

버스트 큐잉 구조에서 실시간 트래픽 QoS에 관한 연구

°김광준*, 장혁*, 나상동*, 배철수**
★ 조선대학교 컴퓨터 공학과
★★ 관동대학교 정보기술 공학과 °

A Study on the QoS of Real-Time Traffic in Bursty Queuing Scheme

Gwang-Joon Kim*, Hyok-Jang*, Sang-Dong Na*, Chul-Soo Bae**
★ Dept. of Computer Eng. Chosun Univ.
★★ Division of Information Comm. Eng. Kwandong Univ.

ABSTRACT

본 논문은 실시간성 서비스인 버스트 트래픽을 갖는 ATM(Asynchronous Transfer Mode)망에서 실시간 연결 수락 제어를 위한 종단간의 지연보장을 알고리즘을 이용하여 실시간 서비스 트래픽 서비스 품질의 성능에 대해 분석하였다. 셀 손실 확률은 QoS(Quality of Service)의 기준이 되므로 제안된 실시간 알고리즘은 일정한 버퍼 메모리 크기와 연결 요청에 대한 수락 여부 즉, 연결 수락 여부의 결정을 위한 하나의 큐잉 서비스를 필요로 한다. 이러한 트래픽 소스 각 타입에 대한 초과된 셀에 대한 셀 손실 확률은 트래픽 소스 타입 전체에 하나만 존재하더라도 요구를 충족시킬 수 없으므로, 본 논문에서 제안된 버스트 트래픽 특성에 따라 지연보장 알고리즘을 이용한 큐 서비스 방식은 기존의 종단간 지연 보장 알고리즘들에서 사용하던 셀 정렬을 사용하지 않고 몇 개의 FIFO 큐로 구성된 라운드 로빈 큐를 이용하여 복잡성을 크게 감소시켰으며, 실시간 서비스와 비실시간 서비스의 서비스방식을 효과적으로 결합하여 다양한 QoS를 보장해야 하는 ATM망에 효과적으로 적용할 수 있고 제안된 알고리즘에 산출되는 각각의 셀 손실 확률에 충분한 상한선을 제공하는 전형적인 매개변수를 사용하여 성능을 분석한다.

I. 서 론

각종 서비스들을 하나의 망으로 통합하려는 시도와는 별도로, 고속 패킷 전송과 멀티미디어 통신 서비스에 대한 수요는 꾸준히 증가해 왔다. 이러한 고속 패킷 전송 및 멀티미디어 통신 서비스는 근본적으로 고속 및 광대역의 통신 채널을 요구하였다. 64 kbps급의 ISDN이 이러한 서비스를 수용하는 것은 불가능하므로 자연히 광대역 서비스를 제공하기 위한 B-ISDN(Broadband Integrated Service Digital Network)의 필요성이 대두되게 되었고 ITU-T는 광대역 서비스들이 갖는 다양한 특징들을 통합 수용하기 위한 새로운 통신 방식으로 ATM(Asynchronous Transfer Mode)을 그 표준으로 채택하게 되었다[2].

ATM은 다양한 종류의 서비스를 동시에 처리할 수 있도록 기존의 회선 교환 방식과 패킷 교환 방식의 장점을 결합하여 셀이라는 일정한 크기의 패킷을 비동기식 시분할 다중화 방식에 의해 전송하는 방식이다. ATM망에서는 이 선적인 트래픽에 대해 개개의 서비스 품질(QoS:Quality of Service)을 보장할 수 있는 트래픽 제어가 핵심 기술 중의 하나이다. 트래픽 제어에는 연결 수락제어, 체중 제어, 흐름 제어, 우선 순위 제어 등이 포함된다. 큐 서비스 방식은 연결 수락 제어 과정에서 결정된 사용자 매개 변수를 이용하여 대역폭과 버퍼 공간을 이용하여 사용자가 요구하는 QoS를 만족시킬 수 있도록 여러 연결로부터 들어 온 셀들의 우선 순위를 결정하고 그 우선 순위에 따라서 전송하는 기법이다. 실시간 멀티미디어 통신 서비스가 B-ISDN의 주된 서비스로 부각됨에 따라 전송 지연이나 지연 변이 등을 보장하기 위해 큐 서비스 방식의 중요성이 커지게 되었다 [4,5,6].

기존 패킷망에서의 큐 서비스 방식은 QoS를 보장하기 위해 사용하는 처리 방법에 따라 정렬 우선 순위 큐 방식과 프레임 방식으로 분류할 수 있다. 정렬 우선 순위 방식은 우선 순위를 보장하기 위하여 큐에 저장된 패킷을 정렬하는 방식이며, 프레임 방식은 우선 순위가 다른 여러 단계의

프레임을 사용하여 QoS를 보장해 주는 방식이다.

본 논문에서는 프레임 방식의 단순성을 이용하여 큐잉 구조에서 지연 보장 알고리즘 이용하여 우선 순위 정렬 방식의 단점을 보완한 새로운 큐 서비스 방식을 제안한다. 제안하는 큐 서비스 방식은 동시성과 셀 교환 방식에 지연 EDD(Earliest Due Date)방식에서와 같은 종단간 마감시간을 이용하여 다양한 지연 한계를 요구하는 연결들을 수락할 수 있으면서 구현을 단순화시킬 수 있을 것이다.

II. 종단간 지연보장을 위한 트래픽 변수 특성

A. 실시간 서비스 품질 특성

멀티미디어 응용을 지원할 때 중요시되는, 네트워크의 성능 기준으로는 처리율, 전송지연, 지연분산, 에러율 등이 있다. 처리율은 비트 율이라고도 부르는데, 두 종단간 시스템이 서로 통신할 때 단위 시간당 네트워크가 받아들이고 전송할 수 있는 비트 수로 정의한다. 전송 율이나 대역폭도 종종 처리율과 같은 의미로 사용된다. 기존의 전화처럼 시간에 따라 비트 율이 일정한 CBR 스트림을 생성하는 응용이 있기도 하지만, 대부분의 응용은 시간에 따라 비트 율이 가변적인 VBR 스트림을 생성한다. 더욱이 모든 네트워크가 과다한 데이터의 손실 또는 지연없이 CBR 스트림을 전송할 능력이 있는 것은 아니다. 데이터 스트림이 가변적인 비트율을 가지는 특성을 버스트성이라고 하는데 이는 평균 비트율/최대 비트율로 측정한다.

ATM 망에서의 서비스 품질은 ATM이 셀 단위로 정보를 전송하므로 셀 레벨에서의 품질을 유지하는 것이 중요하며, 중요한 품질요소로는 셀 전송지연, 셀 지연변이 그리고 셀 손실을 등이다. 셀 전송 지연은 전송 대기 시간을 포함한 하나의 셀이 송신측에서 수신측까지 도달하는데 소요되는 시간으로 두 지점간의 셀 전송지연은 그림 1과 같이 나타낼 수 있다. 그림 1에서 데이터 블록이 전송되는 과정은 접근 지연 전송매체가 가용이거나 네트워크가 전송 준비가 될 때까지 송신측이 기다리는 시간을 말하는데 그

림에서 (1)~(2) 사이의 시간이다.

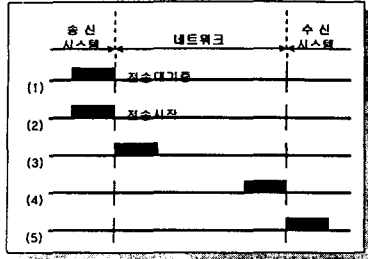


그림 1 두 지점간의 셀 전송 지연
Fig. 1 Cell transmission delay between two nodes

네트워크 전송지연은 송신측이 보낸 데이터 블록의 첫 번째 비트가 수신측에 도달되기까지의 경과시간으로서 (2)~(4) 또는 (3)~(5) 사이의 시간이다. 블록전송지연은 하나의 데이터 블록이 모두 전송될 때까지의 시간으로서 (2)~(3) 또는 (4)~(5) 사이의 시간을 말한다. (1)~(5)까지 경과된 시간을 합하여 종단간(end-to-end delay)이라 부른다. QoS(Quality of Service)는 어떤 응용이 데이터를 전송하기 전에 네트워크에 특별한 요구를 할 때 사용되는 개념으로 네트워크에서는 서비스의 등급을 요구하는 요청들은 사실 상 거의 모두 수용될 수 있지만 통신 네트워크는 응용이 요구한 QoS를 만족시켜줄 수 있는지 여부를 결정할 후 연결을 거절하거나 수락하며 수락한 경우에는 그에 맞게 자원을 예약하는것으로 이를 연결수락제어라고 한다.

ATM 망의 각 통신 매체별 요구품질은 음성이나 영상은 셀 전송지연에 아주 민감한(delay sensitive) 특성을 나타내는데 반하여 음성의 경우 셀의 손실율은 10^{-2} 정도 까지 허용되며 데이터나 차분 압축 영상신호의 경우에는 $10^{-9} \sim 10^{-12}$ 범위의 셀 손실율을 요구한다. 이와는 대조적으로 데이터에 대한 전송지연은 비교적 엄격하지 않은 조건을 갖는다. ATM은 한 호가 연결을 요청할 때 서비스에 따라 요구되는 서비스 품질을 만족할 수 있는 충분한 망 자원이 가능한 경우에만 연결을 허용한다[7,9].

B. 트래픽 왜곡 패턴

트래픽 왜곡은 실시간 채널의 설정 절차에서 여러 가지 응용에 종단간 실행 요구를 근거리 요구로 변환한다. 각각의 노드는 근거리 요구를 충족시켜 근거리 지연 한계를 유도하여 임의의 시간 간격동안 노드에 도달할 수 있는 최대 패킷의 수를 아는 것이 매우 중요하다. 트래픽이 각각의 노드에서 표현자를 만족하면 미리 이를 결정하는 것이 가능하다 따라서 지연한계를 유도할 수 있다.

소스 호스트에서 트래픽 표현자의 매개변수 $(X_{min}, X_{ave}, I, S_{max})$ 를 만족할 수 있는 트래픽을 고려하여 처음 스위치는 나타낸 트래픽 표현자의 매개변수는 $X_{min}, X_{ave}, I, S_{max}$ 이다. 즉, 그림 2 에서 트래픽 왜곡의 패턴은 원래의 표현자에 따르지 않으며 이는 어떻게 X_{min} 이 변하는가를 보여준다. 예를들어 몇 개의 노드를 통과해야만 함으로 최소 상호 도착시간은, X_{min} 보다 더욱 작게된다. 최악의 경우에 두 개의 연속적인 패킷은 0의 상호 도착시간을 갖고 연속적으로 도착한다. 평균 상호도착시간 X_{ave} 가 I 의 상관성 때문에, 이는 연산되는 평균간격에 의존하여 변하므로 지연 한계 분석은 평균 상호도착시간의 인수(factor)로 계산할 한다. 이러한 문제에 대한 접근은 트래픽 소스의 근본적인 특성들로부터 각각의 노드에서 트래픽을 재특성화하는 방법이 제안되어왔다[6,7,10]. 여기서 결정적 흐름 모델은 소스들 특성화하기 위해 이용된다. 이러한 모델에서 만일 간격 u 동안 그 출력의 양이 $\sigma + u \times \rho$ 에 의해 한정되면, 트래픽 소스는 (σ, ρ) 를 만족한다고 말할 수 있다. 여기서 δ 는 최대 버스트 크기이고, ρ 는 트래픽의 평균율이다. 망의 유입점에서 (σ_i, ρ_i) 를 만족하는 접속 트래픽은 스위치 i 의 유입점에서 $(\sigma_j + \sum_{k=1}^i (\rho_k \times d_{kj}) \times \rho_j)$ 로써 특성화된다. 여기서 d_{kj} 는 스위치 k 에서의 연결 j 의 근거리 지연 바운

드로 재특성화된 트래픽을 가짐으로써 근거리 지연 분석은 이루어질 수 있다. 다른 하나의 접근은 비율제어 서비스인 Jitter-EDD(Jitter-Earliest Due Date)라고 하는 비작업 보존(non-working-conserving) 서비스의 클래스를 이용한 각각의 스위치에서 트래픽 패턴을 재구성하는 것이다.

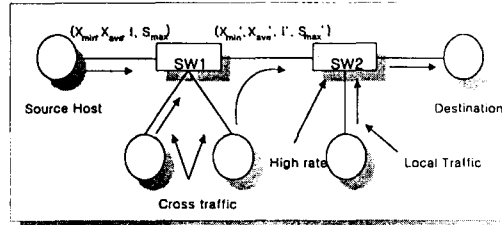


그림 2 트래픽 왜곡 패턴
Fig. 2 Traffic distortion pattern

본 논문에서는 비율 제어를 하기 위해 저장된 우선순위 방식의 하나인 Delay-EDD 방식으로 각각의 노드에 비율 제어가 필요없는 작업보존 서비스(work-conserving)에서 대체 접근 방법을 제안하였다. 그리고 지연 특성을 고려하여 종단간 지연에 추가할 수 있기 때문에 지연 범위 분석에서 전파(propagation)지연과 스위칭 시간은 무시하였다.

III. 제안된 큐 서비스 방식의 전송 지연 특성

지터 특성을 개선하기 위해서 사용하는 가장 일반적인 방법은 작업 비보존 방식을 채택하는 것이다. 그러나 기존의 큐 서비스 방식에서는 연결별로 트래픽을 재구성하는 별도의 조절기 등을 사용하였기 때문에 동작이 복잡하여 구현이 복잡하다는 단점을 가지고 있다. 큐 서비스 방식은 기본적으로 프레임 방식이기 때문에 별도의 트래픽 조절기 없이 지터 특성을 개선할 수 있다. 이를 위해서는 실시간 트래픽을 위한 라운드 로빈 큐와 비실시간 트래픽을 위한 라운드 로빈 큐를 분리시켜야 하기 때문에 복잡해지는 단점이 있으나 연결별로 별도의 조절기 없기 때문에 기존의 큐 서비스 방식들에 비해서는 구현이 용이하다. 지터 특성을 개선하기 위한 큐 서비스 방식의 구조는 그림 3과 같이 나타낼 수 있다.

그림 3에 나타낸바와 같이 제어기에서는 도착한 셀의 마감 시간을 계산하여 라운드 로빈 큐로 보내 주는데, 개선된 라운드 로빈 큐에서는 실시간 트래픽과 비실시간 트래픽을 나누어서 저장한다. 실시간 트래픽 큐와 비실시간 트래픽 큐의 동작은 T_j 가 지날 때마다 우선 순위가 변한다는 점에서는 기본적인 알고리즘과 동일하지만 실시간 트래픽 큐는 작업 비보존 방식을 사용하고 비실시간 트래픽 큐는 작업 보존 방식을 사용한다.

실시간 트래픽 큐에 저장되어 있는 셀 들은 우선 순위가 1이 큐에 저장되어 있는 셀만 전송이 가능하며, 나머지 큐에 저장된 셀들은 그 큐의 우선 순위가 1이 될 때 까지 기다려야 한다.

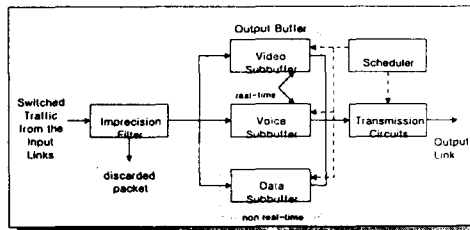


그림 3 출력 버퍼 스케줄러 시스템
Fig. 3 The Output buffer scheduler system

그림 3은 종단간 지연보장을 이용한 실시간 트래픽 제어 를 위해 제안된 큐 서비스형태의 출력 버퍼 스케줄러 시스

템의 구조로서 기존에 제안된 출력 버퍼 시스템과의 차이점은 서버버퍼를 추가한 것으로서 서버버퍼는 실시간 통신을 위해 혼잡제어를 실행할 요소이다. 출력버퍼는 동일한 출력 링크를 공유하는 분할된 세 개의 서버버퍼로 구성되어 있다. 각각의 서버버퍼들은 다른 트래픽 클래스의 패킷들로 저장되는데 이는 영상, 음성, 그리고 데이터이다. 데이터 서버버퍼가 비실시간인 반면에 비디오와 음성 서버버퍼들은 실시간 서버버퍼들이다. 전송회선은 출력 링크 아래로 패킷을 보내므로 서버 버퍼에서의 패킷에 의한 전송과 출력 링크의 사용은 스케줄러에 의해 조정된다. 연결 i 의 셀들이 k 개의 노드를 경유하여 목적지에 도착한다고 하면 연결 i 의 k 번째 셀이 겪게되는 종단간 지연은 식(1)과 같이 나타낼 수 있다.

$$D_i^* = \sum_{j=0}^{k-1} d_{i,j}^* + \sum_{j=1}^k \tau_j \quad \dots\dots\dots (1)$$

단 $d_{i,j}^*$ 는 연결 i 의 k 번째 셀이 j 번째 노드에서 겪는 지연을 τ_j 는 $j-1$ 번째 노드와 j 번째 노드 사이에서 경험하는 전송 선로에서의 전달 지연을 나타낸다. j 번째 노드에서 겪는 대기지연은 제안하는 큐 서비스 방식이 프레임의 사용하여 전송하기 때문에 일반적으로 연결 수락 제어 과정에서 요구하는 전송 지연 한계보다 항상 작게된다. 따라서 연결 수락 제어과정에서 연결 i 에 대해 요구한 전송 지연 한계가 $d_{i,j}$ 라고 하면 연결 i 의 k 번째 셀이 j 번째 노드에서 겪는 대기 지연의 한계는 식(2)와 같다.

$$d_{i,j}^* \leq d_{i,j} \quad \dots\dots\dots (2)$$

위에서 사용된 알고리즘은 작업 보존 방식이기 때문에 셀은 우선 순위가 1이 되기 전에 전송될 수 있다. 즉 라운드 로빈 큐가 비어 있을 때 도착한 셀은 즉시 전송될 수 있다. 따라서 대기 지연의 하한은 j 번째 노드의 한 셀의 전송 시간이 된다. j 번째 노드에서 한 셀을 전송하는 시간이 s_j 이면 식(2)는 식(3)과 같이 수정할 수 있다.

$$s_j \leq d_{i,j}^* \leq d_{i,j} \quad \dots\dots\dots (3)$$

식(2)와 식(3)을 이용하면 종단간 지연의 한계는 식(4)와 같이 표현할 수 있다.

$$\sum_{j=0}^{k-1} s_j + \sum_{j=1}^k \tau_j + D_i^* \leq \sum_{j=0}^{k-1} d_{i,j} + \sum_{j=1}^k \tau_j \quad \dots\dots\dots (4)$$

식(4)는 큐 서비스 방식이 지연 EDD 방식과 마찬가지로 사용자가 요구하는 지연 한계를 보장할 수 있다는 것을 보여 준다.

제안된 큐잉 시스템에서 실시간 트래픽의 지연 특성을 살펴보면 다음과 같다. 우선 연결 i 의 k 번째 셀이 j 번째 노드에서 겪는 대기 지연의 상한은 식(4-9)와 같다. 실시간 트래픽 큐는 작업 비보존 방식이기 때문에 우선 순위가 1인 셀 때에만 전송이 가능하므로 대기지연의 하한은 바로 셀이 저장된 큐의 우선 순위가 1이 되는 때이다. 연결 i 의 k 번째 셀이 저장된 큐의 우선 순위가 1이 되는 순간의 최소는 $d_{i,j-2T}$ 이므로, 연결 i 의 k 번째 셀이 j 번째 노드에서 겪는 대기 지연 한계는 식(5)와 같다.

$$d_{i,j} - 2T \leq d_{i,j}^* \leq d_{i,j} \quad \dots\dots\dots (5)$$

식(1)과 식(5)를 이용하면 종단간 지연의 한계는 식(6)과 같이 표현할 수 있다.

$$\sum_{j=0}^{k-1} (d_{i,j} - 2T) + \sum_{j=1}^k \tau_j \leq d_i^* \leq \sum_{j=0}^{k-1} d_{i,j} + \sum_{j=1}^k \tau_j \quad \dots\dots\dots (6)$$

식(3)을 식(4)와 비교해보면 종단간 지터가 현격하게 감소하고 경유하는 중간 노드의 수와 하위 프레임의 길이에 의해서 결정되므로 쉽게 예측할 수 있다는 것을 알 수 있다. 큐 서비스 방식을 적용하기 위한 연결 수락 제어 기법에 대해 알아보고, (σ, ρ) 모델을 이용하여 기존의 지연 EDD 방식과 연결 가능한 수락범위를 비교한다.

IV. 종단간 지연보장을 통한 실시간 트래픽 연

결수락 제어 분석

(σ, ρ) 모델은 간단하여 분석에 쉽게 사용할 수 있는 장점이 있지만 실제의 트래픽을 너무 과장해서 나타내기 때문에 분석 결과가 실제와 차이가 많이 나는 단점이 있다. (σ, ρ) 모델을 사용했을 때 지연 EDD에서 보장할 수 있는 한계는 연결 지연을 요구하는 지연 한계가 작은 순서부터 연결 번호를 붙이던 식 (7)과 같다.

$$d_i = \frac{\sigma_i + \sum_{j=1}^i (\sigma_j - \rho_j d_j) + \max s_k}{1 - \sum_{j=1}^i \rho_j} \quad \dots\dots\dots (7)$$

큐 서비스 방식에서 보장할 수 있는 지연 한계는 마찬가지로 방법으로 하면 식(8)과 같다.

$$d_i = \frac{\sigma_i + \sum_{j=1}^i (\sigma_j - \rho_j d_j) + \left(\sum_{j=1}^i \sigma_j - 1 \right) \cdot T_j + \max s_k}{1 - \sum_{j=1}^i \rho_j} \quad \dots\dots\dots (8)$$

(σ, ρ) 모델을 사용하여 큐 서비스 방식의 연결 수락 범위를 요구하는 지연 한계가 다른 트래픽을 나타낼 수 있으며, 트래픽의 버스트 특성은 같다고 가정하였다. 일반적으로 버퍼가 스위치의 입력 부분에 있는가 또는 출력 부분에 있는가에 따라 스위치의 구조를 두 가지의 범주로 나눌 수 있다. 버퍼가 스위치의 입력부분에 있는 입력 버퍼 스위치에서 패킷들은 스위치의 입력부에서 저장되고 트래픽 스케줄링 단계와 두 번째, 입력 포트로부터의 선택적 패킷 스케줄링 단계이다. 그리고 버퍼가 스위치의 출력 부분에 있는 출력 버퍼 스위치에서 패킷들은 스위치의 출력 포트에 버퍼되어지는 것으로 간주되며, 패킷들은 스위치에 도달하면 즉시 전송된다.

종단간 지연을 보장하기 위해 연결은 트래픽 특성을 요구하며, 망에서는 지연 한계를 유도하기 위해 버스트성 있는 트래픽을 기술하는 모델 $(x_{min}, x_{avg}, I, S_{max})$ 을 이용한다. 이 모델은 최고 전송율과 일정 기간의 평균 전송율 등의 두가지 파라미터로 버스트성을 나타내는 VBR 트래픽과 같은 경우에 적합한 모델로써 트래픽 특성을 이용한다.

소스 트래픽 표현자는 ATM 트래픽 표현자에 속하는 트래픽 파라미터의 하나의 부분집합으로서 어느 특정한 트래픽 소스에 의해 요구된 연결의 본질적인 트래픽 특성을 포착하기 위해 연결 설정시 사용된다. 연결 트래픽 표현자는 UNI에서 ATM 연결의 트래픽 특성을 명시하기 위한 것으로 소스 트래픽 표현자간의 트래픽 파라미터의 집합과 지터의 허용치 및 ATM 연결에 일치하는 셀들을 정확히 표현하기 위해 사용된다. 본 논문에서는 지연 범위 연산을 구현하고 FCFS 서비스를 위한 새로운 연결 수락제어는 두 가지 단계로 실행된다. 첫 번째, 충분한 자원들이 이용될 수 있는가 하는 단계와 두 번째 채널에 대해 필요한 자원을 비축해 두는 단계로 이루어진다.

처음의 단계에서 다음의 테스트들은 경로를 거쳐 각각의 링크에서 실행되어지는데 이용될 테스트는 식(9)와 같고 종단간 지연테스트는 식(10)과 같으며, 그리고 노드 i 에서 버퍼 요구량은 식(11)과 같다.

$$\sum_{j=1}^i \frac{S_j}{X'_{min}} \times \frac{1}{l} \leq 1 \quad \dots\dots\dots (9)$$

$$d_i = \frac{1}{l} \times \sum_{j=1}^i \left(\frac{d_{k,j} - d_{min}^k}{X'_{min}} \right) S_{max} + \frac{S_{max}}{l} \leq \min(d_{i,j}) \quad \forall j \in n \quad \dots\dots\dots (10)$$

$$B_i \geq \sum_{j=1}^i \left(\left(\frac{d_{i,j}}{X'_{min}} - 1 \right) \right) S_{max} \quad \dots\dots\dots (11)$$

여기서 $d_{i,j}$ 는 연결 j 에 대해 노드 i 에서 받아들일 수 있는 상위 지연 바운드로서 식(8)로서 나타낼 수 있으며, B_i 는 노드 i 에서 이용할 수 있는 버퍼 공간의 양이다. 버퍼 요구는 d_i 에 기반을 두고 있는데 이는 주어진 현재의 채널의 세트에 요구되는 실질적 버퍼들을 나타내기 때문이다. 만일

d_i 가 새로운 트래픽의 추가에 의해 증가된다면 버퍼 요구는 $d_{i,j}$ 를 이용하여 제공할 수 있다. 버퍼의 유용성은 새로운 트래픽이 언제든지 요구되어지고 버퍼 용량이 d_i 에서 어떠한 요인에 따라 증가되면 어느 때라도 체크되기 때문에 채널 j 는 충분한 버퍼 용량을 보장한다. $d_{i,j}$ 의 이용은 또한 비축된 버퍼의 양을 증가시키며, 이는 채널 수락의 가능성을 감소하여 목적 노드에 도달되었을 때 다음 지연 한계 테스트는 식(12)와 같이 실행된다.

$$D \geq \sum_{i=1}^n \bar{d}_i \dots \dots \dots (12)$$

만일 결정 테스트 또는 버퍼 테스트가 어떤 노드에서 실패하거나 마지막 종단간 지연 테스트가 실패하면 채널은 거절된다. 앞 단계에서 노드 지연 바운드마다 노드상의 채널에 주어질 수 있는 최소 지연이 있다. 만일 새로운 채널이 이러한 지연 한계를 예약하면 어떤 새로운 채널도 노드에서 수락되어질 수는 없다. 이것은 더 많은 트래픽 추가가 최악의 지연을 필수적으로 증가시키기 때문이다. 이러한 영향에 따라 채널은 성능이 엄격하게 요구되지 않는 다른 망으로 돌려져야 한다. 지연에 이용할 수 있는 휴지기 Δ_j 는 $\Delta_j = D_j - D_i$ 로 정의된다. 여기서 D_j 는 요구된 종단간 지연 바운드이다.

비율 제어된 망들에서 노드 i 에 대해 할당된 Δ_j 의 일부분으로 $d_{i,j}$ 를 증가시키는 노드들 위에서 휴지기를 분할하는 것이다. 그러나 $d_{i,j}$ 는 노드 i 에서 지터를 증가하기 때문에 이는 망들을 보존하는 작업에 불충분하다. 이러한 지터에서의 증가는 지연을 i 에서 뿐만 아니라 i 노드와 목적 노드 사이에서도 증가한다. 따라서 휴지기의 분할은 초기 노드에서 증가되는 지터에 의해 야기되는 지연뿐만 아니라 현재의 망에서 증가되는 지연도 고려해야 한다.

σ_{i-1} 에 의한 상위 지연바운드의 증가 $d_{i-1,j}$ 는 $\sigma_{i-1} \times S_{max} / (l \times X_{max}^i)$ 로 $d_{i,j}$ 를 증가시킨다. 이는 앞에서 d 에 대한 방정식을 고려함으로써 보여지고 노드 $i+1$ 에서 d_{i+1} 은 $[\sigma_{i-1} + \sigma_{i-1} \times S_{max} / (l \times X_{max}^i)] \times S_{max} / (l \times X_{min}^i)$ 에 의해 증가된다. 노드 $i+1$ 에서의 증가는 이전 노드 즉, 상수에 의한 다중화에서 증가에 대한 합이다. 노드에서 지연 영향에 의한 이러한 증가를 Δ_m 으로 나타낸다. 목적 노드에서 채널 j 에서 발생된 종단간 지연을 다음 식 (13)과 같이 다시 고쳐 쓸 수 있다.

$$D_j + \Delta_j = \sum_{i=1}^n \bar{d}_i + \sigma_n + \sum_{m=1}^n \Delta_m \times \frac{S_{max}}{l} \times \frac{1}{X_{min}^i} \dots \dots \dots (13)$$

여기서 σ_n 은 노드 n 에서 트래픽의 버스트성을 나타내고, Δ_m 은 노드 m 에서 $d_{m,j}$ 이 증가함으로써 발생하는 추가적 지연이다. 이러한 방정식은 종단간 지연을 두 개의 부분으로 분리하는데 이는 최소 종단간 지연과 추가적 종단간 지연이 트래픽의 버스트의 특성에 따라 증가함을 알 수 있다.

V. 시뮬레이션 결과 및 분석

제한된 큐잉 시스템에서 실시간성 트래픽 서비스 품질 평가와 종단간 지연 특성을 분석하기 위해 MATHCAD 툴을 사용하였으며, 또 큐 서비스 방식의 연결 수락 범위를 버스트 트래픽의 크기와 지연한계가 다른 두 종류의 트래픽에 대해서 지연 EDD 방식과 비교하였다.

그림 4는 트래픽의 지연한계를 고려하지 않고 입력되는 소스의 버스트 특성에 따라 요구되는 버퍼의 확률량을 나타내고 있다. 그림에서 나타낸 결과에서 알 수 있듯이 버스트성이 강한 트래픽이 제한된 큐잉 시스템에 입력될 때 버스트 특성이 약한 트래픽보다 버퍼 요구량이 많음을 알 수 있다.

그림 5는 제한된 큐잉 서비스 구조에서 다양한 버스트 특성을 가지고 있는 실시간 VBR 트래픽이 버퍼에서 대기되는 지연 및 지터 특성 즉 셀 지연 변이 허용치를 변화시킴으로써 수락 가능한 망의 부하율을 나타내고 있다. 버스트성 트래픽이 큐잉 시스템에 유입될 때 버퍼 내에서의 누적량을 증가시켜 지연을 증가시킨다.

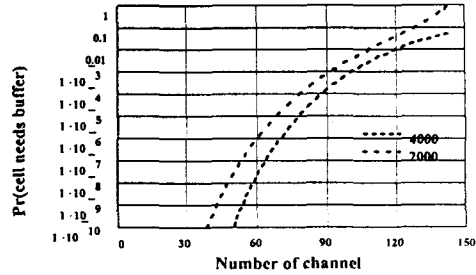


그림 4 버스트 트래픽 특성에 따른 버퍼 요구 확률
Fig. 4 Probability of Buffer required corresponding burst traffic characteristic

그림 5의 결과에서 나타낸바와 같이 버스트성 트래픽 유입은 요구되는 버퍼의 용량을 증가시킴으로써 버퍼 용량 증가에 따른 셀 지연이 증가하여 망내의 트래픽 혼잡이 발생하게 된다. 이러한 혼잡을 회피하기 위해서는 버스트성 트래픽 특성에 따른 셀 지연 변이 허용치를 증가시켜 줌으로써 망의 부하율을 줄일 수 있고 이는 ATM 망의 종단간의 지연보장을 통한 실시간성 서비스의 품질을 보장할 수 있다. 셀 지연 변이가 상대적으로 적을 때 수락 가능한 부하율의 감소가 거의 없음을 알 수 있고 또한 최대 셀 용이 적은 영역에서는 망의 부하율에 비례하여 감소함을 알 수 있다.

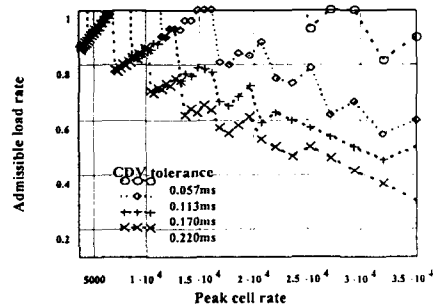


그림 5 셀 지연 변이 허용치에 따른 VBR 트래픽의 수락가능한 부하율
Fig. 5 Admissible load rate for VBR traffic source with different CDV tolerance

그림 6은 지연 EDD 방식의 연결 수락 가능 범위를 나타낸다. 두가지 형태의 트래픽 특성 17ms와 25ms의 전송 지연과 2000과 4000셀의 버스트 크기를 고려하며, 각각 155 Mbps의 전송률을 이용한 컴퓨터 시뮬레이션 결과를 나타내고 있다. 각각 그래프의 아래부분이 연결 수락 가능한 범위이다. 또한 프레임의 길이에 따른 영향을 살펴보기 위해서 제한된 큐 서비스 방식의 전송 프레임의 길이를 6ms, 0.5ms로 변화시켜 가면서 연결 수락범위를 그림 7과 그림 8에서 나타낸 것과 같으며 이러한 결과를 이용하여 지연 EDD방식과 비교하였다. 제한된 큐 서비스 방식을 이용하여 연결 수락 제어는 트래픽 왜곡의 영향을 계산할 수 없는 트래픽 왜곡의 영향이 실제로 존재한다는 가정하에 트래픽 왜곡에 대한 영향에 종단간 지연을 보장함으로써 연결 수락 가능범위를 다양한 프레임의 길이에 대해 나타내었다. 제한된 큐 서비스 방식에서 T_j 가 6ms일 때 지연 EDD방식과 비교해서 많은 차이를 보이지만 T_j 값이 점점 줄어들수록 연결 수락 가능 범위가 지연 EDD 방식에 근접하여 T_j 의 길이가 0.5ms이면 지연 EDD방식의 그래프와 거의 일치하는 것을 알 수 있으며, 지연 EDD 방식과 같은 연결 수락 가능 범위가 차이가 없음을 알 수 있다.

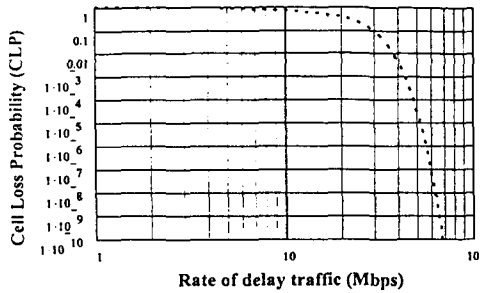


그림 6 지연 EDD 방식의 연결 수락 가능 범위
Fig. 6 Connection admission possibility region of delay-EDD discipline

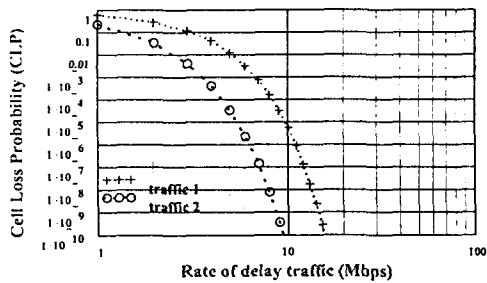


그림 7 제안하는 큐서비스 방식의 연결 수락 가능 범위 ($T_f=6ms$)
Fig. 7 Connection admission possibility region of proposed queue service discipline ($T_f=6ms$)

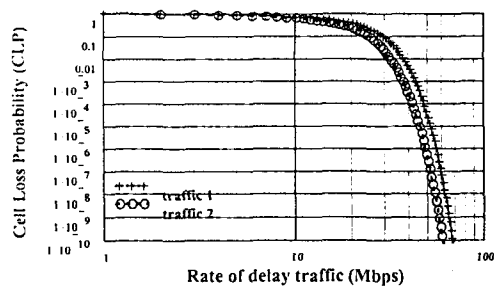


그림 8 제안하는 큐서비스 방식의 연결 수락 가능 범위 ($T_f=0.5ms$)
Fig. 8 Connection admission possibility region of proposed queue service discipline ($T_f=0.5ms$)

VI. 결 론

본 논문은 ATM망에서 효과적으로 중단간 전송 지연 한계를 보장하면서 구현을 단순화 시킬 수 있는 새로운 큐 서비스 방식을 제안하였다. 제안하는 서비스 방식에서는 상위 프레임과 하위 프레임을 따로 정의하여 대역폭 할당과 전송 지연보장을 분리할 수 있기 때문에 유연한 QoS보장이 가능하다.

제안한 큐 서비스 방식은 기본 알고리즘 방식인 지연 EDD 방식이나 RCSP 방식과 같은 지터 조절기를 사용하지 않았기 때문에, 중단간 지연 지터를 보장해 줄 수는 없었다. 그러나 개선된 알고리즘에서는 실시간 트래픽과 비실시간 트래픽을 나누어 저장하기 때문에 연결별 트래픽 조절이 없이도 지연 지터 특성을 개선할 수 있다. 또한 큐 서비스 방식은 기존의 큐 서비스 방식과 달리 실시간 서비스뿐만 아니라 비실시간 서비스의 대역폭을 보장해 줄 수 있다.

또한 큐 서비스에 대한 중단간 지연바운드를 유도하여 제시된 지연분석은 각각의 노드가 비율 제어를 필요로 하지 않았으며, 중단간 지연 한계 연산을 수치적으로 증명되었다. 링크의 모든 접속에 대한 최대 셀율의 합이 링크 속도보다 클 때뿐만 아니라 평균율의 합이 링크 속도보다 작을 때에도 중단간 지연보장이 이루어짐을 알 수 있다. 시뮬레이션 결과는 트래픽 왜곡의 영향이 최악의 경우를 고려하여 모든 채널이 일정한 비율로 셀을 생성하는 기본적인 경우에서 트래픽 왜곡의 영향이 적음을 보였다. 버스트성 트래픽에 대해 이러한 영향은 더욱 증가하며, 모든 경우에 제시된 지연 예측은 정확하다.

제안하는 큐서비스 방식은 기존의 중단간 지연 보장 알고리즘에서 사용하던 셀 정렬을 하지 않고 몇 개의 FIFO 큐로 구성된 라운드 로빈 큐를 이용하여 복잡성을 크게 감소시켰으며, 실시간 서비스와 비실시간 서비스의 서비스 방식을 효과적으로 결합하여 다양한 QoS를 보장해야 하는 ATM망에 효과적으로 적용할 수 있다.

[참고문헌]

- [1] Ian R. Philp and Jane W. S. Liu, "End-to-End scheduling in real-time packet switched networks," In Proceeding International Conference on Network Protocol, Columbus, Ohio, pp. 23-30, October 1996.
- [2] Ian R. Philp, "Scheduling real-time messages in packet-switched network," Ph. D thesis, University of illinois at Urbana-Champaign, 1996.
- [3] J. C. R. Bennett and H. Zhang, "Hierarchical packet fair queueing algorithm," Proceeding of SIGCOMM '96, August 1996.
- [4] S. Jamaloddin Golestani, "A stop-and-go queueing framework for congestion management," In SIGCOMM'90 Symposium, Communication Architecture & Protocols, Philadelphia Pennsylvania, September 1990.
- [5] H. Zhang and S. Keshav, "Comparison of Rate-Controlled Service Disciplines," Proceeding of ACM SIGCOMM'91, Zurich, Switzerland, September 1991.
- [6] H. Zhang and D. Ferrari, "Rate-Controlled Service Disciplines," Journal of High Speed Network : Special Issue on Quality of Service, 3(4), 1994.
- [7] A. Demers, S. Keshav and S. Shenker, "Analysis and simulation of a fair queueing algorithm," In Journal of Internetworking Research and Experience, ACM SIGCOMM'90, pp. 3-12, October 1990.
- [8] L. Zhang, "Virtual Clock : a new traffic control algorithm for packet switching networks," ACM Transactions on Computer Systems, vol. 9, pp. 101-124, May 1991.
- [9] H. Gilert 외 2인, "Developing a cohesive traffic management strategy for ATM networks," IEEE Communication, pp. 36-45, October 1991.
- [10] CCITT SG XVII, TD9, Swiss Geneva Meeting, January 1993.
- [11] C. Woodworth 외 2인, "Congestion Control in ATM Networks," IEEE Globcom'91, pp. 1082-1088, 1991.
- [12] G. Ramamurthy, "A Multidimensional Framework for Congestion Control in B-ISDN," IEEE JSAC, pp. 1511-1520, December 1991.
- [13] J. M. Hyman, A. A. Lazar and G. Pacifici, "Real-time scheduling with quality of service constraints," IEEE JSAC, 9(7) : pp. 1052-1063, September 1991.
- [14] W. K. Shih, J. W. S. Liu and J. Y. Chung, "Fast algorithms for scheduling imprecise computations," In Proceedings of the 10th IEEE Real-Time Systems Symposium, December 1989.
- [15] J. P. Lehoczky, L. Sha and J. Strosnider, "Enhancing aperiodic responsiveness in a hard-real-time environment," In Proceedings of 8th IEEE Real-Time Systems Symposium, pp. 262-270, December 1987.