지고 있으며 이와같은 서비스를 수용하기 위하여 공중 전기통신망도 음성, 비디오, 데이터와 같은 멀티미디어 트래픽을 고속으로 전송할 수 있는 광대역종합정보통신망(BISDN: Broadband Integrated Services Data Network)의 형태로 발전되고 있다. 광대역종합정보통신망은 ATM(Asynchronous Transfer Mode) 기술을 기반으로 하여 사용자가 원하는 정보를 실시간으로 서비스할 수 있으며, 이를 위하여 각각의 응용서비스 연결에 요구되는 서비스 품질(QoS: Quality of Service)을 보장하여야 한다.

현재 ITU-T(International Telecommunication Union)과 ATM Forum 등에서 고려되는 서비스 품질 파라미터는 지연(delay), 셀 손실율(cell loss rate), 지연변이(delay variation) 등이 있으며, 이 중에서 지연은 멀티미디어 실시간 서비스에 있어 정해진 최대범위(maximum bound)를 초과하는 경우 곧바로 셀 손실을 의미하게 되므로 매우 중요한 문제이다.

연결별로 할당된 서비스 품질 요구사항을 만족시키고 공정한 서비스를 제공하기 위해서는 ATM과 같은 고속 스위치 및 대용량 서버에서는 제한된 자원(예, 출력링크 용량)에 대해 입력 패킷(셀)들을 공정하게 스케줄링하여 전송하여야 한다. 효과적인 스케줄링 알고리즘은 호 수락 절차(call admission procedure)를 통하여 할당된 대역폭 및 기타 서비스 품질요구사항에 따라 각 연결에게 자원을 공정하게 배분하고, 할당된 범위를 벗어나 잘못 사용하는 이용자로 부터 각 연결을 보호할 수 있어야 한다.

현재까지의 패킷전송 스케줄링 연구에서, 출력 링크를 공유하는 모든 연결에 대하여 공정하게 대역폭을 할당하는 공정성(fairness) 측면에서는 이미 많은 연구가 이루어 졌으며, 유체 모델에 근거한 WFQ(Weighted Fair Queuing)등과 같이 이상적인 스케줄링에 근사한 방법[1][2][3][4][5]들이 제안되었다. 그러나 이들 방법들은 실제 ATM과 같은 고속의 통신망 환경에서는 계산의 복잡성으로 인해 실제 구현이 어려운 것으로 알려져 있다[6][7].

본 논문에서는 실제 고속통신망 환경에서 구현이 가능한 라운드 로빈 방식을 사용하면서 각 연결별 지

연특성을 고려하며 QoS 요구사항에 적합하게 지연과 지연변이를 제한 시키고 실시간 트래픽의 지연에 따른 패킷 손실을 제어할 수 있는 새로운 스케줄링 기법을 제안한다. 본 논문에서 제안한 지연허용패킷저장 기법(DTPRS: Delay Tolerant Packet Reserving Scheme)은 각 연결에 공정한 대역폭 할당을 제공할 뿐만 아니라, 각 연결의 서로 다른 지연특성에 기반하여 지연이 허락되는 패킷의 출력링크 시간 슬롯을 저장하고 이를 지연에 급박한 패킷에 할당함으로써 지연, 지연변이, 패킷 손실등을 감소 시킬 수 있다.

본 논문의 제 II장에서는 공정성 보장을 위한 기존의 패킷 스케줄링 기법에 대해 살펴보고 그 문제점을 지적한다. 제 III장에서는 제안된 지연허용패킷저장 기법(DTPRS)의 동작을 설명하고, 알고리즘을 제안한다. 제 IV장에서는 제안된 방법의 공정성을 해석하였다. 제 V장에서는 모의실험을 통하여 제안된 방법이 공정성 측면에서 완벽에 가까운 기존의 방법에 거의 근사한다는 결과를 보이고, 지연 품질 요구사항에 따른 연결별 패킷지연과, 패킷손실, 그리고 지연의 변이 등에 대해 우수한 성능을 갖는다는 것을 보인다. 제 VI장에서는 본 논문의 결론을 맺고 향후 연구과제에 대해 기술한다.

## II. 관련 연구

지금까지의 패킷교환망에서는 'best-effort' 위주의 서비스를 지원하였으나 앞으로는 멀티미디어 서비스와 같은 다양한 종류의 품질요구사항을 갖는 트래픽을 지원하여야 한다. 이를 위하여 각 연결별로 일정한 수준의 성능을 보장하여야 하고 통신망에서 단일 출력 링크에 대해 여러 개의 연결 트래픽이 입력될 때 어떻게 입력 패킷들을 출력할 것인가를 결정하기 위해 효과적인 스케줄링 기법이 사용되어야 한다. 본 장에서는 현재까지 연구된 주요 스케줄링 기법들을 간략히 소개하고 장단점을 비교분석할 것이다.

WFQ(Weighted Fair Queuing)[1][2], WF<sup>2</sup>Q(Worst-case Fair Weighted Fair Queuing)[5]은 데이터 트래픽을 무한히 쪼갤 수 있는 유체 모델에서 정의되는 이상적인 스케줄링 알고리즘인 GPS(Generalized Processor Sharing)를 패킷환경에 근사화 하여 각 연결에 공정한 대역폭을 할당한다. WF<sup>2</sup>Q는 서비스 완료 시간 뿐

만 아니라 서비스 시작시간을 이용함으로써 Fair Queuing 환경을 패킷환경에 가장 잘 근사화 시킨다. 그러나 이 방법들은 거의 완벽한 공정성을 제공하지만 각 패킷 도착 시 가상시간(virtual time)의 계산이 복잡하고, 출력링크를 통해 패킷을 전송할 때 다음에 보내야 할 패킷을 결정하기 위하여 각 연결별로 큐에 쌓여있는 첨두패킷의 가상시간을 정렬하여 가장 작은 값을 갖는 패킷을 전송하여야 하므로 실시간 정렬의 문제가 있다. 이러한 계산의 복잡성과 실시간 정렬 문제는 연결의 수가 많을 경우 ATM과 같은 고속의 통신망 환경에서는 구현이 불가능한 것으로 알려져 있다[6][7]. SCFQ(Self Clocked Fair Queuing)[4]는 내부적으로 생성하는 가상시간을 활용함으로써 가상시간 계산에 소요되는 작업 복잡도를 줄였다. 시분할 다중화 시스템에 기초한 가상시간을 이용하는 방법으로는 Virtual Clock 방법[3]이 있으나 각 연결들 사이에 완전한 공정성을 보장하지 못하는 단점이 있다. 이상과 같은 방법들은 기본적으로 실시간 정렬의 문제로 실제 환경에서의 사용이 어렵다[6].

위의 가상시간 방식을 이용하는 방법들이 N개의 연결에 대해  $O(\log N)$ 의 복잡도를 갖는데 비해 거의 같은 공정성을 보장하면서도 패킷처리 작업의 복잡도를 활성화된 연결의 수에 관계없이  $O(1)$ 으로 유지할 수 있는 방법인 DRR(Deficit Round Robin) 스케줄링 알고리즘[7]이 제안되었다. DRR은 그림 1의 예에서와 같이 각 연결들에 대해 할당된 대역폭에 해당하는 quantum( $Q_i$ )을 배정하고, 매 라운드에서 이에 해당하는 패킷양 만큼을 서비스할 수 있다. 패킷길이가 고정되지 않은 패킷 서비스에서 만약 해당 라운드에서 전송하고 남아있는 quantum 양이 다음에 전송할 패킷길이 보다 작아 해당 패킷을 처리할 수 없을 경우에는, 다음 라운드에서 새로이 할당될 quantum에 남아있는 deficit counter  $D_i$ 을 더한 패킷양 만큼의 서비스를 제공할 수 있다. 그러나 DRR은 지연이나 패킷 손실 측면의 서비스 품질 요구사항을 고려하지 않았고, 패킷 지연도 라운드의 길이에 비례하여 길어지는 단점이 있다.

본 논문에서 제안한 지연 서비스 품질요구사항을 고려한 효과적인 라운드 로빈 패킷 스케줄링 기법은 WFQ등의 Fair Queueing의 공정성에 근사하면서  $O(1)$ 의 복잡도로 구현이 가능하며, DRR등에서 보인

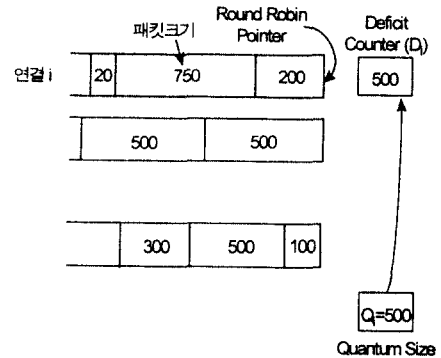


그림 1-(a). 라운드로빈의 시작점에서  $D_i$ 는 0으로 초기화되고, 연결  $i$ 의 서비스 시점에  $Q_i = 500$  만큼이  $D_i$  값에 할당된다.

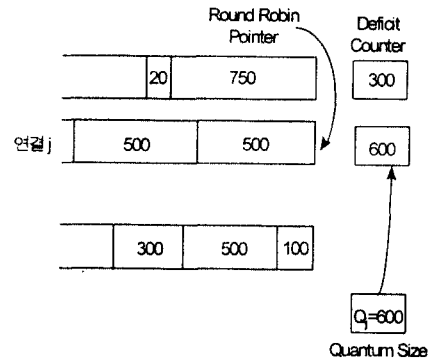


그림 1-(b). 연결  $i$ 의 두번째 패킷크기가 남아 있는  $D_i = 300$  보다 크므로 다음 연결  $j$ 로 포인터를 이동시킨다. 다음 라운드에서 연결  $i$ 의  $D_i$  값은 남아있는 300에  $Q_i = 500$ 을 더해 800이 된다.

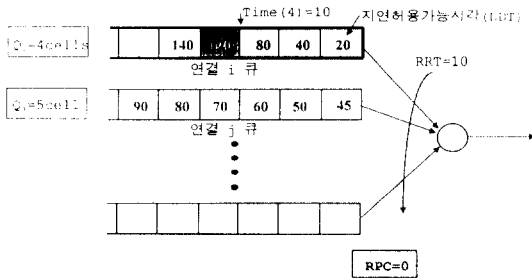
지연관련 문제점을 보완하여 실시간 멀티미디어 서비스에 적용가능한 구조를 갖는다.

### III. 지연 서비스 품질을 고려한 지연허용 패킷저장기법(DTPRS) 스케줄링

본 장에서는 패킷기반의 통신망에서 단일 출력 링크에 대해 여러 개의 연결이 입력될 때 연결 별로 할당된 대역폭에 해당하는 트래픽을 공정하게 서비스하고 동시에 지연 관련 서비스 품질을 최대한 만족시키기 위한 새로운 스케줄링 기법을 제안한다.

3.1 DTPRS 스케줄링의 연구 배경 및 기본 구조

제 II장에서 설명한 바와 같이 DRR을 이용한 패킷 스케줄링은 호 설정 단계에서 연결별로 사용자와 망이 호 수락 제어를 통하여 현상된 할당 대역폭에 근거하여 미리 한번의 라운드에 전송 가능한 패킷양(quantum)을 계산하여, 매 라운드 마다 연결별로 quantum 범위 내에서 큐에 대기중인 패킷을 출력링크를 통하여 전송하게 된다. 이러한 방법은 대역폭을 공정하게 할당할 수 있지만 약속된 지연품질에 따른 효과적인 스케줄링을 제공하지는 않는다. 본 논문에서 제안하는 스케줄링 기법은 DRR에 기초한 라운드 로빈을 사용하면서, 각 연결별 지연허용 특성에 근거하여 동적으로 출력링크 시간슬롯을 각 연결에 할당하게 된다. 다음 그림에서는 패킷의 크기가 고정된 ATM 셀의 경우를 설명하고 패킷 크기가 고정되지 않은 경우는 3.2절의 알고리즘에서 일반화 한다.



DDT : Deadline Departure Time, RRT : Round Robin Time, RPC : Reserved Packet Counter

그림 2. 라운드 로빈을 이용한 DTPRS 구조(연결 i 서비스 전)

그림 2에서와 같이 제안된 지연허용패킷저장기법(DTPRS)은 기본적으로 DRR과 같은 라운드 로빈 방식을 사용한다. 각 연결별로 할당된 Quantum 크기(Qi) 만큼 매 라운드 마다 패킷을 전송할 수 있다. 그림 2의 예에서는 연결 i는 4개의 셀을 전송할 수 있으며, 연결 j는 5개의 셀을 한꺼번에 매 라운드에서 전송할 수 있다. 제안된 방식에서는 DRR과는 달리 각 셀(패킷)은 해당 연결의 큐에 도착한 때 지연 서비스 품질 요구사항에 근거하여 해당 큐에서 최대로 지연시킬 수 있는 지연 허용시간 스텝트를 갖고 입력된다. 그림에서 큐의 각 셀에 표시된 숫자는 큐에서의 지연 허용가능 시각을 의미한다. 즉 해당시각을 초과하여 계속해서 큐에 있게 되면 지연품질을 망에서 지

키지 못하게 되며 실시간 트래픽의 경우 셀(패킷) 손실로 이어지게 된다. 그림 2에서 볼 수 있듯이 연결 i의 네번째 셀은 연결 i의 quantum이 4개 셀이기 때문에 이번 라운드에서 송신될 수 있지만 다음 번 라운드에 서비스 되어도 지연 허용치를 초과하지 않기 때문에, 즉 셀 4 전송직전 현재시간(Time(4)=10)의 기준에서 1회 라운드 로빈 시간(RRT=100)을 더한 110보다 큰 지연 허용 시각(120)을 갖기 때문에 금번 라운드에서 서비스 하지 않고 다음 번 라운드로 송신을 연기시키며 연결 i의 4번째 셀을 위해 사용될 예정이었던 출력링크 시간슬롯은 저장된다.

그림 3은 연결 i의 서비스를 끝내고 연결j를 서비스하기 위한 상태에서의 큐 스케줄링 구조를 보여준다. 연결 i의 송신 시 4번째 셀 슬롯이 사용되지 않았기 때문에 저장패킷카운터(RPC: Reserved Packet Counter) 값이 연결 i의 4번째 셀 크기 만큼 증가 되었다. 연결 j는 이번 라운드에서 5개 셀을 전송할 수 있는데 6번째 셀의 경우 이번 라운드에서 전송되지 않고 다음 라운드에서 전송될 때, 다음 라운드 서비스 시작 시각은 140이고 연결 j의 6번째 셀의 지연허용시각은 90 이므로 지연허용치를 초과하게 되며 실시간 트래픽의 경우 셀 손실로 이어진다. 이때 RPC의 값에 1 셀 크기만큼의 출력링크 시간슬롯이 저장되어 있으므로 뒤의 다른 연결의 셀에 지연을 미치지 않고도 이신에 저장된 패킷슬롯을 사용하여 연결 j의 6번째 셀을 전송할 수 있다. 셀 전송 후의 RPC는 다시 0이 된다.

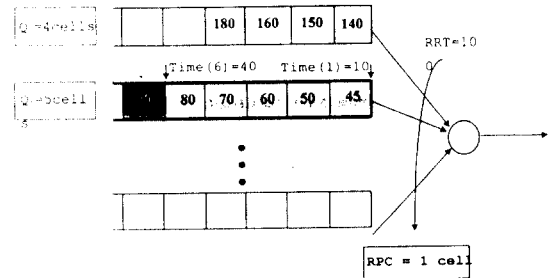


그림 3. 연결 j 서비스 시의 큐 스케줄링 구조

따라서 그림 3에서 볼 수 있듯이 지연요구사항에 비해 여유가 있는 패킷들을 허용 가능한 범위에서 지연시키고, 지연에 급박한 패킷들을 먼저 서비스 해

주게 된다. 제안된 방법을 사용하면 DRR을 이용하여 대역폭 할당의 공정성을 보장하면서 지연서비스 품질 요구사항까지 고려한 효율적인 스케줄링을 수행할 수 있다. 또한 제안된 방법은 지연을 요구된 범위 안에 제한 시키기 위해 노력함으로써 전체적으로 지연의 변이(delay variation) 역시 제한할 수 있다.

본 논문에서 제안한 방법은 기본적으로 전송되어야 할 패킷이 큐에 있는 한 서버가 휴지(idle)일수 없는 작업보존 서비스원칙(work conserving service discipline)으로 동작하기 때문에 그림 2의 예에서, 강제 지연된 패킷을 가진 연결에 뒤이은 연결들의 큐에 전송될 패킷이 대기하고 있지 않은 경우, 강제 지연된 패킷은 실제로는 지연없이 다음 라운드에서 서비스 된다.

제안된 방법에 대한 알고리즘을 설명하기 전에 몇 가지 변수에 대한 정의와 수식을 제시한다.

**정의**

**DDT : Deadline Departure Time**

연결  $i$ 의 새로운 패킷  $k$ 가 도착했을 때 해당 패킷이 연결  $i$ 의 큐에 있다가 서비스 받아야 할 최후 시간.

$$DDT_k = \text{Packet Arriving Time}(PAT_k) + \text{Maximum Delay in Queue } i(MDT_i)$$

**MDT : Maximum Delay Time in Queue**

연결  $i$ 의 지연 서비스품질 요구사항에 따라 해당 스위치 노드  $l$ 의 연결  $i$  큐에서 지연 가능한 최대값

$$MDT_i = \text{Delay QoS value for Connection } i \times \text{weighing value of node } l$$

**RPC : Reserved Packet Counter**

지연에 민감한 연결에 사용하기 위해 저장된 총 패킷 크기

**RPCmax : Maximum Reserved Packet Counter**

RPC가 가질 수 있는 최대값

**RRT : Round Robin Time**

$m$ 번째 라운드의 연결  $i$  큐 서비스 개시 시 다음 라운드  $m+1$ 에서 연결  $i$ 의 큐가 다시 서비스 받을 수 있

을 때까지의 라운드 로빈 지연시간.

**NRRT : Next Round Robin Time**

$m$ 번째 라운드의 연결  $i$ 의 패킷  $k$  전송 시점에서 이번 라운드에 전송하지 않고 다음  $m+1$ 번째 라운드에서 전송할 경우 라운드 로빈 지연시간.

명제 1 출력 링크 속도를  $C$ ,  $m$ 번째 라운드 로빈 시작 시점에서 각 연결별 잔여 패킷양(deficit counter)을  $D_{i(m)}$ 이라 하고  $m-1$ 라운드까지 저장된 RPC 값을  $RPC_{(m)}$ 이라 할때, 한 라운드의 길이는 다음과 같이 제한된다.

$$RRT_{(m)} \leq \frac{\sum_i Q_i + RPC_{(m)} + \sum_i D_{i(m)}}{C} \tag{1}$$

증명: RRT는 해당 라운드의 각 연결별로 할당된 패킷을 전송하는데 소요되는 시간을 의미하므로, 각 연결별로 해당 라운드에서 보낼 수 있는 최대 패킷양인  $\sum Q_i$ 을 모두 전송하고, 이전 라운드에서 저장된  $RPC_{(m)} + \sum D_{i(m)}$ 을 모두  $m$ 번째 라운드에서 사용할 때 최대가 된다.  $m$ 번째 라운드에서 RPC 값을 저장하거나  $D_i$ 를 남길 경우에는 RRT는 명제 1의 오른쪽 항 보다 작게 된다. ■

명제 2  $S_{i(m)}$ 을  $m$ 번째 라운드에서 연결  $i$ 의 quantum  $Q_i$  중 이미 전송된 패킷양이라 할때  $NRRT_{(m)}$ 의 길이는 다음과 같이 제한된다.

$$NRRT_{(m)} \leq \frac{\sum_i Q_i + RPC_{(m)} + \sum_i D_{i(m)} - S_{i(m)}}{C} \tag{2}$$

증명: NRRT 값은 현재 연결  $i$ 의 패킷  $k$  서비스 시점에서 다음 연결로 이동 시 해당 연결  $i$ 의 패킷  $k$  전송시점으로 되돌아 오는데 걸리는 시간이므로, 해당 라운드에서의 최대 라운드 로빈 시간  $RRT_{(m)}$ 에  $S_{i(m)}$ 을 전송하는 데 소요된 시간을 뺀 값이  $NRRT_{(m)}$ 의 최대값이 된다. 이때  $S_{i(m)}$ 은 연결  $i$ 가  $m$ 번째 라운드에서 서비스 받은 양이다. ■

3.2 DTPRS 스케줄링 알고리즘

3.1 장에서 설명한 제안된 스케줄링 구조와 정의를 사용하여 본 장에서는 패킷의 수신과 송신 시 수행되는 절차를 단계별로 설명한다. 제안된 알고리즘은 고정길이 패킷을 갖는 ATM 환경에서도 똑같이 적용될 수 있는 일반화된 형태를 가진다.

**Packet arriving part**

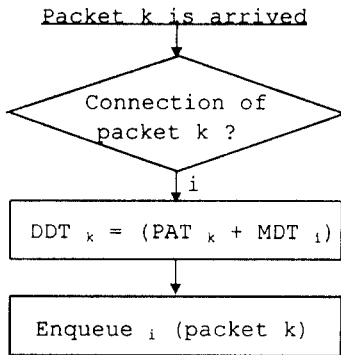


그림 4. 패킷 도착 시 절차

**Packet departure part**

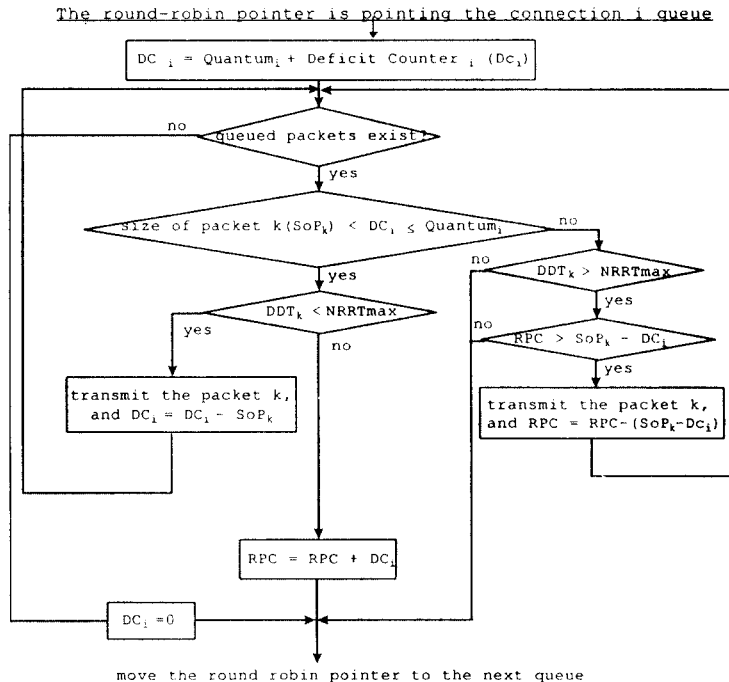


그림 5. 패킷 서비스 시 DTPRS 절차

DTPRS 스케줄링 알고리즘에서 NRRTmax는 해당 라운드에서 계산된 NRRT의 최대값이다.

그림 4와 그림 5에서 볼 수 있듯이 제안된 DTPRS 스케줄링 기법은 DRR을 기반으로 하여 비교적 간단한 동작으로 각 연결의 지연특성을 고려한 효율적인 패킷 전송 스케줄링을 수행한다. 그림 4의 “패킷 수신부”는 하나의 패킷 도착 시 패킷의 연결(ATM의 경우 VPI/VCI)을 식별하여 해당 큐에 넣는다. 이때 도착 패킷은 지연 서비스품질 요구사항을 만족시키기 위하여 큐에서 서비스 되어야 할 최종시점인 DDT를 가지게 된다. 그림 5의 “패킷 송신부”에서는 큐에서 서비스할 패킷이 존재할 경우 다음에 서비스될 패킷의 DDT 값이 다음 라운드 로빈 시간보다 작을 경우 이를 전송한다. 만약 다음 라운드에 전송하여도 지연 품질값을 만족시킬 수 있을 경우에는 해당 연결에서 전송할 수 있는 패킷양 중 전송하지 않은 패킷양 만큼을 RPC에 추가한다. 또한 임의의 연결 큐에서 다음 라운드 시간보다 작은 DDT를 갖고 있는 패킷의 패킷양이 서비스 받을 수 있는 할당된 패킷양 보다

많아지게 되면, 해당 연결 큐는 RPC에 저장된 패킷 양의 범위내에서 추가적으로 서비스를 받는다.

실제 구현에 있어 RPC 값을 제한 시키는 것이 필요한데 그 이유는 RPC 값을 크게 갖도록 허용할 경우 짧은 시간구간에서 공정성이 보장되지 않을 수 있고, 각 라운드의 최대 라운드 로빈 시간이 길어져서 지연이 커질 수 있기 때문이다. 제 V장의 모의실험에서는  $RPC_{max}$  값을  $\sum Q_i$ 와  $\sum Q_i/2$ 로 제한하여 그 결과를 보였다.

### VI. 공정성(Fairness) 해석

본 논문에서 제안한 지연허용패킷저장기법(DTPRS)은 기본적으로 할당된 대역폭을 공정하게 서비스하는 DRR(Deficit Round Robin)에 기반을 두고 있으므로 공정성을 일정한 이하로 제한시킬 수 있다. 패킷 전송 스케줄링에서 공정성 척도(FM: Fairness Measure)는 임의의 시간구간( $t1, t2$ ) 시간구간 동안 전송된 연결  $i$ 의 패킷양이고,  $f_i$ 는 할당된 대역폭이다.

$$FM(t1, t2) = \left| \frac{Sent_i(t1, t2)}{f_i} - \frac{Sent_j(t1, t2)}{f_j} \right| \quad (3)$$

공정성 척도 값이 작다는 것은 할당된 대역폭에 맞게 출력링크 자원을 모든 연결에 대해 공정하게 서비스 하였다는 것을 의미한다. 위의 공정성 비교 대상의 연결들에 대해 임의의 구간에서 큐가 비어있지 않다는 것은 DRR의 경우 임의의 구간 라운드에서 처리할 수 있는 할당된 quantum 이상의 패킷이 해당 구간에 도착한 경우 (해당 구간에서 할당된 대역폭을 초과하는 트래픽이 입력된 경우)의 큐상태를 의미한다.

제안된 방법과 같이 지연 QoS 요구사항을 고려하여 강제적 지연등 수행하는 알고리즘에 대해 기존의 공정성 비교 개념을 그대로 사용할 경우 지연에 여유가 있어 강제적 지연을 시키는 경우, 입력 트래픽이 할당된 것보다 적게 들어오는 시간구간에서도 큐에 패킷이 전송되지 않고 대기 상태를 유지할 수 있으므로 트래픽이 할당된 것 이상으로 들어오는 연결의 큐 대기 상태와의 공정성 비교는 적절치 않다. 따라서 다음 정의를 사용하여 기존의 DRR 과 같은 방법에서의 공정성 비교대상 상태와 같은 트래픽 입력 상태에서 공정성을 비교한다.

정의 DTPRS에서 지연을 고려한 RPC의 저장 및 사용없이, 순수하게 매 라운드에서 할당된 quantum을 전송하는 상황에서 해당 연결의 큐에 계속 패킷이 대기 상태에 있는 구간을 해당연결의 연속작업구간 이라 한다. 이때 연속작업구간에서의 전체 입력 트래픽은 할당된 대역폭을 초과하게 된다.

공정성 분석은 임의의 시간 구간에서 같은 연속작업구간 ( $t1, t2$ )를 갖는 모든 연결  $i, j$ 에 대하여 식 (3)과 같이 전송된 패킷총량을 각 연결의 전송률로 나눈 값의 차이가 최대값으로 어떤 값에 제한되는 가를 분석한다.

먼저 같은 연속작업구간을 포함하는 두연결  $i, j$ 에 대해, 연속작업구간에서 연결  $i$ 가 지연 서비스 품질을 고려한 강제적인 지연을 통해 할당된 전송율보다 최소로 패킷을 전송하는 경우와 다른 연결  $j$ 가 저장된 RPC까지 모두 사용하여 전송 가능한 최대값을 갖는 경우에 대해 공정성 척도인  $\left| \frac{sent_i(t1, t2)}{f_i} - \frac{sent_j(t1, t2)}{f_j} \right|$  값을 구하면 최악의 공정성을 갖을 때의 공정성 비교를 수행할 수 있다.

제안된 DTPRS 방법을 사용할 경우의 공정성이 최악이 되는 상황에서의 공정성 해석을 위하여 연속작업구간에서 할당된 연결의 전송률에 비해 가장 적게 전송되는 상황을 살펴보면 그림 6과 같다.

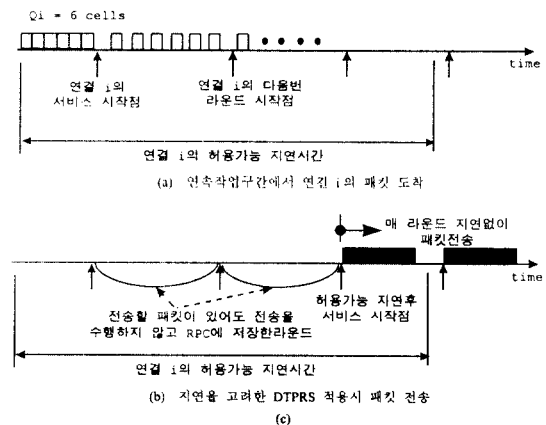


그림 6. 연결 i의 연속작업구간에서 공정성비교를 위한 연결 i의 전송량이 가장 적은 상태

그림 6에서 알수 있듯이 연속작업구간에서 타연결의 전송량에 비해 연결 i의 전송량의 차이가 가장 크게 되는 구간은 연결 i의 큐에 전송 되어야 할 패킷이 있더라도 최대 지연허용 까지 패킷을 전송하지 않고 다른 지연에 급박한 연결의 패킷전송에 사용하기 위해 패킷을 저장한 라운드의 합이된다. 최대 지연허용 시간을 지나면 패킷을 더이상 지연시킬 수 없기 때문에 연결 i 큐의 연속작업구간이 끝날 때까지 패킷의 저장 없이 계속 할당된 만큼 전송하게 된다. 전송을 개시한 라운드 이후부터 다른 연결과의 공정성 차이는 계속 작아지게 된다.

연결 i의 패킷을 할당된 것보다 적게 보내 공정성이 최악이 되는 구간의 길이는 다음과 같이 구해진다.

$$Round_i(|t1-t2|) = \left\lfloor \frac{MDT_i}{RRT} \right\rfloor \quad (4)$$

모든 연결에 대해 패킷을 할당된 것보다 적게 보내 공정성이 최악이 되는 구간인 식 (4)의  $Round_i(|t1-t2|)$  구간이 가장 긴 연결은 최대허용지연시간이 다른 연결에 비해 최대값을 갖는 연결로, 이때  $Round_{max}(|t1-t2|)$  는 다음과 같다.

$$Round_{max}(|t1-t2|) = \left\lfloor \frac{MDT_{max}}{RRT} \right\rfloor, \quad (5)$$

$$MDT_{max} = \text{Max}(MDT_i, \forall i)$$

명제 3 식 (5)에 의해 결정된 최악의 전송을 갖는 연결이 i이고 연결 j의 최대 패킷 크기가  $L_{jmax}$ 라 할때, 식 (5)에 의해 결정된 (t1, t2) 구간에서 전송될 수 있는 패킷양은 다음식과 같다.

$$\begin{aligned} Sent_i(t1, t2) &\geq \text{Max} \left\{ 0, (Q_i - RPC_{max}) \times \left\lfloor \frac{MDT_i}{RRT} \right\rfloor - D_i(t2) \right\} \\ &\geq \text{Max} \left\{ 0, (Q_i - RPC_{max}) \times \left\lfloor \frac{MDT_i}{RRT} \right\rfloor - L_{i,max} \right\} \end{aligned} \quad (6)$$

증명: MDT가 가장 긴 연결 i에 대해 식 (5)로 표현되는 구간에서 연결 j의 전송량의 최소값은 다음의 경우에 의해 결정된다.

- ① 할당된 quantum  $Q_i$ 가  $RPC_{max}$ 보다 작아 매 라운드에서 할당된 quantum을  $RPC$ 로 저장

할 때에는  $Round_i(|t1-t2|)$  구간에서 전송된 연결 j의 최소 패킷 전송량은 0이된다.

- ② 할당된  $Q_i$ 가  $RPC_{max}$ 보다 큰경우 그 차이인  $(Q_i - RPC_{max})$ 는 매 라운드 전송되어야 하며, 여기에 t2 시점에서 남을 수 있는 전송된 값에서 마지막 t2 시점에 남은  $D_i$ (deficit counter) 값을 뺀 값이 최소 전송량이 된다.  $D_i \leq L_{i,max}$  이므로 식 (5)의 구간에서 연결 j 전송량의 최소값은 식 (6)과 같이 표현된다. ■

반면에 식 (5)의 구간에서 임의의 연결 j의 전송량이 최대가 되는 경우는 연결 j가 다른 연결이 저장한  $RPC$ 를 매 라운드에서 추가로 사용하는 경우이다. 이때 연결 j에게 할당된 양 이외에 추가로 사용할 수 있는 패킷양은 식 (7)과 같다.

$$\begin{aligned} Sent'_j(t1, t2) &\leq \text{Min} \left\{ \sum_{i,j} \left\lfloor \frac{MDT_{max}}{RRT} \right\rfloor \times Q_i, \right. \\ &\quad \left. RPC_{max} \times \left\lfloor \frac{MDT_{max}}{RRT} \right\rfloor \right\} \end{aligned} \quad (7)$$

명제 4  $L_{jmax}$ 를 연결 j의 최대 패킷크기라 할때, 연결 j가 식 (5)에 의해 결정된 (t1, t2) 구간에서 최대한으로 전송할 수 있는 패킷양은 다음과 같다.

$$\begin{aligned} Sent_j(t1, t2) &\leq \left\lfloor \frac{MDT_{max}}{RRT} \right\rfloor \times Q_j + D_j(t1) + RPC(t1) \\ &\quad + \text{Min} \left\{ \sum_{i,j} \left\lfloor \frac{MDT_{max}}{RRT} \right\rfloor \times Q_i, RPC_{max} \times \left\lfloor \frac{MDT_{max}}{RRT} \right\rfloor \right\} \\ &\leq \left\lfloor \frac{MDT_{max}}{RRT} \right\rfloor \times Q_j + L_{j,max} + RPC_{max} \\ &\quad + \text{Min} \left\{ \sum_{i,j} \left\lfloor \frac{MDT_{max}}{RRT} \right\rfloor \times Q_i, RPC_{max} \times \left\lfloor \frac{MDT_{max}}{RRT} \right\rfloor \right\} \end{aligned} \quad (8)$$

증명: 연결 j가 (t1, t2) 구간에서 최대한으로 전송할 수 있는 패킷양은 다음항들의 합에 의해 결정된다.

- ① 식 (7)에 의해 다른 연결들이 저장한  $RPC$  값을 모두 사용하는  $Sent'_j(t1, t2)$



- ② (t1, t2) 구간의 라운드에서 연결 j에 할당된 모든 패킷을 전송
- ③ t1 시간 이전에 저장되었던  $D_j(t1)$ ,  $RPC(t1)$  를 모두 사용할 경우이다.  
 $D_j$  값은 최대로  $L_{jmax}$ 로 제한되고,  $RPC$ 는  $RPCmax$ 로 제한 되므로 식 (8)을 얻는다. ■

정리 1 연속작업구간 (t1, t2)를 갖는 모든 연결에 대해 최대지연 허용시간으로 인해 최악의 전송률을 갖게되는 연결을 i라 할때, 임의의 연결 j에 대하여 다음의 관계가 항상 성립한다.

$$\left| \frac{Sent_i(t1, t2)}{f} - \frac{Sent_j(t1, t2)}{f} \right| \leq \left\{ \begin{array}{l} \frac{\left[ \frac{MDT_{max}}{RRT} \times Q_i + L_{jmax} + RPC_{max} \right]}{f} + \\ \text{Min} \left\{ \sum_{(n,j)} \left[ \frac{MDT_{max}}{RRT} \right] \times Q_i, RPC_{max} \times \left[ \frac{MDT_{max}}{RRT} \right] \right\} \\ \text{Max} \left\{ 0, (Q_i - RPC_{max}) \times \left[ \frac{MDT_{max}}{RRT} \right] - L_{imax} \right\} \end{array} \right\} \quad (9)$$

증명: 명제 3, 4에 의해 참이다. ■

이상과 같은 공정성 해석을 통하여, 제안된 방법은 그림 6에서 알 수 있듯이 지연 허용치 만큼 강제적 지연을 시킴으로써 전체 연속작업구간에서 공정성은 일정하게 보장되지 않고 구간별로 변화가 있음을 알 수 있다. 식 (5)를 통하여 공정성이 최악이 되는 구간을 도출하였고, 이 구간에서의 공정성 척도는 정리 1에 의해 제한됨을 알 수 있다. 정리 1의 공정성 척도는 연결의 최대지연허용시간이 길어질수록 MDTmax에 비례하여 공정성이 나빠짐을 알 수 있는데, 이는 지연허용시간 구간에서는 지연이 허락되는 연결을 QoS 보장 한도 내에 지연시키고, 다른 지연에 민감한 연결을 위해 추가적으로 시간 슬롯을 사용하였기 때문에 당연한 결과이다.

본 논문에서 제안한 DTPRS는 지연이 허용되는 한

해당 연결의 패킷을 전송하지 않음으로써 RPC 값을 증가시키고 다른 지연에 급박한 패킷을 가진 연결에 RPC 를 사용하게 허락함으로써 지연에 민감하지 않은 연결이 전체 트래픽 전송시간에 비해 아주 짧은 구간인 최대 지연가능 구간(수 msec) 동안에는 공정성을 보장받을 수 없을 수 있다. 그러나 제안된 방법은 사용자가 망과의 QoS 협상을 통하여 망에서의 최대지연허용시간을 허락하였으므로 이 구간에서의 해당 연결의 공정성 제공은 중요하지 않다. 반면에 이 구간에서 저장된 출력 링크 시간슬롯을 지연에 급박한 패킷에 할당함으로써 보다 효율적인 자원의 사용을 통해 전체적인 스케줄링 성능을 향상시킬 수 있다.

#### IV. 모의 실험 결과

본 장에서는 컴퓨터 모의실험을 통하여 본 논문에서 제안한 DTPRS 패킷 스케줄링 방법에 대한 공정성 분석 및 지연, 지연변이, 손실 등에 대한 성능평가를 수행한다. 컴퓨터 모의실험의 구조는 그림 7과 같이 구성된다. 트래픽 발생기는 원하는 수만큼의 연결에 대한 정의된 트래픽을 발생시키며, DRR 과 제안된 DTPRS 스케줄링 알고리즘이 구현된 스케줄러단에서는 연결별 QoS를 고려한 스케줄링을 수행한다. 마지막으로 성능분석기에서는 스케줄러를 통과한 출력 트래픽을 연결별로 저장하여 스케줄러의 성능을 분석 평가한다.

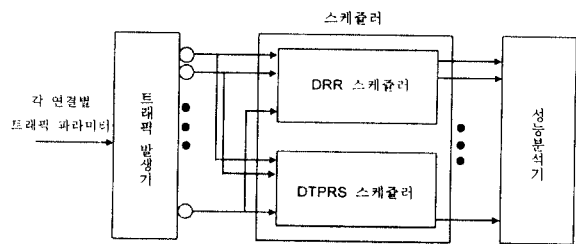


그림 7. 모의실험 구성도

##### 5.1 공정성(fairness 분석)

공정성 분석을 위하여 다음 표 1과 같은 트래픽을 연결별로 발생시켜 적용하였다. 모의실험 환경은 53 byte의 고정된 패킷크기를 갖는 ATM 환경에서 수행

하였다. 모두 10개의 연결이 슬롯당 1개의 ATM 셀을 전송하는 출력단을 공유하게 된다. 표 1의 두번째 열은 호 수락 절차 후 할당된 연결별 대역폭으로 출력단의 대역폭을 1로 하여 정규화된 값을 갖는다. 세번째 열은 실제 입력된 정규화된 트래픽양이고 마지막 열은 QoS 협상을 통하여 할당된 해당 노드에서의 각 연결별 최대 지연가능 값(슬롯)이다. 악조건에서의 공정성 분석을 위하여 지연에 민감하지 않은 연결 5(지연값 500 슬롯)의 다음에 지연에 민감한 잘못된(할당된 것보다 많은 양의 트래픽을 발생) 연결 6을 배치하여 전체적 공정성에 얼마만큼 영향을 미치는가 검토하였다.

표 1. 공정성 분석에 사용된 연결별 트래픽 및 지연값

| 연결 | 할당된 대역폭 | 실제 입력된 대역폭 | 지연 QoS 값(슬롯) |
|----|---------|------------|--------------|
| 1  | 0.05    | 0.25       | 50           |
| 2  | 0.15    | 0.25       | 40           |
| 3  | 0.05    | 0.25       | 20           |
| 4  | 0.15    | 0.25       | 30           |
| 5  | 0.05    | 0.25       | 500          |
| 6  | 0.15    | 1.00       | 50           |
| 7  | 0.05    | 0.25       | 40           |
| 8  | 0.15    | 0.25       | 70           |
| 9  | 0.05    | 0.25       | 15           |
| 10 | 0.15    | 0.25       | 20           |

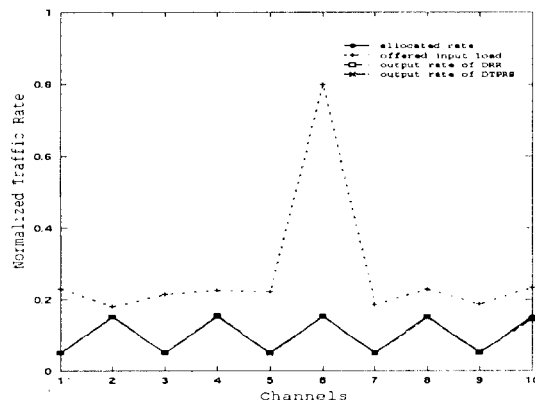


그림 8. 공정성 분석

각 연결의 입력단으로 들어오는 트래픽은 ON/OFF 모델을 따른다고 가정하였다. 트래픽 모델의 ON 기간과 OFF 기간은 지수분포(exponential distribution)를 가지며, 셀 전송율이  $\lambda$ 인 (출력링크 대역폭을 1로 생각하여 정규화한 값) 경우 지수분포 발생을 위한 평균 ON 기간은  $100\lambda$  슬롯이고, 평균 OFF 기간은  $100(1-\lambda)$  슬롯이다. 모의실험 시간은 10,000 슬롯시간 동안 수행하였고,  $RPC_{max} = \sum_i Q_i$ 로 제한하였다.

그림 8의 결과에서 알 수 있듯이 DRR 과 제안된 DTPRS 모두 원래 할당된 대역폭과 거의 일치하게 공정한 서비스를 하는 것을 알 수 있다. 이는 제안된 DTPRS가 공정성 측면에서 실제 ATM 환경에서 문제없이 사용될 수 있음을 의미한다.

### 5.2 기타 지연관련 QoS에 대한 평가

제안된 DTPRS은 공정성 분석을 통하여 연결별로 할당된 대역폭을 공정하게 서비스 하는 것을 알 수 있었다. 효과적인 셀 스케줄링을 제공하기 위해서는 공정성 뿐만 아니라 요구된 지연 QoS 값에 대한 평균 지연, 지연의 변이 (delay variation), 지연으로 인한 셀 손실, 최대 지연값 등이 사용자가 요구한 범위내에서 서비스 되어야 한다. 본 장에서는 이러한 지연과 관련된 파라미터에 대해 제안된 방법과 DRR과의 모의실험을 통한 성능평가를 수행하였다.

#### (1) 평균지연 및 지연으로 인한 셀손실 분석

표 2 및 표 3은 할당된 대역폭으로 입력되는 여러 연결에 대한 평균지연 및 지연으로 인한 셀 손실 비교를 나타낸다. 정규화 값은 QoS에 병기된 지연허용값으로 평균 지연 값을 나눈 것이다. 이는 각 연결별 패킷 지연이 요구된 값에 비해 어느 정도 비율로 만족시키는가를 나타낸다.

위 결과에서 알 수 있듯이 제안된 방법이 평균지연 측면에서 호 설정 단계에서 요구된 지연값에 더 근접하는 평균지연을 나타내었다. 그 이유는 지연에 비교적 덜 민감한 연결 5의 셀들을 지연 허용범위 안에서 지연시키고 이를 다른 지연에 민감한 셀의 전송에 사용하였기 때문이다. 또한 지연으로 인한 셀 손실도 제안된 방법은 DRR 보다 많은 수가 감소함을 알 수 있다.  $RPC_{max}$ 의 비교에서는  $RPC_{max} = \sum_i Q_i/2$ 로 제한

표 2. 평균 지연

| 연결 | 할당/입력<br>대역폭 | Q <sub>i</sub><br>(셀수) | 지연<br>QoS 값<br>(슬롯) | 평균지연<br>(DRR) |      | 평균지연(DTPRS)<br>( $RPC_{max} = \sum_i Q_i$ ) |      | 평균지연(DTPRS)<br>( $RPC_{max} = \sum_i Q_i / 2$ ) |      |
|----|--------------|------------------------|---------------------|---------------|------|---|------|---|------|
|    |              |                        |                     | 지연            | 정규화  | 지연  | 정규화  | 지연  | 정규화  |
| 1  | 0.10         | 5                      | 50                  | 42.73         | 0.85 | 29.50                                       | 0.59 | 23.33   | 0.46 |
| 2  | 0.18         | 9                      | 40                  | 32.87         | 0.82 | 21.72                                       | 0.54 | 22.07   | 0.55 |
| 3  | 0.16         | 8                      | 20                  | 30.80         | 1.54 | 23.17                                       | 1.15 | 23.57   | 1.17 |
| 4  | 0.14         | 7                      | 30                  | 43.97         | 1.46 | 27.76                                       | 0.92 | 28.23   | 0.94 |
| 5  | 0.12         | 6                      | 500                 | 33.29         | 0.06 | 220.22                                      | 0.44 | 202.94  | 0.40 |
| 6  | 0.10         | 5                      | 50                  | 38.02         | 0.76 | 18.40                                       | 0.36 | 17.95   | 0.35 |
| 7  | 0.08         | 4                      | 40                  | 66.01         | 1.65 | 25.38                                       | 0.63 | 24.11   | 0.60 |
| 8  | 0.06         | 3                      | 70                  | 76.86         | 1.09 | 53.89                                       | 0.76 | 55.09   | 0.78 |
| 9  | 0.04         | 2                      | 15                  | 26.48         | 1.76 | 18.67                                       | 1.24 | 16.29   | 1.08 |
| 10 | 0.02         | 1                      | 20                  | 59.49         | 2.97 | 26.94                                       | 1.34 | 26.06   | 1.30 |

표 3. 지연으로 인한 셀 손실

| 연결 | 셀손실<br>(DRR) | 셀손실 (DTPRS)<br>( $RPC_{max} = \sum_i Q_i$ ) | 셀손실 (DTPRS)<br>( $RPC_{max} = \sum_i Q_i / 2$ ) |
|----|--------------|---|---|
| 1  | 283          | 191   | 121   |
| 2  | 413          | 232   | 232   |
| 3  | 576          | 521   | 537   |
| 4  | 549          | 369   | 363   |
| 5  | 0            | 94  | 67  |
| 6  | 271          | 67  | 67  |
| 7  | 305          | 97  | 136   |
| 8  | 213          | 111   | 114   |
| 9  | 168          | 143   | 122   |
| 10 | 143          | 93  | 74  |

한 것이  $\sum_i Q_i$ 로 제한한 것보다 약간 우수하게 나타났다.

(2)지연 변이(delay variation) 및 최대 지연값 비교

현재 국제 표준화 기구에서 논의되고 있는 또 하나의 QoS 성능평가 지표는 지연의 변이 (delay variation)의 정도로 이는 이용자 응용서비스에도 중요한 성능

변수가 될뿐아니라 망 내부에서의 트래픽 변형을 유발하는 요소 중의 하나이기도 하므로 이를 억제할 필요가 있다.

스케줄링 단에서의 지연변이는 망 내부에서 여러 교환기를 거쳐감에 따라 점점 입력 트래픽이 버스티해지는 현상을 초래한다. 이러한 현상은 정확한 망 내부에서의 트래픽 모델링을 통한 호 수락 제어나 혼잡제어를 어렵게 한다[6]. 본 논문에서 제안한 방법은 지연값을 지연 QoS에 맞추기 위해 동적인 스케줄링 기법을 사용함으로써 지연변이 값역시 전체적으로 DRR에 비해 작은 지연변이 값으로 제한된다.

표 4의 결과를 살펴보면 제안된 방법이 DRR보다 지연 표준편차를 적은 값으로 제한시키고 있으며 이는 사용자 응용서비스에 안정적 운용과 망내부에서의 혼잡등을 억제할 수 있다. DTPRS에서  $RPC_{max}$ 를  $\sum_i Q_i / 2$ 로 제한한 것이  $\sum_i Q_i$  보다 약간 더 우수한 성능을 보였다.

최대 지연값 역시 제안된 방법을 사용할 경우 DRR에 비해 작은 값으로 제한되고 있음을 알 수 있으며, 단 연결 5의 경우 지연 QoS 값이 500 슬롯임을 생각할때 DRR에 비해 증가한 최대 지연값은 수용할 수 있는 값이라 생각된다.

표 4. 지연변이(지연 표준편차) 비교

| 연결 | 지연값<br>(슬롯) | 지연표준편차<br>(DRR) | 지연표준편차 (DTPRS)<br>( $RPC_{max} = \sum_i Q_i$ ) | 지연표준편차 (DTPRS)<br>( $RPC_{max} = \sum_i Q_i / 2$ ) |
|----|-------------|-----------------|--|--|
| 1  | 50          | 45.92           | 39.18  | 33.78  |
| 2  | 40          | 33.55           | 30.26  | 27.33  |
| 3  | 20          | 39.06           | 29.16  | 28.19  |
| 4  | 30          | 48.13           | 37.97  | 36.96  |
| 5  | 500         | 39.97           | 189.55   | 184.74   |
| 6  | 50          | 39.02           | 18.35  | 18.76  |
| 7  | 40          | 64.37           | 27.36  | 24.78  |
| 8  | 70          | 105.75          | 77.90  | 76.08  |
| 9  | 15          | 29.16           | 23.26  | 19.50  |
| 10 | 20          | 82.02           | 40.88  | 42.48  |

표 5. 최대 지연값 비교

| 연결 | 최대 지연값 (DTRR)<br>(슬롯) | 최대 지연값 (DTPRS)<br>( $RPC_{max} = \sum_i Q_i$ ) | 최대 지연값 (DTPRS)<br>( $RPC_{max} = \sum_i Q_i / 2$ ) |
|----|-----------------------|--|--|
| 1  | 212                   | 228  | 193  |
| 2  | 158                   | 140  | 148  |
| 3  | 196                   | 144  | 131  |
| 4  | 211                   | 189  | 189  |
| 5  | 197                   | 626  | 611  |
| 6  | 158                   | 82   | 113  |
| 7  | 250                   | 162  | 111  |
| 8  | 440                   | 315  | 320  |
| 9  | 152                   | 131  | 112  |
| 10 | 404                   | 221  | 225  |

(3) 입력 load의 변화에 따른 평균지연 및 셀 손실 비교

입력 트래픽의 load를 출력링크 대역폭 대비 0.4에서 1.0 까지 증가 시키면서 모든 연결의 평균 지연 및 셀 손실 수의 변화를 실험하였다.

입력 트래픽의 load를 증가시키에 따라 제안된 방법이 DRR 방법에 비해 우수함을 알 수 있다. 특히 지연으로 인한 셀손실의 경우 평균 500-1000개의 셀이 털 손실 되는 것을 볼 수 있다. 또한 두가지 RPCmax에 대해 거의 차이가 없이 같은 성능을 보임을 알 수

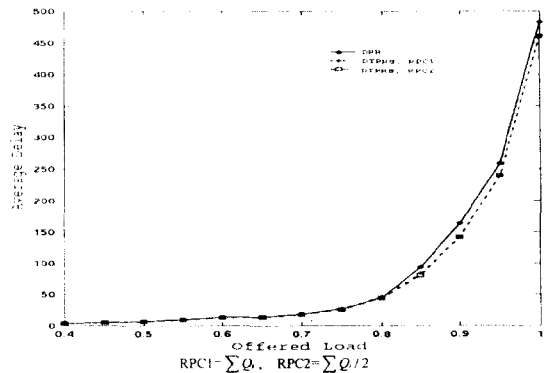


그림 9. 입력 load 변화에 따른 평균지연 비교

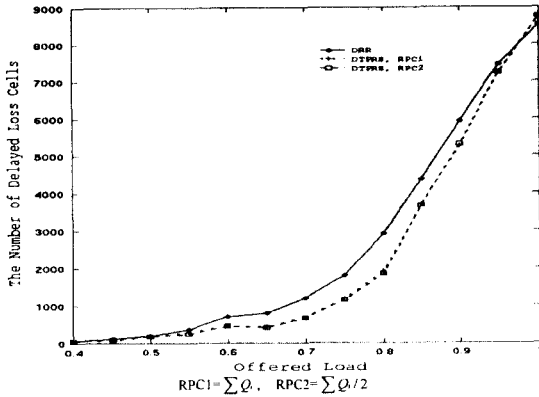


그림 10. 입력 load 변화에 따른 지연으로 인한 셀 손실 비교

있다.

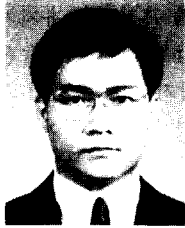
## VI. 결 론

본 논문에서는 실제 고속통신망 환경에서 구현이 가능한 라운드 로빈 방식을 사용하면서 지연 QoS 요구사항에 따라 지연특성을 고려하며 실시간 트래픽 지연에 따른 셀손실을 제어할 수 있는 새로운 스케줄링 기법을 제안하였다. 본 논문에서 제안하는 방법은 해당 연결의 서비스 시점에서 지연 QoS를 고려하여 지연이 가능한 패킷(ATM의 경우 셀)들을 해당 라운드에서는 전송하지 않고, 이 패킷들을 서비스 할 예정이었던 출력 링크 슬롯을 다른 연결의 지연허용 값에 임박하여 해당 라운드에 꼭 전송되어야 할 다른 연결의 패킷에게 할당하게 된다. 제안된 지연허용패킷저장기법(DTPRS: Delay Tolerant Packet Reserving Scheme)은 지연에 덜 민감한 연결의 지연을 할당된 QoS 값 이내로 유지하면서, 지연에 민감한 다른 연결의 패킷 지연을 줄일 수 있으며 지연의 변이도 기존의 방법에 비해 억제된다. 모의실험을 통하여 제안된 방식이 DRR에 비해 지연 및 지연변이, 셀손실 등에 좋은 성능을 보이는 것을 알 수 있었다. 이러한 결과를 통해 제안된 방식이 구현의 용이성을 제공하고 또한 QoS 측면에서도 우수한 성능을 갖고 실제 ATM 통신망과 같은 멀티미디어 고속통신망에 사용될 수 있다는 결론을 얻을 수 있다. 본 논문에서 제안한 방식은 스케줄링 결과로 정확한 QoS를 항상 보장하지는 못하지만 연결별로 QoS 값을 고려하여 실제 고속

망 환경에서 구현이 가능한 O(1)의 복잡도를 유지기 때문에 정확한 QoS 보장없이도 효과적으로 성능을 향상시킬 수 있는 멀티미디어 단말기 등에 응용될 수 있을 것이다. 향후 고속통신망에서 구현이 용이하면서도 여러 서비스 품질 요구사항을 만족시킬 수 있는 공정한 스케줄링 기법에 관한 연구가 계속될 예정이다.

## 참 고 문 헌

1. A. Demer, S. Keshav, and S. Shenker, "Analysis and simulation of a fair queuing algorithm", ACM SIGCOM '89, pp. 3-12.
2. Abhay K. Parekh, and Robert G. Gallager, "A generalized processor sharing approach to flow control in integrated services networks: the single-node case", IEEE/ACM Trans. on Networking, Vol. 1, No. 3, pp. 344-357, 1993.
3. L. Zhang, "Virtual clock: A new traffic control algorithm for packet switching networks", ACM Transactions on Computer Systems, Vol. 9, No. 2, pp. 101-124, 1991.
4. S. Jamaloddin Golestani, "A self-clocked fair queuing scheme for broadband applications", IEEE INFOCOM '94, pp. 636-646.
5. Jon C.R. Bennett, and Hui Zhang, "WF<sup>2</sup>Q: Worst-case fair weighted fair queuing", IEEE INFOCOM '96, pp. 120-128.
6. Hui Zhang, "Service disciplines for guaranteed performance services in packet-switching networks", Proceeding of IEEE, Vol. 83, No. 10, pp. 1374-1396, 1995.
7. M. Shreedhar, and George Varghese, "Efficient fair queuing using deficit round-robin", IEEE/ACM Trans. on Networking, Vol. 4, No. 3, pp. 375-385, 1996.



유 상 조(Sang Jo Yoo) 정회원  
1984년 3월~1988년 2월: 한양대  
학교 전자통신공학  
과 졸업(공학사)  
1988년 3월~1990년 2월: 한국과  
학기술원 전기 및  
전자공학과 졸업(공  
학석사)

1990년 3월~현재: 한국통신 연구개발본부  
1997년 3월~현재: 한국과학기술원 전기 및 전자공학  
과 박사과정  
※주관심분야: 멀티미디어 통신, 영상압축 및 전송,  
무선망 QoS



박 수 열(Soo Yeal Park) 정회원  
1983년 3월~1987년 2월: 고려대  
학교 전자공학과 졸  
업(공학사)  
1988년 3월~1990년 2월: 한국과  
학기술원 전기 및  
전자공학과 졸업(공  
학석사)

1990년 3월~1994년 9 월: 한국통신 통신방연구소  
1994년 10월~1995 년 8월: Hankuk Microtec Research  
1995년 9월~현재: 한국과학기술원 전기 및 전자공학  
과 박사과정  
※주관심분야: 멀티미디어 통신, 영상압축 및 전송,  
Networking, RealTime OS



김 휘 용(Hui Yong Kim) 정회원  
1972년 5월 8일생  
1990년 3월~1994년 8월: 한국과  
학기술원 전기 및  
전자공학과 졸업(공  
학사)  
1996년 3월~현재: 한국과학기술  
원 전기 및 전자공  
학과 석사과정

※주관심분야: QoS 보장 통신



김 성 대(Seong-Dae Kim) 정회원  
1953년 12월 26일생  
1977년: 서울대학교 전자공학과  
졸업(공학사)  
1979년: 한국과학기술원 전기 및  
전자공학과 졸업(공학석  
사)  
1983년: 프랑스 INPT ENSEEIHT  
졸업(공학박사)

1984년~현재: 한국과학기술원 전기 및 전자공학과  
교수  
※주관심분야: 영상처리, 영상통신, 컴퓨터 비전, VLSI  
구현등