

효율적인 자원 할당을 위한 사전 예약과 즉석 예약 간 공유 자원 관리

준회원 이 동 훈*, 정회원 김 종 원*

Shared Resource Management Scheme in Advance and Immediate Reservations for Effective Resource Allocation

Dong-Hoon Yi* Associate Members, JongWon Kim* Regular Members

요 약

많은 양의 대역폭을 사용하는 실시간 멀티미디어 응용들은 일정 수준의 서비스 품질을 보장받기 위해 필요한 자원을 사전에 예약하고 이를 사용할 수 있어야 한다. 자원의 사전 예약을 위해서는 해당 자원에 대한 사용 기간의 명시와 이를 이용해 자원을 사전에 예약하고 할당하는 자원 관리자가 필요하다. 따라서 자원의 사용 기간에 관계없이 예약 요청 발생지점에서 가용 자원의 양만을 고려하는 전통적인 자원 할당 방식과는 달리 사전 예약을 제공하는 자원 관리자는 사전에 예약된 자원과 사용 기간을 명시하지 않는 즉석 예약을 위해 할당된 자원이 서로 충돌하지 않도록 예약 요청들을 처리해야 한다. 일반적으로 자원 예약의 충돌은 예약 형태에 따라 할당 자원을 분리함으로써 해결할 수 있지만 이는 자원의 낭비를 초래하게 된다. 본 논문은 즉석 예약과 사전 예약 요청을 위해 준비된 자원 공간을 변동하는 자원 경계를 이용해 효율적으로 공유하는 자원 관리 방식을 제안한다. 제안하는 자원 관리 방식은 네트워크 비용 함수를 이용해 자원 간 경계를 설정하고, 예약 자원의 충돌에 따라 비용 함수의 가중치 변수를 조정해 경계를 변화시킴으로써 공유 자원을 효과적으로 할당하고 관리한다. 또한, 사용자 효용 함수의 정의를 통해 예약된 자원을 사용하는 서비스의 품질로부터 사용자가 얻는 효용을 정량적으로 측정하며, 자원의 사전 예약이 사용자 효용 및 전체 자원 활용에 미치는 영향을 고찰한다. 네트워크 시뮬레이터(NS-2)를 이용한 실험은 제안한 자원 관리 방식이 자원 공간의 고정 분할과 같은 방식과 비교해 높은 자원 활용율과 더불어 안정적인 서비스 품질을 제공할 수 있음을 보여준다.

Key Words: Resource management, advance reservation, resource sharing, dynamic resource boundary, and quality of service.

ABSTRACT

Real-time multimedia applications that require large amount of bandwidth need resource reservation before starting service for providing the QoS(i.e., Quality of Service). To reserve resources in advance, each reservation request has to notify its expectation on the required amount of resources and service duration. Using this information, a resource manager can schedule advance reservations. However, most existing resource management systems are adopting straightforward call admission control process (i.e., only immediate reservation) by checking currently available resources without considering the service duration. Hence, the resource management system that supports advance reservation has to manage confliction caused by indefinite service duration of immediate reservation. Even though the separation of resource pool according to type of reservation can prevent the confliction, it causes low resource utilization. In this paper, we propose an effective resource management scheme that supports both immediate and advance reservations by sharing resources dynamically. Using network cost function, the proposed scheme determines and adaptively adjusts resource boundary according to the confliction rate by varying weight parameters. And also, we define user utility function to quantify user satisfaction based on how well the reserved resource is guaranteed during service time. Simulation results using NS-2 network simulator show that the proposed scheme can achieve better resource utilization with preferable QoS than other schemes like static resource partitioning.

* 광주과학기술원 정보통신공학과 네트워크미디어 연구실. ({dhyi, jongwon}@gist.ac.kr)

논문번호 : 040080-0211, 접수일자 : 2004년 2월 16일

※본 연구는 대학 IT연구센터 육성*지원 사업과 광주과학기술원의 지원에 의하여 수행되었음.

I. 서론

차세대 네트워크의 고속화와 더불어 많은 양의 대역폭을 소비하는 실시간 멀티미디어 응용들이 보편화되고 있다. 또한, 이들 응용을 이용한 다양한 형태의 서비스들이 특정 사용자를 대상으로 유료로 제공되는 경우가 많아지고 있다. 일반적으로 인터넷 방송, 인터넷 강의 및 주문형 비디오와 같은 유료 서비스는 사용자로부터 일정 수준 이상의 서비스 품질(QoS)을 요구받게 된다. 하지만, 현재 인터넷의 최선형 서비스에서는 자원을 독점해 사용할 수 없기 때문에 사용자가 증가하는 경우 서비스의 품질이 저하되는 것을 피할 수 없다. 서비스 품질은 그것이 시작되는 시점과 사용 기간 동안 가용한 대역폭의 양에 따라 많은 영향을 받으므로 안정된 서비스 품질을 유지하기 위해서는 한정된 네트워크 자원을 동시에 사용하는 사용자 혹은 트래픽의 수를 일정 수준으로 제한해야 한다. 자원 관리자는 호 수락 제어(i.e., Call Admission Control)를 이용해 자원을 필요로 하는 사용자로부터 자원 예약 요청을 받아 가용한 자원의 양을 바탕으로 수락 여부를 결정하는 방식으로 동시 사용자의 수를 제한한다^[1]. 이 때, 각 사용자의 자원 예약 요청은 자원을 할당받기 위해 자신이 필요로 하는 자원의 양을 명시해야만 한다. 하지만, 호 수락 제어가 예약 요청이 발생한 시점에서 가용한 자원의 양만을 고려하기 때문에 사용 기간은 일반적으로 명시하지 않는다. 이러한 예약 요청은 수락된 후 바로 자원을 할당받아 서비스를 시작할 수 있기 때문에 즉석 예약 형태로 구분되어 진다. 또한, 수락된 예약을 위해 할당된 자원은 해당 서비스가 종료 될 때까지 다른 예약 요청을 위해 사용되지 않는다. 이것은 가용한 자원이 부족한 경우, 즉석 예약을 통해 필요한 시점에 자원을 예약할 수 없는 상황이 빈번히 발생할 수 있음을 뜻한다.

많은 양의 대역폭 자원을 사용하는 비디오 컨퍼런싱, 비디오 스트림 서비스와 같은 실시간 멀티미디어 응용들의 경우에는 서비스 품질을 위한 자원의 예약이 필요한 시간 내에 이루어져야 한다. 자원의 사전 예약에 대한 개념은 중요한 회의를 위해 미리 회의의 장소를 예약하는 것처럼 네트워크를 통한 다자간 실시간 원격 화상 회의를 위해 필요한 대역폭을 사전에 예약함으로써 안정적인 서비스 품질을 제공받도록 하기 위해 제안되었다^[2,3]. 사전 예

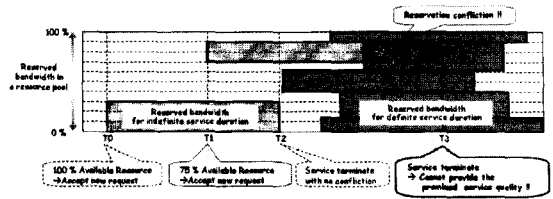


Figure 1. Infinite service duration and advance reservation.

약은 필요한 자원을 미리 예약함으로써 필요한 시간에 예약 실패를 겪지 않고 원하는 품질의 서비스를 안정적으로 받을 수 있다는데 의의가 있다. 이것은 전통적인 호 수락 제어를 통한 즉석 예약과는 달리 미래의 자원을 예약하는 것이므로, 자원 관리자는 자원의 예약 정보(시작시간, 사용기간, 필요 자원 양)를 정확히 알아야 하며 이를 통해 미래의 자원 할당 상황을 파악하고 사전 예약을 수행할 수 있다. 따라서 각 사전 예약 요청은 위와 같은 자신의 예약 정보를 명시해야 한다.

앞에서 설명한 바와 같이, 즉석 예약을 위한 호 수락 제어는 예약 요청이 발생한 시점에서 자원의 가용 여부를 이용하기 때문에 즉석 예약은 사용 기간을 명시하지 않는다. 그림 1은 사용 기간을 명시하지 않는 즉석 예약의 수락으로 발생할 수 있는 예약 자원의 충돌을 설명하고 있다. 하나의 자원 공간을 이용해 두 가지 형태의 자원 예약 요청을 처리하는 자원 관리자는 수락된 즉석 예약에 할당된 자원과 미래에 미리 예약되어 있는 자원 사이에서 발생할 수 있는 충돌을 막을 수 없다. 이는 자원 사용 기간의 명시 유무에서 발생하는 문제이며 효율적인 자원의 할당과 관리를 어렵게 한다. 따라서 사전 예약을 고려하는 자원 관리 방식은 두 자원 예약 간 차이점을 고려해 효율적인 자원의 할당을 행할 수 있는 방법에 초점이 맞추어져 있다.

본 논문에서는 즉석 예약과 사전 예약 요청을 하나의 자원 공간을 이용해 효율적으로 처리하기 위한 자원 관리 방식을 제안한다. 제안한 자원 관리 방식은 각각의 예약 요청에 필요한 자원 공간을 변동하는 경계를 통해 제어한다. 이는 사전 예약과 즉석 예약을 처리하는 자원 관리에서 쉽게 고려되는 자원 공간의 고정 분할로부터 발생하는 불필요한 자원의 낭비를 줄일 수 있다. 또한, 즉석 예약에 할당된 자원과 미래에 예약된 자원과의 충돌 비율에 따라 경계를 유동적으로 변화시킴으로써 예약 자원의 충돌에 따른 서비스 품질의 저하를 막는다. 이를

위해 본 논문에서는 네트워크 비용 함수와 사용자 효용 함수를 정의한다. 네트워크 비용 함수는 기중치 변수의 조정을 통해 변동하는 자원의 경계를 설정하는데 이용되며, 사용자 효용 함수는 예약된 자원을 사용하는 서비스로부터 사용자가 얻는 효용을 정량화하는데 이용된다. 제한한 자원 관리 방식은 네트워크 시뮬레이터(NS-2)¹⁵⁾를 이용한 모의실험으로 평가되며, 결과의 분석과 고찰을 통해 제안한 자원 관리 방식이 공유 자원의 효율적인 할당으로 높은 자원 활용율과 더불어 안정적인 서비스 품질을 제공함을 확인한다.

본 논문은 다음과 같이 구성되어 있다. 2절에서는 사전 예약을 고려하는 자원 관리 방식들의 장단점과 사전 예약을 위해 고려해야 하는 요소들을 고찰하고 네트워크 비용 함수와 사용자 효용 함수를 정의한다. 3절에서 변동하는 경계를 통해 사전 예약과 즉석 예약을 하나의 자원 공간을 이용해 처리하고 관리하는 자원 관리 방식을 제안한다. 제안된 자원 관리 방식은 4절에서 네트워크 시뮬레이터를 이용한 모의실험을 통해 평가된다. 5절에서는 사전 예약을 고려하는 자원 관리와 관련된 여러 연구들을 살펴보고, 이들이 가지는 문제점들을 고찰한 뒤 마지막으로 6절에서 결론을 맺도록 한다.

II. 사전 예약과 자원 관리

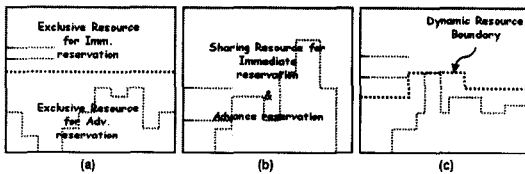


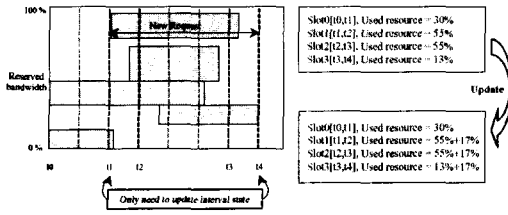
Figure 2. Resource management schemes ((a) static partitioning, (b) full-sharing, and (c) dynamic boundary).

사전 예약과 자원 사용 기간을 명시하지 않는 즉석 예약 사이에서 효율적인 자원의 할당을 달성하기 위해 제안된 여러 방식들 중 가장 간단한 방식은 즉석 예약과 사전 예약을 위한 자원 공간을 고정적으로 나누어 관리하는 고정 경계 방식(그림 2.(a))이다. 고정 경계 방식은 각 자원 공간의 공유를 원칙적으로 막기 때문에, 사용 기간을 명시하지 않는 즉석 예약의 수락으로 인한 사전 예약과의 충돌은 발생하지 않는다. 하지만, 자원의 고정 경계는 각 예약을 위한 자원의 최고 사용량을 예측해 결정하므로 변동하는 네트워크 상황과 한정된 자원의

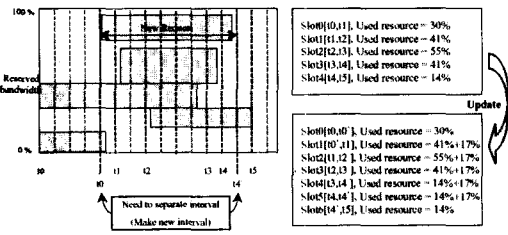
양을 고려하면 적절하지 않다. 이는 사용되지 않는 자원이 다른 서비스의 부족한 자원을 위해 사용되는 것을 불가능하게 하므로 자원의 낭비를 초래하기 때문이다. 고정 경계 방식과는 반대로 자원 공간을 완전히 공유하는 방식(그림 2.(b))은 하나의 자원 공간을 이용하기 때문에 고정 분할 방식에 비해 상대적으로 자원 효율성은 높지만, 사용 기간을 명시하지 않는 즉석 예약으로 인한 예약 자원 간 충돌을 피할 수 없기 때문에 가용 자원이 부족해지면 품질의 저하를 겪는 서비스가 많아지게 된다. 따라서 완전 공유 방식은 자원의 충돌을 최소화하기 위해 정확한 자원 할당 스케줄링과 즉석 예약의 사용 시간 예측이 필요하다. 하지만, 이는 현실적으로 매우 어려운 문제이기 때문에 자원 관리에 있어 큰 단점으로 작용하게 된다. 본 논문에서는 자원 공간의 효율적인 할당을 위해 자원 공간 사이에 변동하는 경계를 설정하며, 예약 자원의 충돌율에 따라 경계를 변화시키는 적응형 변동 경계 방식(그림 2.(c))을 제안한다. 변동하는 자원 경계를 이용하기 때문에 고정 경계 방식에 비해 향상된 자원 활용을 제공하며, 예약 자원 충돌율에 따라 경계가 유동적으로 변화하므로 완전 공유 방식에 비해 간단하면서도 안정적인 서비스 품질을 제공할 수 있다.

1. 사전 예약을 위한 요구사항

자원을 사전에 예약하기 위해서는 자원 관리자와 예약을 요청하는 사용자 간에 반드시 고려해야 하는 몇 가지 문제들이 있다. 첫째, 자원 관리자는 사전 예약이 실행될 시점과 자원의 사용 기간을 알아야 한다. 자원 관리자는 사전 예약 요청이 명시한 예약의 시작 시간과 사용 기간을 시간 테이블을 이용해 시간에 따른 자원 할당량을 기록하고 관리할 수 있게 되는 것이다. 둘째, 시간 테이블의 관리 부하를 조절하기 위해 시간 테이블의 최소 시간 단위를 결정해야 한다. 그림 3은 시간 슬롯의 크기에 따른 슬롯의 생성과 관리에서 나타나는 차이를 설명한다. 정확한 자원 관리를 위해서는 최소 시간 단위가 작아야 하지만, 이로 인해 관리해야 할 슬롯의 수는 많아지게 된다. 셋째, 사전 예약을 즉석 예약과 구별 짓기 위한 기준으로 최소 예약 가능 시간(minimum booking time⁴⁾)을 고려해야 한다. 사전 예약은 자원 관리자의 입장에서 자원의 효과적인 할당과 관리를 가능하게 하기 때문에 즉석 예약에 비해 높은 우선권을 부여 받게 된다. 따라서 어떤 자원 예약 요청이 사전 예약으로서 처리되기 위해



(a) Large interval timeslot



(b) Small interval timeslot

Figure 3. Effect on the timeslot interval to the reservation.

서는 최소 예약 가능 시간 이전에 예약 요청이 이루어져야 한다. 이로써 예약 자원에 충돌이 생겼을 경우 사전 예약은 즉석 예약보다 우선 할 수 있게 되는 것이다.

본 논문에서는 사전 예약을 지원하는 자원 관리자와 환경을 다음과 같이 가정한다. 각 자원 예약 요청은 즉석 예약과 사전 예약 두 가지 형태가 있으며, 즉석 예약은 사용할 자원의 양만을 명시하고 사용 기간은 명시하지 않는다. 사전 예약의 경우 사용할 자원의 양을 포함한 시작 시간과 사용 기간의 정보를 명시하되 사용 기간은 상황에 따라 유동적이라고 가정한다. 시간 테이블의 최소 시간 단위는 1 s.t. (i.e., Simulation Time)이며, 최소 예약 가능 시간은 15 s.t.으로 정의한다.

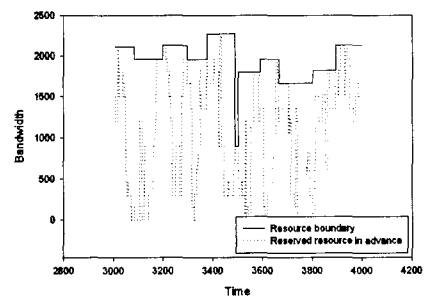
2. 네트워크 비용 함수

사전 예약과 즉석 예약을 동시에 처리하는 자원 관리자는 사용 기간을 명시하지 않는 즉석 예약으로 인한 자원 충돌을 최소화하면서 동시에 한정된 자원을 효과적으로 할당 할 수 있어야 한다. 따라서 각 예약 요청을 위한 자원 공간 사이에 경계를 두어 공유 자원의 양을 제한해야 하며, 자원 할당 상황에 따라 경계를 적절히 변화시켜 공유 자원의 유동성을 제공해야 한다. 본 논문에서는 즉석 예약과 사전 예약에 대한 자원 경계를 유동적으로 결정하기 위해 네트워크 비용(C) 함수(식 1)를 정의한다.

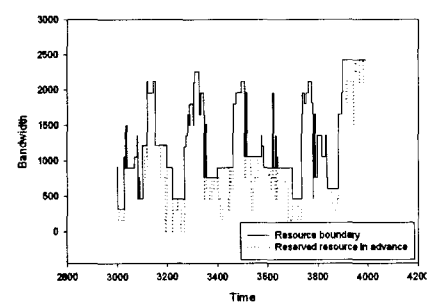
$$C(\tilde{R}(t), R(t)) = \mu \times n_{R(t)} + \sigma \times \int_{t_i}^t [\tilde{R}(t) - \tilde{R}(t)] dt,$$

where, $t \in [t_s, t_e], \mu + \sigma = 1$ (1)

네트워크 비용 함수는 특정 시 구간(i.e., t_s, t_e)내에서 시간에 따른 자원 경계의 변동 횟수($n_{R(t)}$)와 결정된 자원 경계($\tilde{R}(t)$)와 실제 예약된 자원($R(t)$)사이의 차이를 고려한다^[5]. 결정된 경계에 의한 네트워크 비용은 위의 각 요소에 가중치 변수를 적용해 계산한다. μ 는 자원 경계의 변동 횟수에 대한 가중치이고 σ 는 결정된 경계와 실제 할당 자원 간 차이로 인해 낭비되는 자원에 대한 가중치이다. μ 가 σ 에 비해 크게 되면, 네트워크 비용을 결정하는 주요인이 자원 경계의 변동 횟수가 되며 네트워크 비용의 최소화를 위해 자원 경계는 평탄하게 결정된다. 이것은 자원 간 경계가 정적으로 분할됨을 뜻하며, 즉석 예약으로 인한 자원 예약 충돌은 줄어들게 된다. 반대로 σ 가 μ 에 비해 커지게 되면, 네트워크 비용은 결정된 경계로 인해 초과 할당된 자원의 양에 따라 결정되므로, 자원 경계는 실제 할당 상황과 매우 가깝게 결정된다. 이는 즉석 예약으로 인한 예약 자원의 충돌율을 증가시키게 된다. 그림 4는 가



(a) Toward static boundary ($\mu=0.9$)



(b) Toward dynamic boundary ($\sigma=0.9$)

Figure 4. Resource boundary with different weight parameter in network cost function.

중치 변수에 따라 결정된 자원 경계를 나타낸다.

본 논문에서는 네트워크 비용 함수의 가중치에 따라 변화하는 자원 경계를 이용한 자원 관리 방식을 제안한다. 자원 관리자는 사전에 예약된 자원의 할당 상황을 파악하기 위해 시간 슬롯 테이블을 이용하며, 시간 슬롯 테이블에는 시간에 따라 미리 예약되어 사용될 자원의 양이 명시되어 있다. 전술한 바와 같이 사전 예약은 최소 예약 가능 시간 이전에 이루어져야하기 때문에, 이 시 구간 내에서 시간 슬롯 테이블의 자원 할당 정보는 더 이상 변하지 않게 된다. 따라서 자원 관리자는 시간 슬롯 테이블에 저장된 자원 할당 정보와 네트워크 비용 함수를 이용해 최소 예약 가능 시간 내에서 가장 작은 비용을 가지는 자원 경계를 결정한 후 이를 즉석 예약의 자원 할당에 필요한 정보로 사용하게 되는 것이다.

3. 사용자 효용과 유동적 자원 예약

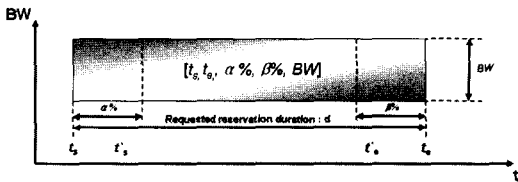


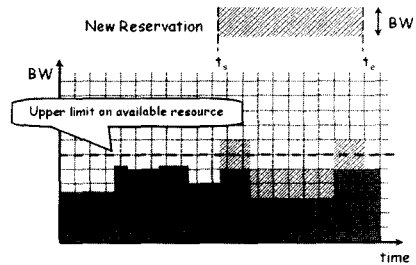
Figure 5. Advance reservation with flexible service duration.

일반적으로 즉석 예약의 경우 사용 기간을 명시하지 않기 때문에 VBR(i.e., variable bite rate)과 같이 자원의 사용량에 대해서만 유동성을 고려할 수 있었다. 하지만, 사전 예약의 경우 시작 시간과 더불어 사용 기간을 명시하기 때문에 서비스 시간에 대해서도 유동성을 고려할 수 있다(그림 5). 자원 관리자는 예약이 가지는 유동성을 이용해 자원을 효율적으로 이용할 수 있게 된다. 사용자 효용(U) 함수는 요청한 자원 예약과 실제 할당된 자원 사이에서 사용자가 느끼는 만족도를 정량화하기 위해 정의되었다. 일반적으로 즉석 예약의 경우, 사용 기간을 알 수 없기 때문에 예약 요청이 발생한 순간에 자원의 가용 여부가 사용자의 효용을 결정하게 된다. 하지만, 사전 예약은 사용 기간을 정확하게 알 수 있으므로 사용자 효용은 미리 예약된 자원이 서비스 기간 동안 얼마나 잘 제공되었는가에 따라 결정된다. 식 2는 변경된 예약으로부터 사용자가 얻는 효용을 정의한다. 이 식에서 사용자의 효용은 요청한 예약의 사용 기간(t_s, t_e)과 실제 할당된 사

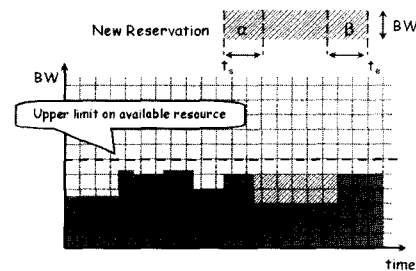
용 기간(t'_s, t'_e)의 차이에 따라 다음과 같이 지수적으로 감소한다고 가정한다. 사용자 효용의 최소값은 U_{min} 이며, 예약이 명시한 유동 구간에 대해서도 예약이 수락될 수 없는 경우에 사용자 효용은 0으로 처리한다.

$$U = \frac{1}{2} \left(e^{\frac{(t'_s - t_s) \ln(U_{min})}{\alpha \cdot d}} + e^{\frac{(t_e - t'_e) \ln(U_{min})}{\beta \cdot d}} \right),$$

where, $t_s \leq t'_s \leq t_s + \alpha \cdot d$, $t_e - \beta \cdot d \leq t'_e \leq t_e$. (2)



(a) Without time flexibility



(b) With time flexibility

Figure 6. Possibility on the resource reservation with time flexibility.

제안하는 자원 관리 방식에서는 사전 예약이 가지는 유동성을 이용해 예약이 가능한 범위에서 최대한으로 가능한 자원 할당을 수행한다. 이 때, 자원 관리자는 사용자 효용 함수를 이용해 변경된 예약으로부터 사용자가 얻을 수 있는 효용을 계산한다. 이로써 사전 예약의 변경을 통해 자원 관리자가 전체 사용자에게 제공할 수 있는 평균 사용자 효용을 비교한다. 그림 6은 사전 예약이 자원의 사용 기간에 유동성을 가짐으로써 예약의 성공과 더불어 향상된 자원 활용을 가능하게 한다는 것을 보여준다.

1 본 논문에서 최대 사용자 효용은 1, 최소 사용자 효용(U_{min})은 0.5로 가정한다.

III. 적응형 자원 관리 방식

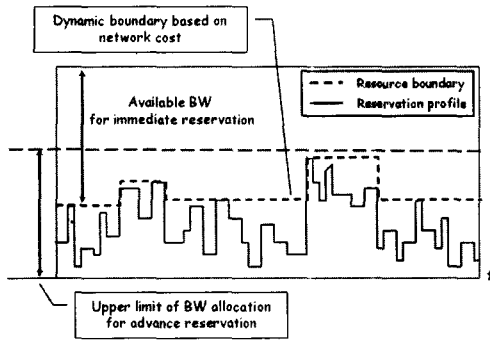


Figure 7. Resource sharing with dynamic resource boundary.

제안하는 적응형 자원 관리 방식은 사전 예약을 통해 얻어진 시간 테이블의 자원 할당 정보와 앞에서 정의한 네트워크 비용 함수를 이용해 자원 간 변동 경계를 설정한다. 결정된 자원 경계는 즉석 예약을 수락하기 위해 필요한 가용 자원의 양을 결정하는데 쓰이게 된다(그림 7). 즉석 예약의 자원 할당에 필요한 자원이 부족한 경우 자원 관리자는 자원 경계를 이용해 현재 사전 예약을 위한 자원 공간 중 사용하지 않는 부분을 빌려오게 된다. 이 때, 사용 기간을 알 수 없는 즉석 예약의 수락과 자원

의 공유로 인해 발생하는 예약 자원의 충돌을 최소화하기 위해, 제안하는 적응형 자원 관리 방식에서는 자원 경계를 결정하는 네트워크 비용 함수의 가중치 변수를 예약 자원 충돌율에 따라 조정함으로써 자원 경계를 적절히 변화시킨다(그림 8). 이로써, 적응형 자원 관리 방식은 높은 자원의 활용도와 더불어, 일정 수준으로 조정된 자원 예약 충돌율을 가지게 되는 것이다.

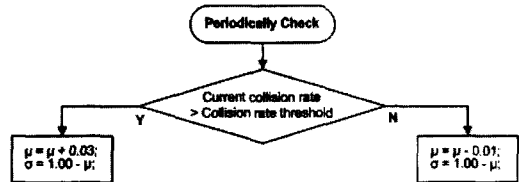


Figure 8. Weight parameter adaptation scheme according to collision rate.

그림 9는 제안한 자원 관리 방식의 전체 흐름을 나타낸다. 즉석 예약 요청이 들어오면, 자원 관리자는 현재 가용한 자원의 양을 확인한 후 예약 요청의 수락 여부를 결정한다. 이 때, 가용 자원이 부족하면 바로 거절하지 않고 사전 예약을 위해 준비된 자원에서 공유를 통해 빌려올 수 있는 자원의 양을 확인한 후 다시 수락 여부를 결정한다. 자원의 공유를 통해 수락된 즉석 예약은 사전 예약과 자원의

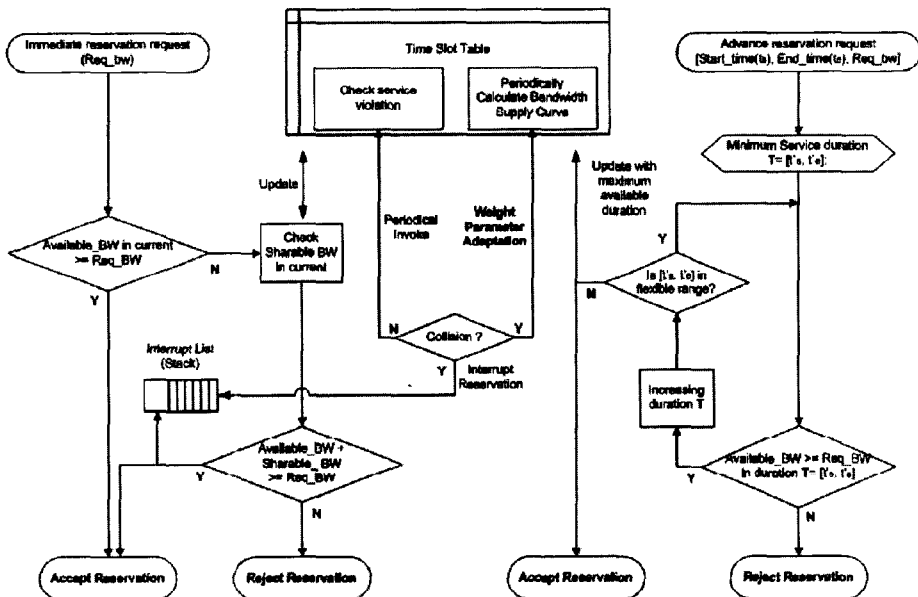


Figure 9. Flow diagram of the proposed resource management scheme.

중복 할당을 발생시킬 수 있기 때문에, 예약 자원 충돌이 발생한 경우 즉석 예약에 할당된 자원을 회수해야 한다. 따라서 공유를 통해 수락된 즉석 예약은 서비스 중단 리스트에 올라가게 된다. 자원 관리자는 주기적으로 자원 예약 충돌율을 확인하고, 그것이 미리 정해진 한계 값 보다 커지게 되면 서비스 중단 리스트에 있는 즉석 예약에 할당된 자원을 사용하는 서비스들을 순차적으로 중단시킨다. 이로써 자원 충돌로 인한 전체 서비스 품질의 저하를 일정 수준으로 유지할 수 있게 되는 것이다. 사전 예약의 경우, 자원 관리자는 사전 예약이 명시한 유동적인 서비스 기간을 이용해 최소 사용 기간 동안 자원이 가용한지를 확인한다. 만약 가용한 자원이 있다면 자원 관리자는 사용 기간을 점차 늘려가면서 최대 가능한 예약 기간을 찾아낸 후 사전 예약을 수락한다. 하지만, 최소 사용 기간 내에도 필요한 자원이 가용하지 않았다면 사전 예약은 거절된다. 수락된 사전 예약의 정보는 시간 테이블에 저장되고, 자원 관리자는 네트워크 비용 함수와 시간 테이블을 이용해 주기적으로 자원 간 경계를 결정한다. 이 때, 네트워크 비용 함수에서 사용되는 두 개의 가중치 변수는 주기적으로 확인되는 자원 충돌율에 따라 조정된다.

IV. 모의실험 결과 및 고찰

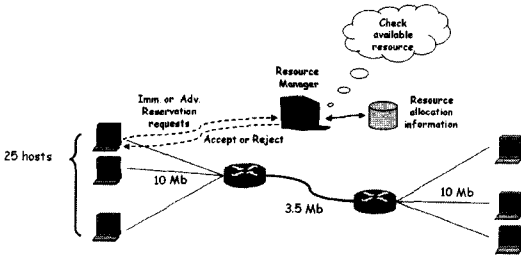


Figure 10. Simulation topology.

QoS(i.e., quality of service)를 제공할 수 있는 네트워크 환경에서 사용자 혹은 멀티미디어 응용들은 자신이 원하는 서비스의 품질을 위해 필요한 자원으로써 네트워크 대역폭을 예약해 사용할 수 있으며, 예약된 자원은 해당 시간에 요청한 대역폭 내에서 서비스를 실행하기 위해 사용되어진다. 따라서

2 앞 절에서 설명했듯이, 자원 예약 간 충돌이 발생했을 때 사전 예약은 즉석 예약에 대해 우선권을 가진다.

네트워크는 이러한 서비스를 위해 대역폭을 미리 준비해 놓아야 하며, 사용자의 예약 요청을 처리하고 대역폭을 관리할 수 있는 자원 관리자가 필요하게 된다. 본 논문에서는 제한한 자원 관리 방식들이와 같은 환경에서 적용하고 평가하기 위해 네트워크 시뮬레이터(NS-2)를 이용한다. 네트워크는 일반적으로 사용되는 덤벨(dumbbell) 토폴로지(그림 10)로 구성하였으며, 양쪽 라우터에는 25개의 호스트들이 10Mbps로 연결되어 일정한 시간 간격으로 사전 혹은 즉석 예약 요청을 발생시킨다. 표 1에는 실험 파라미터들이 표시되어 있다.

Table 1. Simulation parameters.

Simulation parameters		Values
Traffic source		UDP/CBR
Mean reservation time - exponential distribution	Imm. reservation	Varing
	Adv. reservation	40 s.t.
Simulation time		100000 s.t.
Number of hosts		25
Bottleneck link capacity		3.5 Mbps
Upper limit capacity for Adv. reservation		2.0 Mbps
Mean requested bandwidth	Imm. reservation	300 kbps
	Adv. reservation	450 kbps

실험 결과는 즉석 예약의 평균 사용 기간을 변화시켜 트래픽 밀도에 따른 사용자 활용도, revenue, 자원 충돌율, 패킷 분실율 및 예약 요청 거절율을 다른 자원 관리 방식들과 비교해 평가한다. 자원 관리를 통한 자원 활용도는 제공된 revenue로써 정량적으로 계산할 수 있는데, 일반적인 호 수락 제어와 같이 즉석 예약만을 다루는 경우에는 식 33 과 같이 정의된다. 즉 revenue는 성공한 예약들로부터 얻어낼 수 있는 전체 효용으로 생각될 수 있다.

$$Revenue \quad R = (1 - p_b)(D_{avg} \cdot C_d + C_r) \cdot \lambda \quad (3)$$

하지만, 사전 예약과 즉석 예약이 동시에 고려되는 경우에는 예약 자원 충돌로 인한 즉석 예약의 서비스 중지를 고려해야 하기 때문에 식 44와 같이

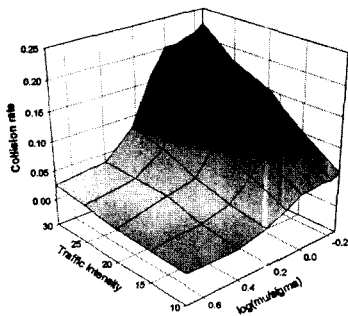
3 p_b : Reservation blocking rate, D_{avg} : Average amount of data transmitted, C_d : Cost for data unit, C_r : Cost for each reservation, λ : Mean holding time of reservation

정의된다. 즉, 전체 revenue는 각 예약이 가지는 revenue에서 예약의 충돌로 인해 즉석 예약에서 없어진 revenue와 서비스의 중단으로부터 발생하는 비용을 뺀 값이 된다.

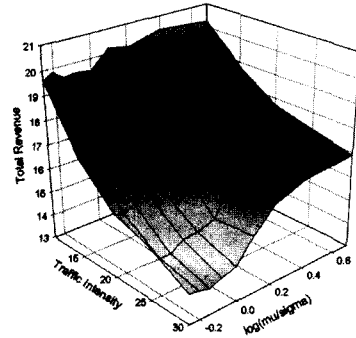
$$R_{total} = (1 - p_c) \cdot R_{imm} + R_{adv} - (1 - p_{b,imm}) \cdot p_c \cdot C_i \cdot \lambda_{imm}. \quad (4)$$

1. 변동 경계와 자원 예약 충돌율 및 Revenue

그림 11은 네트워크 비용 함수의 가중치 파라미터들의 변화와 트래픽 밀도에 따른 자원 충돌율과 제공된 전체 revenue를 보여준다. μ 값이 σ 값에 비해 큰 경우 네트워크 비용 함수는 평탄한 자원 간 경계를 결정하게 된다. 이것은 높은 트래픽의 밀도에 도 낮은 자원 충돌율을 보이며, 향상된 revenue를 제공하게 된다. 반대로, μ 값이 σ 값 보다 작은 경우에는 네트워크 비용 함수는 변동량이 큰 자원 경계를 결정하게 된다. 이는 트래픽 밀도가 증가할 때, 큰 자원 충돌율을 가져온다는 것을 알 수 있다. 이것은 네트워크 비용 함수의 가중치 파라미터를 자원 충돌율에 따라 조정함으로써, 트래픽의 밀도와 관계없이 일정한 수준의 자원 충돌율과 제공할 수 있는 전체 revenue를 향상시킬 수 있음을 뜻한다. 그림 12는 자원 충돌율에 따라 변화하는 네트워크 함수의 가중치 변수의 예이다. 제한하는 자원 관리 방식은 이와 같은 조정을 통해 유동적인 자원 간 경계를 결정하며 이를 통해 낮은 자원 충돌율과 높은 revenue를 제공한다.



(a) Collision rate



(b) Total revenue

Figure 11. Effect on collision rate and total revenue with different weight parameters.

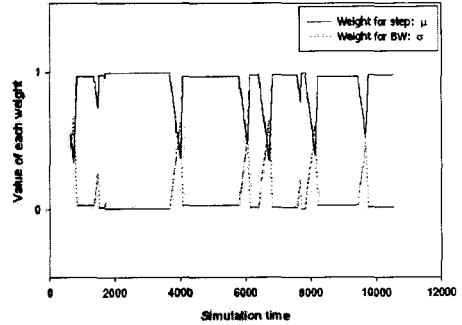


Figure 12. Example on weight parameter adaptation.

2. 변동 경계를 이용한 예약 자원 공유

모의실험에서는 즉석 예약과 사전 예약을 위한 자원 경계를 설정하는 방식에 따라 다음 네 가지 자원 관리 방식을 비교한다. 고정 경계 방식(Static boundary)은 자원 간 경계를 결정한 후 움직이지 않으며, 각 예약은 분할 된 자원 공간의 자원만을 이용해 수락된다. 완전 공유 방식(Full sharing)은 자원 간 경계를 설정하지 않고 전체 자원 공간에서 가용한 자원을 바탕으로 두 예약이 동시에 처리되는 방식이다. 변동 경계 방식(Movable boundary)은 네트워크 비용 함수의 가중치 변수 중 μ 와 σ 의 값을 고정된 상태에서 자원 간 경계를 설정한 경우이다. μ 의 값이 0.9인 경우 자원 간 경계는 고정 경계 방식으로 간주 될 수 있다. 마지막으로, 제안하는 자원 관리 방식은 현재 발생하고 있는 자원 충돌율을 설정된 한계 값과 비교해 네트워크 비용 함수의 가중치 변수를 조정하는 적응형 변동 경계 방식이다. 이는 네트워크 상태에 따라 고정 경계 방식과

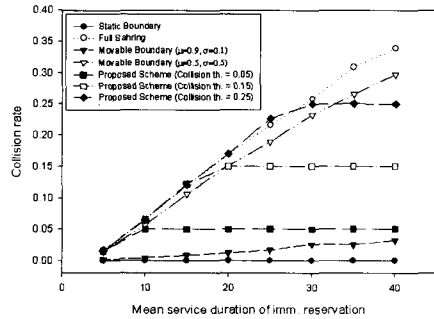
4 p_c : Collision rate, C_i : Cost for each interrupted reservation, R_{imm} : Revenue from imm. reservation, R_{adv} : Revenue from adv. reservation.

완전 공유 방식을 절충하는 자원 관리 방식으로 생각 될 수 있다.

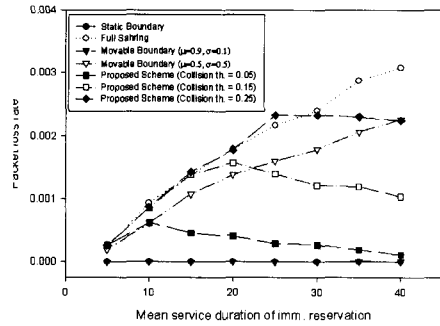
그림 13.(a)는 즉석 예약의 자원 사용 기간에 따른 예약 자원 충돌율이다. 고정적인 자원 경계를 설정하게 되면 자원을 공유하지 않기 때문에 예약 자원의 충돌은 일어나지 않음을 알 수 있다. 완전 공유 방식의 경우에는 트래픽 밀도가 증가함에 따라 예약 자원 충돌율은 포화되지 않고 계속 증가하게 된다. 하지만, 제안한 자원 관리 방식에서는 예약 자원 충돌율이 설정된 한계 값에 포화됨을 알 수 있다. 이것은 제안한 자원 관리 방식이 적절한 수준의 예약 충돌율을 미리 결정하고, 이를 유지할 수 있음을 보여주는 것이다. 그림 13.(b)는 예약 자원의 충돌로 발생하는 패킷 손실을 비교한 것이다. 완전 공유 방식인 경우에는 패킷 손실이 계속 증가하지만, 제안한 방식에서는 패킷 손실 역시 어떤 한계 값에 포화됨을 알 수 있다. 예약 된 자원을 사용하는 멀티미디어 응용이 어느 정도의 패킷 손실을 처리할 수 있다면, 예약 자원의 충돌이 곧바로 서비스의 품질 저하로 이어지는 않기 때문에 자원 관리 시스템에서 예약 충돌율의 한계값을 결정하는데 이용될 수 있을 것이다.

그림 13.(c)는 자원 공유 방식에 따른 즉석 예약의 요청 거절율(예약 실패율)이다. 고정 경계 방식은 완전 공유 방식에 비해 거절율에 있어 일정한 차이를 보이고 있음을 알 수 있다. 이러한 차이는 고정 경계 방식의 경우, 고정된 자원 경계로 인해 즉석 예약을 위해 사용할 수 있는 공유 자원이 없기 때문이다. 제안한 방식은 낮은 트래픽 로드에서는 비교적 낮은 거절율을 보이며, 트래픽 로드가 높아질수록 높은 거절율을 보이는 것을 알 수 있다. 이것은 적응형 변동 경계를 이용해 트래픽 로드가 낮을 때는 가능한 많은 자원을 공유하고, 트래픽 로드가 높아질수록 공유 할 수 있는 자원의 양을 줄이기 때문이다. 그림 13.(d)는 제공된 전체 revenue를 비교하고 있다. 낮은 트래픽 로드에서 제안한 자원 관리 방식은 예약 자원 충돌율의 한계값을 이용해 적응형 변동 경계를 설정함으로써 높은 revenue를 가짐을 알 수 있다. 하지만, 트래픽 로드가 높아질수록 자원 관리로부터 얻을 수 있는 revenue가 고정 경계 방식이 제공하는 revenue와 같아지게 된

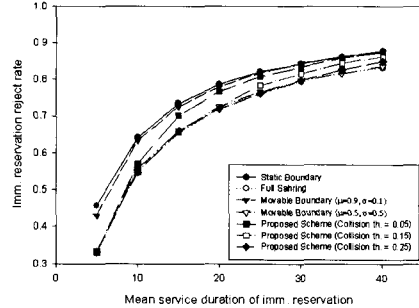
5 각 호스트들은 같은 시간 간격으로 자원 예약 요청을 발생하므로 사용 시간의 증가는 트래픽 밀도의 증가로 고려될 수 있다.



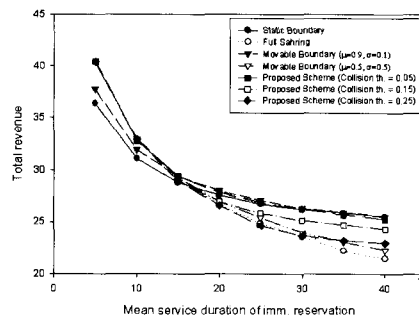
(a) Collision rate



(b) Packet loss rate



(c) Imm. reservation reject rate



(d) Total revenue

Figure 13. Simulation results.

다. 이는 고정 경계 방식은 예약 충돌이 발생하지 않기 때문에 전체 revenue를 결정할 때 예약 자원 충돌로 인해 없어지는 revenue가 없어 높은 예약 거절율에도 불구하고 그림과 같은 revenue를 제공할 수 있는 것이다. 하지만, 앞서 서술한 바와 같이 예약 자원의 충돌율이 실제 해당 서비스의 중단을 의미하지는 않기 때문에, 이를 고려하면 제한한 자원 관리 방식이 가장 많은 revenue를 제공할 수 있음을 알 수 있다.

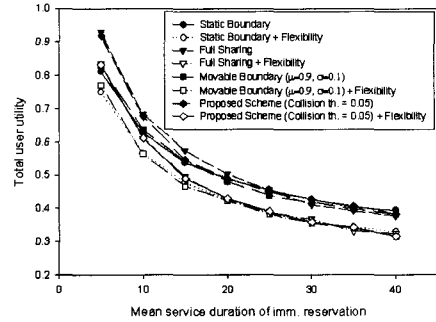
3. 자원 예약 기간의 유동성과 변동 경계

사전 예약이 자원의 사용 기간에 유동성을 가지는 경우 전체 성능에 미치는 영향을 분석하기 위해 모의실험에서는 모든 사전 예약은 사용 기간의 시작과 끝에 10%의 유동 구간을 갖는다고 가정한다. 그림 14.(a)는 평균 사용자 효용으로 사용자가 자신의 예약 요청에 유동성을 상정했을 때 희생해야 하는 사용자 효용을 예약을 요청한 전체 사용자 입장에서 계산한 것이다. 그림 14.(b)는 제공된 revenue로 변경된 예약으로 인한 사용자 활용도의 희생이 전체 revenue에 미치는 영향을 확인할 수 있다. 전반적으로 높아진 revenue를 확인할 수 있는데, 이는 사전 예약에 있어 유동성을 고려하는 경우, 예약 기간의 변화를 통해 사용자 효용을 희생함으로써 좀 더 많은 예약 요청을 수락할 수 있게 됨을 뜻한다.

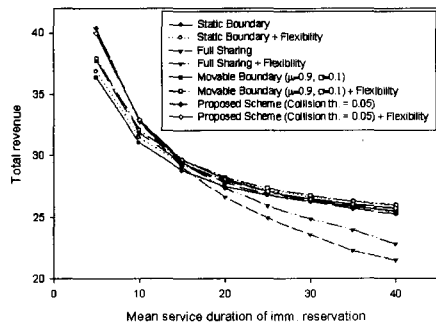
V. 관련 연구

L. C. Wolf et. al.^[6]는 사전 예약을 제공하기 위한 제반 요소(i.e., 시작 시간, 사용 기간, 최소 사전 예약 가능 시간, 자원 할당 실패 시 필요한 기능 등)와 이를 실행하기 위해 필요한 모듈들을 디자인했다. 또한, 일반적인 자원 관리 구조에의 적용을 위한 몇 가지 시나리오를 제안했는데, 제안한 방식은 사용 기간을 명시하지 않는 즉석 예약을 위해 자원을 분할해 관리하는 고정 경계 방식을 택하고 있다. A. Schill et. al.^[7]은 ATM의 IPng, RSVP 프로토콜을 기반으로 사전 예약을 제공하기 위한 시스템의 구조와 이에 필요한 요소들에 대해 논의했다. 이 논문의 사전 예약을 위한 시그널링 모델은 가장 중요한 요소로써 호 수락 제어와 예약 요청이 가지는 사용 기간의 문제를 다루고 있고 이를 해결하기 위해 자원 공간의 고정적인 분할을 제안하고 있다.

보다 향상된 자원의 활용을 위해 D. Ferrari et.^[9]



(a) Total user utility



(b) Total revenue

Figure 14. Effect of advance reservation time flexibility on total user utility and total revenue.

al.는 사전 예약과 즉석 예약을 위한 자원 공간의 제한적인 공유를 허용하는 방식을 제안했으며, Tenet^[8] 프로토콜을 이용하여 실제 사전 예약을 제공하는 자원 관리 시스템을 구성하였다. 그러나 자원 할당의 충돌로 인한 서비스 품질의 저하는 고려되지 않았다. F. Sallabi et. al.^[10]는 미리보기 시간(look-ahead time)을 이용해 즉석 예약의 수락으로 인한 자원 할당의 충돌을 줄이는 방식을 제안했다. 하지만, 미리보기 시간의 결정 방식과 이의 변화가 자원 할당에 미치는 영향에 대해서는 언급하지 않았다. A. G. Greenberg et. al.^[11]는 사전에 예약된 자원의 정보를 이용해 요청된 즉석 예약이 수락되었을 때 이것이 사전에 예약된 자원과 충돌할 확률을 미리 계산한 후, 어떤 한계 값을 기준으로 수락 여부를 결정하는 방식을 제안했다. 이 방식은 각 요청이 일반적으로 자신이 예약한 자원 전부를 사용하지 않는다는 점을 고려하면, 예측 값의 정확성이 자원 활용도에 큰 영향을 주게 된다. 이것은 충돌 확률이 미리 예약된 자원의 양을 고려해 계산되었기 때문에 실제로 발생하는 자원 할당 충돌과는 같

지 않기 때문이다. Degermark *et. al.*^[12]는 사전 예약을 고려하는 측정 기반 수락 제어를 확장한 방식을 제안했다. 이는 현재 활성화되어 있는 서비스들의 과거 자원 사용량을 바탕으로 가까운 미래에 즉석 예약을 위해 사용될 수 있는 자원의 양을 예측하는 방식이다. 이 논문에서 즉석 예약으로 할당된 자원은 완벽한 보장이 필요 없는 서비스를 대상으로 하고 있기 때문에 예약 자원의 총량이 서비스 품질에 미치는 영향이 적으며, 총량이 일어나도 모든 서비스는 정상적으로 시작된다고 가정하고 있다. 하지만, 사전에 예약된 자원을 사용하는 서비스의 경우에는 모든 자원을 보장해 주기 위해 해당 서비스가 시작되기 이전에 즉석 예약 요청을 모두 거절해 자원의 중복을 피하고 있다. 이는 사전 예약이 많아지게 되면 가용 자원이 있어도 즉석 예약을 위한 자원 할당이 이루어지지 않는 문제가 있다. 또한 이 논문에서는 인터넷에서 사용되고 있는 자원 예약 프로토콜들을 이용해 여러 노드에 걸쳐 사전 예약을 구성할 수 있는 방법에 대해서도 논의했다.

자원의 사전 예약 개념은 근래에 많이 연구되고 있는 그리드 컴퓨팅에서도 많은 관심을 끌고 있다. 그리드 컴퓨팅은 지역적으로 분산된 컴퓨팅 자원, 데이터 등을 고속의 네트워크를 통해 공유할 수 있도록 하는 기술로 사용자는 필요한 여러 형태의 자원을 동시에 접근해 사용할 수 있게 된다. 따라서 분산된 컴퓨팅 자원을 효율적으로 관리하고 할당하기 위해 사전 예약을 기반의 자원 관리 방식이 고려되고 있다. [13]과 [14]는 그리드 컴퓨팅에서 사전 예약을 위해 고려해야 할 요소들과 시스템의 디자인을 제안했다.

VI. 결론

많은 양의 대역폭을 필요로 하는 실시간 멀티미디어 응용들은 필요한 시간에 필요한 자원을 사용하기 위해 사전에 자원을 예약 할 수 있어야 한다. 또한, 자원을 사전에 예약하기 위해서 사용자는 자원의 사용 기간을 명시해야 하며 자원 관리자는 사전 예약과 사용 기간을 명시하지 않는 즉석 예약 사이에서 예약의 총량을 최소화하면서 효율적인 자원 할당을 수행해야 한다. 본 논문에서는, 사전 예약과 즉석 예약을 위한 자원 공간 사이에 적응성 있는 변동 경계를 설정하는 자원 관리 방식을 제안했다. 네트워크 비용 함수에서 사용되는 가중치 파라미터의 영향을 고찰하고, 상황에 따라 변화하는

적응형 변동 경계를 이용함으로써 일정 수준의 예약 자원 총출율과 향상된 자원 활용도를 다른 자원 관리 방식과 비교해 확인했다. 또한, 사전 예약이 가지는 사용 기간의 유동성을 고려하기 위해 사용자 효용 함수를 정의하였고 제안한 자원 관리 방식을 통해 사용자가 희생한 효용이 자원 활용도의 향상으로 이어짐을 확인했다.

참고 문헌

- [1] S. Sargento and R. Valadas, "Call admission control in IP networks with QoS support," in *Proc. of the IEEE INFOCOM'00*, Mar. 2000.
- [2] A. Gupta, "Advance reservation in real-time communication services," in *Proc. of the IEEE 22nd Annual Conference on Local Computer Networks (LCN '97)*, 1997.
- [3] W. Reinhardt, "Advance reservation of network resources for multimedia applications," in *Proc. of the 2nd International Workshop on Advanced teleservices and High-Speed Communication Architectures*, pp. 23-33, Mar. 1994.
- [4] O. Schelen and S. Pink, "Sharing resources through advance reservation agents," in *Proc. of the IFIP Fifth International Workshop on Quality of Service (IWQoS'97)*, May. 1997.
- [5] J. Schmitt, O. Heckmann, M. Karsten, and R. Steinmetz, "Decoupling different time scales of network QoS systems," in *Proc. of the 2001 International Symposium on Performance Evaluation of Computer and Telecommunication Systems*, Jul. 2001.
- [6] L. C. Wolf, L. Delgrosii, R. Steinmetz, S. Schaller, and H. Wittig, "Issues of reserving resource in advance," in *Proc. of International Workshop on Network and Operating System Support for Digital Audio and Video (NOSSDAV'95)*, 1995.
- [7] A. Schill, S. Kuhn, and F. Breiter, "Resource reservation in advance in heterogeneous networks with partial ATM infrastructures," in *Proc. of the INFOCOM '97*, pp. 612-619, Apr. 1997.
- [8] A. Banerjee, D. Ferrari, B. Mah, M. Moran, D. Verma, and H. Zhang, "The Tenet real-time protocol suits: Design, implementation, and experiences," *IEEE/ACM Transactions on Networking*, vol. 4, pp. 1-10, Feb. 1996.

[9] D. Ferrari, A. Gupta, and G. Ventre, "Distributed advance Reservation of real-time connection," in *Proc. of the Fifth International Workshop on Network and Operating System Support for Digital Audio and Video*, pp. 15-26, Apr. 1995.

[10] F. Sallabi, and A. Karmouch, "Resource reservation admission control algorithm with user interactions," in *Proc. of the GLOBECOM'99*, vol.4, pp. 2086-2090, Dec. 1999.

[11] A.G. Greenberg, R. Srikant, and W. Whitt, "Resource sharing for book-ahead and instantaneous-request calls," *IEEE/ACM Transactions on Networking*, vol.7, pp. 10-22, Feb. 1999.

[12] M. Degermark, T. Kohler, S. Pink, and O. Schelen, "Advance reservations for predictive service in the Internet," *ACM/Springer Journal of Multimedia Systems*, vol. 5, pp. 177-186, May 1997.

[13] W. Smith, I. Foster, and V. Taylor, "Scheduling with advanced reservation," in *Proc. of the 14th International Parallel and Distributed Processing Symposium (IPDPS'00)*, pp. 127-132, May 2000.

[14] I. Foster, C. Kesselman, C. Lee, B. Lindell, K. Nahrstedt, and A. Roy, "A distributed resource management architecture that supports advance reservation and co-allocation," in *Proc. of the International Workshop on Quality of Service (IWQoS '99)*, pp. 27-36, Jun. 1999.

[15] UCB/LBNL/VINT, *Network simulator - ns (ver. 2)*. <http://www.isi.edu/nsnam/ns>, 1998.

이 동 훈 (Dong-Hoon Yi)

준회원



2001년 2월: 인하대학교
전자공학과 학사
2003년 2월: 광주과학기술원
정보통신공학과 석사
2003년 3월~현재:
광주과학기술원
정보통신공학과 박사과정

<주관심분야> Internet QoS, Network Resource Management, Call Admission Control. Network Calculus

김 종 원 (JongWon Kim) 정회원



1987년 2월: 서울대학교
제어계측공학과 학사
1989년 2월: 서울대학교
제어계측공학과 석사
1994년 2월: 서울대학교
제어계측공학과 박사
2001년 9월~현재:

광주과학기술원 정보통신공학과 부교수
2000년 7월~2001년 6월: 미국 InterVideo Inc., Fremont, CA, 개발자문
1998년 12월~2001년 7월: 미국 Univ. of Southern California, Los Angeles, CA, EE-Systems Department 연구조교수
1994년 3월~1999년 7월: 공주대학교 전자공학과 조교수
<주관심분야> Networked Media Systems and Protocols focusing "Reliable and Flexible Delivery for Integrated Media over Wired/Wireless Network"