부하제어형 서비스 지원을 위한 이더넷 대역폭 관리기의 설계 및 구현

(Design and Implementation of Ethernet Bandwidth Manager for Supporting Controlled-Load Service)

김 병 식 † 이 재 국 †† 전 우 직 †††
(Byeongsik Kim) (Jaekuk Lee) (Woojik Chun)

요 약 인터넷에서 기존의 최선형 서비스 외에 QoS를 제공하기 위한 노력으로 IETF에서는 새로운 서비스 모델로 통합서비스 모델과 이 서비스의 지원을 위한 자원 예약 시그널링 프로토콜로 RSVP를 제시하였다. 그러나 RSVP는 링크 계층의 기술과는 독립적으로 정의되었기 때문에 실제 사용하는 링크 계층의 기술에 맞는 자원 예약 방법과 서비스 매핑 방법이 필요하다. 이를 위해서 IETF에서는 RSVP를 IEEE802 형태의 망에 매핑 시키기 위한 방법으로 SBM을 제시하였다. 그러나 IETF의 SBM은 RSVP의 시그널링 메커니즘에 종속적이기 때문에 항상 RSVP와 함께 사용되어야 한다는 제한점과 RSVP의 시그널링 경로를 데이타 경로도 그대로 따라야함으로써 데이타 전달이 늦어지는 단점이 있다. 본 논문에서는 RSVP와 독립적이면서 시그널링 경로와 데이타 전달 경로를 분리한 새로운 메커니즘을 제시한다. 또한, 현재 가장 널리 쓰이고 있으면서 특성상 대역폭 관리가 어렵다고 알려져 있는 링크 계층 기술인 이더넷에서 대역폭 관리를 수행할 수 있는 이더넷 대역폭 관리기에 대해서 기술한다. 이더넷 대역폭 관리기의 설계를 위해서 이더넷 대역폭 특성을 파악하기 위한 시뮬레이션을 수행하였으며, 이 결과를 이용해 실제 이더넷 대역폭 관리기를 구현한 결과를 보인다. 또한 본 논문에서 제시하는 메커니즘은 네트워크 계층의 시그널링 프로토콜과는 독립적으로 설계되었기 때문에 SBM의 제한점 및 한계점을 극복할 수 있으며 이 메커니즘을 이용해서 이더넷 상에서 기존의 최선형 서비스 이외에 통합서비스 중에서 부하제어형 서비스를 제공할 수 있음을 확인할 수 있었다.

Abstract To support Integrated Services in the Internet, the Internet Engineering Task Force(IETF) has developed new mechanisms. These new mechanisms include Resource Reservation Protocol(RSVP) and Integrated Services model. RSVP and Integrated Services are independent of the underlying link layer technologies and it is necessary to define the mapping of RSVP and Integrated Services specifications onto specific link layer technologies. IETF proposed a Subnet Bandwidth Manager(SBM) providing a method for mapping RSVP onto IEEE802 style networks. However, the SBM has several constraints and problems. For solving these constraints and problems, we design and implement a new Ethernet Bandwidth Manager(EBM). This paper discusses (1) new bandwidth management mechanism independent of RSVP, (2) simulation of Ethernet bandwidth to analyze the characteristics of Ethernet and accomplish Ethernet bandwidth management, (3) design and implemetation of EBM that allows Controlled-Load Service, and (4) results of our experiments on EBM.

· 본 연구는 '99년도 한국전자통신연구원의 연구비 지원으로 이루어진 것임
† 정 회 원 : 충남대학교 컴퓨터공학과
bskim@ce.cnu.ac.kr
†† 비 회 원 : (주)디지털라이프 연구원
jklee@ce.cnu.ac.kr
††† 종신회원 : 충남대학교 컴퓨터공학과 교수
chun@ce.cnu.ac.kr
논문접수 : 2000년 1월 11일
심사완료 : 2000년 5월 25일

1. 서 론

최근 데이타 링크 계층 구현 기술의 발전으로 인해서 다양한 QoS(Quality of Service)를 요구하는 멀티미디어 응용들의 발전이 가속화되고 있다. 그러나 현재의 인터넷은 최선형(Best-Effort) 서비스만을 제공하는 아키텍쳐로 되어 있기 때문에 멀티미디어 응용들의 다양한

QoS 요구사항을 충족시킬 수 있는 새로운 메커니즘의 개발에 관심이 집중되고 있다. 그 결과 IETF(Internet Engineering Task Force) 에서는 기존의 최선형 서비스와 부가적인 새로운 QoS 서비스가 통합된 새로운 서비스 모델인 통합서비스(Integrated Services) 모델을 제시하였다[1].

인터넷에서 통합서비스를 지원하기 위해서 IETF 및 관련 연구 기관에서는 전반적인 골격과 프로토콜들 그리고 새로운 기술에 대한 표준들을 개발하고 있다. 그 중의 하나가 RSVP(Resource Reservation Protocol)로 특정 데이타 플로우에 대해서 일정한 양의 자원이나 특정 서비스를 요청하도록 하고 그 플로우에 대해서 필요한 자원들을 예약해 놓을 수 있는 메커니즘을 제공하는 시그널링 프로토콜이다[2]. 그러나 RSVP는 IP 계층에서의 자원 예약을 위한 프로토콜로 데이타 링크 계층의 기술과는 독립적으로 정의되어 있고, RSVP 라우터들 사이에 여러 개의 서로 다른 구현 기술을 사용하는 서브넷이 존재하는 경우에는 부가적으로 고려해야 할 사항이 많이 존재한다. 즉, RSVP를 이용해 실제 자원을 예약해서 통합서비스를 지원하기 위해서는 실제 사용하는 링크 계층의 기술에 맞는 자원 예약 방법이나 서비스 매핑 방법을 필요로 한다[3][4][5].

IETF에서는 여러 가지 데이타 링크 계층 구현 기술로 RSVP의 자원 예약 메커니즘을 매핑 시키는 연구의 일환으로 SBM(Subnet Bandwidth Manager)을 제시하였다[3]. 그러나 SBM은 기존의 RSVP 시그널링 프로토콜의 메커니즘을 그대로 이용하는 방법을 취하고 있기 때문에 항상 RSVP와 같이 사용되어야 한다는 문제 이외에도 부수적인 여러 가지 문제를 가지고 있다.

본 논문에서는 SBM의 간단한 개요 및 문제점을 제시하고 이를 개선한 새로운 이더넷 대역폭 관리기인 EBM(Ethernet Bandwidth Manager)을 제안한다. 또한 EBM을 이용해서 서브넷의 대역폭을 관리함으로써 통합서비스 모델 중에서 부하제어형(Controlled-Load) 서비스와 기존의 최선형 서비스를 특정 플로우에게 제공할 수 있는 가능성을 보인다. 본 논문의 구성은 다음과 같다. 제 2절에서는 인터넷 통합서비스와 RSVP에 대한 소개와 IETF의 SBM에 대해서 기술한다. 제 3절에서는 SBM의 단점과 이더넷에서 부하제어형 서비스 제공을 위한 이더넷의 특성을 시뮬레이션한 결과 및 이더넷에서의 부하제어형 서비스 제공을 가능하게 해주는 이더넷 대역폭 관리기인 EBM의 동작 메커니즘과 EBM이 사용하는 메시지들에 대해서 기술한다. 제 4절에서는 EBM의 구현 방법에 대해서 그리고 제 5절에서는 EBM을 사용해서 이더넷에서도 부하제어형 서비스를 제공할 수 있는 가능성을 보인다. 마지막으로 제 6절에서는 결론 및 향후 연구 방향에 대해서 기술한다.

2. 관련 연구

과거에는 QoS 제공을 위해서 프로토콜을 구현할 때 고려해야 할 사항[6], 운영체제와 관련된 프로세스들의 스케쥴링 및 프로세스간의 통신에 관련된 사항[7][8], 그리고 프로토콜 스택 중에서 수송 계층(Transport Layer)에서 고려해야 할 사항[9][10][11]등에 관련된 연구들이 각각 독립적인 형태로 진행되었다. 그러나 종단간의(End-to-End) QoS를 지원하기 위해서는 그림 1에서 보는 바와 같이 End System내의 입출력 장치, 입출력 버퍼, 운영체제 프로세스들, 프로토콜 스택 및 네트워크에서의 QoS 제어 및 관리가 통합적으로 이루어져야 한다[12]. 그러나 본 논문에서는 데이타 링크 계층의 기술로 이더넷(Ethernet)을 사용하는 서브넷에서 QoS 지원 방법에 관한 부분으로 범위를 한정하기로 한다.

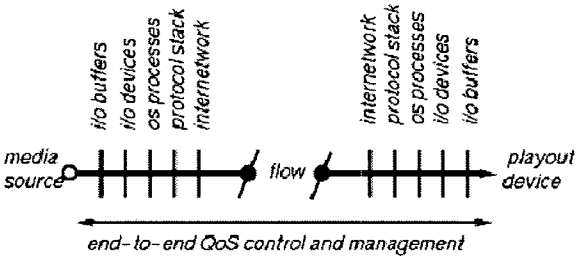


그림 1 종단간 QoS

2.1 인터넷 통합서비스와 RSVP

인터넷 통합서비스(Integrated Services) 모델[1]은 기존의 최선형 서비스(Best-Effort Service) 외에 엄격한 지연 한계를 필요로 하는 응용들을 위한 보장형 서비스(Guaranteed Service)[13]와 망의 부하가 많이 걸려 있을 때에도 부하가 많이 걸려 있지 않은 상태에서 기존의 최선형 서비스가 제공하는 정도의 QoS를 제공할 수 있는 서비스로 정의되는 부하제어형 서비스(Controlled-Load Service)[14]로 구성된다.

보장형 서비스는 특정 패킷들에게 일정 수준의 대역폭 보장, 엄격한 종단간(End-to-End) 지연 한계, 큐잉으로 인한 패킷의 손실이 없도록 제공되는 서비스를 의미한다. 이 서비스는 엄격한 실시간 전송을 필요로 하는 응용들을 위해서 제공되는 서비스이며 이 서비스의 목적은 패킷 스위칭 망에서 전용선에 의해서 제공되는 서

비스의 보장성을 달성하고자 하는데 목적이 있다.

IETF에서는 부하제어형 서비스란 부하가 많이 걸려 있지 않은 상황에서 기존의 최선형 서비스가 제공하는 정도의 서비스를 망의 어떠한 상황에서도 제공할 수 있는 서비스로 정의한다. 일반적으로 기존의 최선형 서비스를 기반으로 하는 인터넷은 망에 부하가 많이 걸려 있지 않을 때에는 성능이 어느 정도 유지되지만 부하가 많이 걸리는 경우에는 성능이 급격히 저하되는 성질을 가지고 있다. 이런 망에 부하제어형 서비스를 도입하면 항상 일정하게 어느 정도의 망 성능을 필요로 하는 응용들에게 적합한 서비스를 제공할 수 있다. 부하제어형 서비스의 정의는 이 서비스에 관련된 명확한 지연 한계에 대한 정의가 없으며 심지어는 패킷 손실률에 대한 제한 사항도 포함하고 있지 않기 때문에 승인 제어, 스케쥴링, 버퍼 관리에 있어서 상당한 융통성을 허용한다.

RSVP는 응용들을 위해서 자원을 예약하기 위한 시그널링 프로토콜로 개발되었다[2]. 그림 2는 RSVP의 시그널링 절차를 보여 준다. 송신자(Sender)는 자기가 생성할 트래픽의 특징을 포함한 *Path* 메시지를 수신자(Receiver)로 보낸다. 이때, *Path* 메시지가 따라가는 경로에 위치하는 모든 라우터들은 라우팅 프로토콜에 의해서 결정되는 다음 홉으로 *Path* 메시지를 전달한다. 수신자는 *Path* 메시지를 수신한 후 자기가 필요로 하는 플로우를 위해서 자원을 예약할 목적으로 *Resv* 메시지로 응답한다. *Resv* 메시지가 송신자 쪽으로 전달되면서 라우터들은 *Resv* 메시지의 자원 예약 요구를 승인할 수도 있고 거절할 수도 있다. 이때 만일 예약 요구가 거절되는 경우에 라우터는 자원 예약 요구에 대한 실패를 지시하는 *ResvErr* 메시지를 수신자 쪽으로 전달하고 시그널링 과정이 종료된다. 그 밖의 메시지들은 *Path* 메시지 전달시 오류가 발생한 경우에 송신자 쪽으로 전달되는 *PathErr* 메시지, 라우터에 유지되는 상태를 제거하기 위한 *PathTear* 메시지와 *ResvTear* 메시지, 자원 예약 요구에 대한 확인을 요청하는 *ResvConf* 메시지 등이 있다.

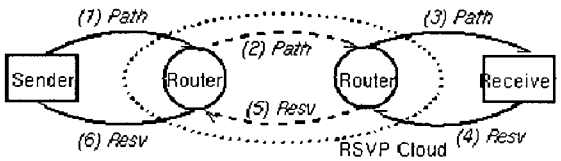


그림 2 RSVP 시그널링

2.2 IETF의 SBM

인터넷은 기본적으로 최선형 서비스만을 제공한다. 그러나 최근 링크 계층의 기술 발전과 실시간 응용들의 증가로 인해서 인터넷상에서 실시간 서비스를 제공할 수 있는 메커니즘의 개발에 관심이 증가하고 있다. 이런 실시간 응용들이 최선형 서비스 외에도 다양한 품질의 서비스를 요청할 수 있도록 새로운 인터넷 아키텍처와 서비스 모델의 확장에 관한 연구가 진행되었다. 이 결과로 자원 예약 프로토콜인 RSVP와 새로운 서비스 클래스를 정의한 통합서비스 모델이 제시되었다. 그러나 RSVP와 통합서비스 모델은 링크 계층의 기술과는 독립적으로 정의되었기 때문에 이더넷, 토큰 링, 프레임 릴레이, ATM등과 같은 여러 가지 형태의 링크 계층의 기술에 맞는 자원 예약 방법과 서비스 매핑 방법을 필요로 한다[3][4][5]. 이를 위해서 IETF issll(Integrated Services over Specific Link Layers) 워킹 그룹에서는 RSVP를 IEEE802 형태의 랜에 매핑 시키기 위한 방법으로 SBM을 제시하였다[3].

하나의 랜 세그먼트 안에서 이 세그먼트를 관리하는 기능을 담당하는 SBM은 SBM 기능을 갖는 장치(호스트 혹은 스위치) 중에서 선출된다. SBM의 선출은 각각의 SBM 기능을 갖는 장치들 중에서 우선 순위를 기반으로 하는 선출 알고리즘에 의해 동적으로 이루어진다. SBM이 되지 못한 그 밖의 장치들은 SBM에게 대역폭 관리 및 할당 요청을 하는 클라이언트로 동작한다. 선출된 SBM이 수행하는 기본적인 승인 제어 절차는 다음과 같다.

- **SBM** 초기화 : 각 랜 세그먼트에서 예약될 수 있는 대역폭 양에 대한 정보를 얻어서 초기값을 설정하는 단계
- **클라이언트** 초기화 : 각 클라이언트는 자기에게 붙어있는 각각의 인터페이스에 대해서 SBM이 존재하는지를 검사하고 자기 자신이 SBM 기능을제공할 수 있는 경우에는 SBM 선출 작업에 후보자로 참여
- **SBM의 승인 제어 절차** : 그림 3에서 보는바와 같이 클라이언트는 RSVP의 *Path* 메시지를 기존의 세션 주소로 전달하는 것이 아니라 SBM으로 전달한다. SBM은 *Path* 메시지를 받아서 필요한 경우에는 Adspec을 수정하고 해당 세션에 대한 PATH 상태(Path State)와 이 *Path* 메시지를 전달한 이전 홉의 주소를 저장하고 RSVP의 시그널링 메커니즘에 따라 *Path* 메시지를 라우터로 전달한다. 이런 방식으로 *Path* 메시지는 중간의 라우터들을 거쳐 최종적으로 수신측 호스트까지 전달된다. 이 *Path* 메시지를 받은 수신측 호스트는 RSVP 메커니즘에 의해서 자원 예약을 위한

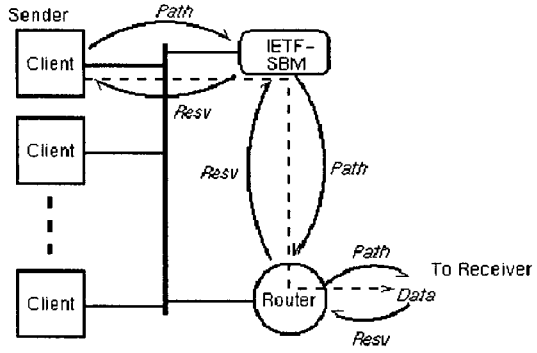


그림 3 SBM의 대역폭 요청 및 할당 메커니즘

Resv 메시지를 생성해서 전달하게 되며 이 *Resv* 메시지는 *Path* 메시지가 전달된 역 방향으로 전달된다. 최종적으로 송신측 호스트가 위치한 서브넷의 SBM이 받았을 때 수신측에서 요청한 대역폭의 허용 가능 여부를 판단하고 그 결과에 따라 2.1절에서 설명한 바와 같이 허용 가능한 경우에는 *Resv* 메시지를 송신측 호스트로 전달하고 허용이 불가능한 경우에는 *ResvErr* 메시지를 수신측 호스트로 전달한다.

이와 같은 승인 제어 및 수신자가 초기화하는 (Receiver-Initiated) 자원 예약 절차를 거친 후 데이타 전송 승인을 받은 송신측 호스트는 *Path* 메시지가 전달되었던 경로로 데이타 전송을 시작한다.

3. 이더넷 대역폭 관리기

IETF에서 제안한 SBM은 몇 가지 제한사항이 있다. 본 절에서는 먼저 SBM의 제한사항 및 단점에 대해서 기술하고 이런 제한사항을 해결한 새로운 시그널링 메커니즘에 대해서 제안한다. 또한 여러 가지 링크 계층 구현 기술 중에서 현재 가장 널리 쓰이고 있으며ATM과 달리 QoS 제공이 어려운 특성을 가지고 있는 이더넷에서의 자원 관리 방법과 이 방법을 이용해서 실제로 한 서브넷에서 부하제어형 서비스와 기존의 최선형 서비스의 제공을 가능하게 해주는 이더넷 대역폭 관리기인 EBM(Ethernet Bandwidth Manager)의 설계 및 구현에 대해서 기술한다. 인터넷 통합서비스 중에서 보장형 서비스는 본 논문에서 대상으로 삼은 링크 계층 구현 기술인 이더넷이 공유 미디어 방식인 CSMA/CD[15] 방식을 기반으로 하고 있기 때문에 근본적으로 제공이 불가능하다고 판단하였다. 따라서 EBM이 제공할 수 있는 서비스는 부하제어형 서비스와 기존의 최선형 서비스로 한정하였다.

3.1 SBM의 단점 및 제한사항

SBM은 IEEE802 형태의 랜(LAN) 상에서 RSVP 메시지들 내에 몇 개의 객체를 추가해서 대역폭의 승인 제어를 하는 시그널링 프로토콜이므로 RSVP 기능이 없는 노드들을 위해서는 사용이 불가능하다. 이처럼 RSVP에 지나치게 종속적인 성질로 인해 나타나는 부가적인 제한사항으로는 그림 4에서 보는 바와 같이 자원 예약을 위한 시그널링 메시지들이 항상 SBM을 통해 전달되므로 시그널링 경로와 항상 같은 경로를 따르는 데이타의 전달 경로도 항상 SBM을 거쳐가야 한다는 단점이 있다. 즉, 송신자로부터 수신자까지의 패스에 라우터의 수가 n 이고 송신자(Sender)와 첫번째 라우터(*R1*) 사이, 각 라우터들(*R1*~*Rn*)사이, 그리고 최종 라우터(*Rn*)와 수신자(Receiver) 사이에 각각 하나의 SBM이 존재한다고 가정할 때, *Path* 메시지와 이의 역방향으로 진행되어야 하는 *Resv* 메시지가 거쳐가야 하는 전체 홉수는 $2(n+1)$이 된다. 또한 시그널링 종료 후 전달되는 데이타는 *Path* 메시지가 전달된 경로를 따라가야 하므로 데이타 패스도 역시 $2(n+1)$ 만큼의 홉수를 거쳐가야 한다. 이런 영향으로 데이타 패킷 전달의 지연 및 시그널링에 필요한 시간도 길어지는 단점이 있다.

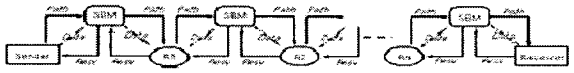


그림 4 SBM의 데이타 전달 경로

SBM은 이상에서 언급한 단점 외에도 몇 가지 제한사항을 가지고 있다. 첫째, SBM은 링크 계층을 위한 어떠한 트래픽 제어 메커니즘도 제시하지 못하고 있다. 따라서 링크 계층에서의 트래픽 제어 메커니즘이나 우선 순위 큐잉 메커니즘이 제공되지 않는 경우에 SBM은 단지 RSVP 플로우들이 생성하는 전체 트래픽의 양을 제한하는 승인 제어 메커니즘으로만 사용이 가능하다[3]. 둘째, 통합 서비스 트래픽들에 대해서 최선형 서비스 트래픽들이 미치는 영향으로 인한 성능 저하를 막을 수 있는 메커니즘의 제시가 없다. 즉, 통합 서비스 트래픽들이 최선형 트래픽들보다 더 향상된 서비스를 보장받을 수 있는 방법이 결여되어있다.

3.2 이더넷에서 부하제어형 서비스 제공을 위한 시뮬레이션 모델

본 논문에서는 공유 미디어 방식의 이더넷에서 호스트들이 생성하는 최선형 트래픽의 양을 적절히 제어하여 부하제어형 서비스를 요구한 호스트의 플로우에게

요구한 서비스를 제공하는 방법을 기본 아이디어로 하고 있다. 이를 위해서 네트워크 시뮬레이터인 *ns-2*[16]를 이용하여, CSMA/CD랜 (이더넷)을 시뮬레이션해서 최대 처리율(Throughput)을 알아내고 다른 트래픽들에게 할당된 대역폭에 영향을 주지 않음과 동시에 호스트들이 일정 수준을 유지하면서 사용할 수 있는 대역폭이 얼마인지를 알아보았다.

그림 5에서는 이더넷의 대역폭을 10Mbps로 지연을 2ms로 했을 때, 1Mbps로 데이타를 생성하는 호스트의 수를 증가시킬 때 나타나는 이더넷의 처리율을 패킷 크기별로 나타내었다. 이 결과로 10Mbps 이더넷에서 각 호스트가 1,024바이트의 패킷을 1Mbps로 계속해서 생성할 때 전체 이더넷 처리량은 7Mbps에서 8Mbps정도임을 알 수 있다. 이 그래프에서 패킷의 크기가 커질 수록 처리율이 커지는 이유는 패킷의 크기가 클 수록 초당 전송되는 패킷의 갯수가 적어지므로 호스트에서 패킷이 전송되는 시점에 충돌이 발생하여 백오프(Back-Off)될 확률이 적어지기 때문이다. 또한 호스트의 갯수가 증가할 수록 처리량이 작아지는 현상은 패킷당 충돌이 일어날 확률이 높아지기 때문이다.

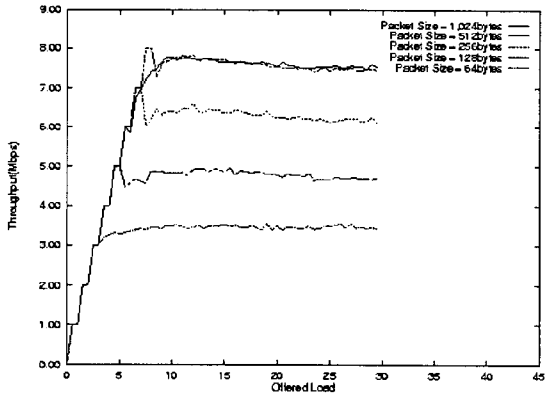


그림 5 이더넷에서 패킷 크기에 따른 처리율의 변화

이더넷에서 부하제어형 서비스를 제공하기 위해서는 부하제어형 서비스를 필요로 하는 트래픽(부하제어형 트래픽)들이 최선형 서비스를 필요로 하는 트래픽(최선형 트래픽)들로부터 영향을 받지 않고 일정 수준의 대역폭을 유지할 수 있어야 한다. 이를 위해서 임의의 트래픽들을 부하제어형 서비스를 받기 위해서 대역폭을 예약받아야 할 예약형 트래픽(Reserved Traffic)으로 또, 예약이 필요 없는 트래픽들을 최선형 트래픽(BestEffort Traffic)으로 구분해서 선정하였다. 그리고, 예약형 트래픽을 각각 1Mbps, 2Mbps, 3Mbps, 4Mbps, 5Mbps로 유지시킨 상태에서 최선형 트래픽을 1Mbps씩 단계적으로 증가시키면서 예약형 트래픽의 변화를 측정하였다. 그 결과 그림 6에서와 같은 그래프들을 얻었다. 이 그래프들의 의미는 예약형 트래픽을 1Mbps로 유지하고 최선형 트래픽을 1Mbps씩 단계적으로 증가시킬 때, 최선형 트래픽이 5Mbps에서 6Mbps로 증가되는 시점에 예약형 트래픽이 1Mbps로 유지되지 못하고 떨어지는 현상을 볼 수 있다 (그림 6(a)). 그림 6(b)의 경우는 예약형 트래픽을 2Mbps로 유지하고 최선형 트래픽을 1Mbps씩 단계적으로 증가시킬 때, 최선형 트래픽이 4Mbps에서 5Mbps로 증가되는 시점에 예약형 트래픽이 2Mbps로 유지되지 못하고 떨어지는 현상을 볼 수 있다. 그림 6(c)에서 처럼 예약형 트래픽을 4Mbps로 유지하고 최선형 트래픽을 1Mbps씩 단계적으로 증가시킬 때, 최선형 트래픽의 사용량이 3Mbps가 되어도 예약형 트래픽의 4Mbps가 유지됨을 볼 수 있지만 전반적으로 예약형 트래픽과 최선형 트래픽의 합이 총 6Mbps를 넘지 않는 경우에 예약형 트래픽의 대역폭을 유지할 수 있음을 알 수 있다. 이를 표 1에 정리하였다.

표 1 예약형 트래픽과 최선형 트래픽의 관계

예약형 트래픽	1Mbps	2Mbps	3Mbps	4Mbps	5Mbps
최선형 트래픽	5Mbps	4Mbps	3Mbps	2Mbps	1Mbps
합	6Mbps	6Mbps	6Mbps	6Mbps	6Mbps

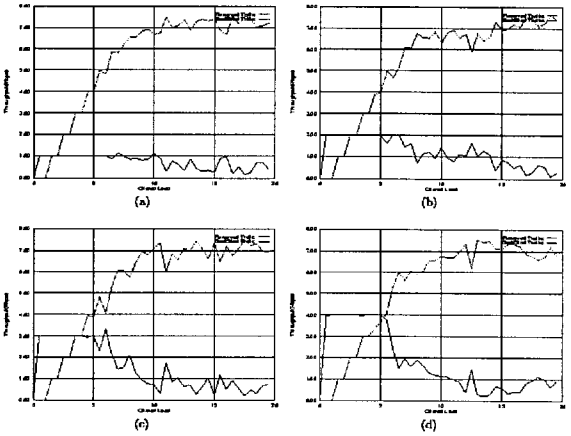


그림 6 예약형 트래픽과 최선형 트래픽의 관계

3.3 EBM의 기본 시그널링 메커니즘

EBM은 이미 언급한 SBM의 단점을 극복하고, 이더넷으로 구성된 서브넷에서 최선형 서비스와 부하제어형

서비스를 제공하기 위해서 설계되었다. EBM을 사용해서 서브넷의 대역폭을 관리하는 메커니즘은 다음과 같다. 이 중에서 EBM을 선출하는 과정은 [3]에서 SBM을 선출하는 알고리즘을 그대로 이용한다. 한 서브넷에서 EBM으로 선출된 호스트에게 대역폭 할당을 요청하는 나머지 호스트들은 클라이언트라 한다. 본 논문에서는 한 서브넷내의 모든 호스트들은 모두 EBM 모듈을 가지고 있어서 EBM 선출 작업에 참여할 수 있음을 가정으로 한다. 선출 알고리즘에 의해서 선출된 EBM은 다음과 같은 절차로 대역폭 관리를 수행한다.

- **EBM 초기화** : 서브넷내의 한 호스트가 EBM으로 선출되면 현 서브넷 내에서 사용 가능한 대역폭의 정보를 알아오기 위해서 서브넷 내의 링크에 대한 트래픽을 모니터링하기 시작한다.

- **클라이언트 초기화** : EBM 이외의 호스트들은 클라이언트가 되며, 이 클라이언트들은 EBM으로부터 오는 *IAmEBM* 메시지내의 주소를 보고 EBM의 주소를 알아내서 저장한다.

- **대역폭 요청, 할당, 관리** : 그림 7에서 대역폭을 요청하고 할당하는 메커니즘을 보여 준다. 클라이언트는 EBM에게 자신이 필요로 하는 대역폭의 양을 *Band_Resv_Req* 메시지를 통해서 EBM에게 요청하고 EBM은 대역폭 할당 정책에 따라 가용 대역폭이 있으면 할당 여부를 *Band_Resv_Rply* 메시지를 통해서 클라이언트에게 전달한다.

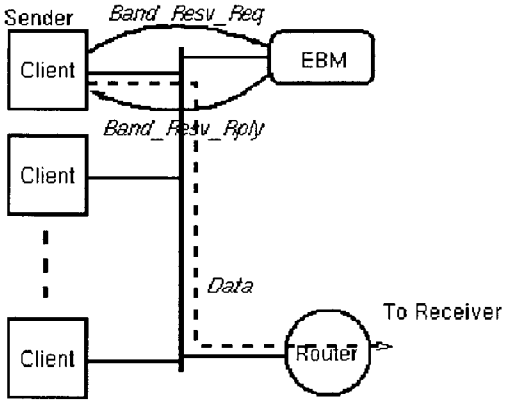


그림 7 EBM의 대역폭 요청 및 할당 메커니즘

위의 절차가 모두 끝나게 되면 자원을 예약한 상태가 되어 데이타를 전송할 준비가 완료된다. 이때 데이타의 전달 경로는 라우팅 정보에 의해서 라우터로 전달된다. 본 논문에서 제안하는 EBM의 시그널링 메커니즘이 여러 세그먼트로 확장되는 경우에, 그림 8에서와 같이 송신자로부터 수신자까지의 패스에 라우터의 수가 n 이고 송신자(Sender)와 첫 번째 라우터(*R1*) 사이, 각 라우터들(*R1*~*Rn*)사이, 그리고 최종 라우터(*Rn*)와 수신자(Receiver) 사이에 각각 하나의 EBM이 존재한다고 가정할 때, *Band_Resv_Req* 메시지와 *Band_Resv_Rply* 메시지의 홉수는 모두 *n+1*이 된다. 또한 데이타 전달 경로도 *n+1* 만큼의 홉수만을 거치면 되므로 IETF의 SBM에 비해서 데이타 패킷 전달에 걸리는 시간 및 시그널링에 필요한 시간도 반으로 줄일 수 있으며 RSVP와는 독립적으로 수행될 수 있다는 장점을 가지고 있다. 또한, EBM은 SBM과는 달리 송신자가 초기화하는 (Sender-Initiated) 자원 예약 절차를 거쳐 데이타 전달이 이루어진다.

3.4 EBM의 메시지들

EBM과 클라이언트들이 사용하는 메시지에는 EBM의 선출 및 관리에 관련된 메시지와 대역폭의 할당, 변경, 철회에 관련된 메시지로 나눌 수 있다. EBM 선출 메시지는 224.0.0.16 멀티캐스트 주소를 사용하여 EBM이 될 수 있는 장치들끼리 주고받으며, EBM의 관리 메시지는 224.0.0.17 멀티캐스트 주소를 이용하여 IETF SBM의 리프레쉬 시간인 5초 간격으로 리프레쉬 한다. 대역폭의 할당, 변경, 철회 메시지는 EBM의 유니캐스트 주소로 EBM과 클라이언트가 통신한다.

그림 8 EBM의 데이타 전달 경로

3.4.1 EBM 선출 및 대역폭 관리 메시지

그림 9는 EBM의 선출 및 대역폭 관리 메시지들의 구조를 보여 주고 있으며 각 필드의 의미는 다음과 같다.

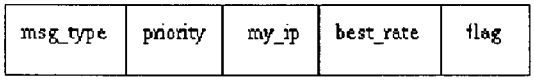


그림 9 EBM 선출 및 대역폭 관리 메시지들의 구조

- **msg_type** : 메시지의 종류를 구별하는 필드로 *EBM_willing* 메시지와 *IAmEBM* 메시지를 구별한다. *EBM_willing* 메시지는 서브넷 안에 호스트들 중에서 동적으로 EBM을 선출하기 위한 메시지로 이 메시지를 받은 호스트들은 자신의 우선 순위와 비교하면서 최종적으로 가장 높은 우선 순위를 갖는 장치가 선출되게

된다. *IAmEBM* 메시지는 EBM으로서 5초 간격으로 자신이 EBM임을 알리는 역할을 함으로써 서브넷 안의 클라이언트들은 EBM이 누구인지를 알고 대역폭 할당 요청을 할 수 있게 한다. 또한 EBM은 네트워크 내의 트래픽 상태를 모니터링해서 best_rate와 flag 파라미터를 가지고 서브넷 안의 최선형 트래픽을 조절한다.

- **priority** : 자신의 우선 순위를 표시하고 EBM 선출에 이용된다.

- **my_ip** : EBM을 선출하고자 할 때 혹은 자신이 EBM임을 알리는 필드로서 자신의 유니캐스트 주소를 삽입한다.

- **best_rate** : 이 필드는 *IAmEBM* 메시지에만 사용되고 *EBM_willing* 에는 사용되지 않는다. 자신이 EBM이 되었을 때 자신이 관리하는 서브넷의 호스트들이 최선형 트래픽으로 생성할 수 있는 최대 대역폭 값을 알리기 위한 필드이다.

- **flag** : 이 필드역시 *IAmEBM* 메시지에서만 사용되며, EBM이 네트워크을 모니터링 하는 도중에 관리하고자 하는 대역폭 임계치를 넘었을 때 자신이 관리하는 호스트들에게 최선형 트래픽의 생성을 반으로 줄이도록 하기 위한 플래그이다. 이로써 네트워크내의 최선형 트래픽을 조절할 수 있다.

3.4.2 대역폭 할당, 변경, 철회 메시지

그림 10은 대역폭 할당, 변경, 철회에 관련된 메시지들의 구조를 보여 주고 있으며, 각 필드의 의미는 다음과 같다.

msg_type	serv_type	flag	flow_id	Tspec

그림 10 대역폭 할당, 변경, 철회에 관련된 메시지들의 구조

- **msg_type** : 메시지의 종류 표시하는 필드로서 대역폭 할당을 요청하는 *Band_Resv_Req* 메시지, 대역폭 할당 요청에 대한 응답을 위한 메시지로 *Band_Resv_Rply*, 대역폭 변경 요청을 위한 *Band_Alt_Req* 메시지와 이에 대한 응답을 위한 *Band_Alt_Rply* 메시지, 대역폭 철회 요청과 이에 대한 응답을 위해서 사용하는 *Band_Retr_Req* 와 *Band_Retr_Rply* 메시지 등이 있다.

- **serv_type** : 요청하는 서비스의 종류를 나타내는 필드로 현재는 사용되지 않고 부하제어형 서비스와 보장형 서비스를 요청하기 위해서 사용될 수 있다.

- **flag** : 메시지의 요청에 대한 응답으로 요청의 수락 여부를 표시하는 필드

- **flow_id** : 각 플로우 별로 대역폭을 할당받아 트래픽을 생성하기 때문에 각 플로우를 구별할 필요가 있다. 이를 위하여 송신자의 주소와 포트, 수신자의 주소와 포트, 프로토콜ID로 플로우를 정의한다. 이 필드의 정보로 패킷을 분류하여 특정 서비스를 제공한다.

- **Tspec** : 트래픽의 생성을 조절하기 위한 정보로 이를 구성하는 파라미터는 토큰 버켓 파라미터(r,b)와 피크율(p), 패킷의 최대 크기인 MTU(M)와 최소 크기(m)로 구성되어 있다. 이 값들을 이용하여 트래픽 생성을 조절하며 할당받은 대역폭을 사용할 수 있게 한다. 삭제 요청 및 응답 메시지에서 이 필드는 사용되지 않는다.

3.5 EBM의 선출 및 대역폭 관리 메커니즘

EBM은 한 서브넷 안에서 대역폭 관리 능력이 있는 호스트들 중에서 동적으로 선출되며, 선출 기준은 현재 선출 과정에 참여하고 있는 호스트들 중에서 가장 우선 순위가 높은 호스트가 EBM이 된다. 각 호스트에 디몬(Daemon) 형태로 존재하는 EBM 모듈은 *Idle, Detect, Elect, EBM* 등의 네 가지 상태중의 한 상태로 유지된다. EBM 모듈이 처음 동작을 시작하면 *Detect* 상태가 되고 일정 시간(ListenInterval)동안 이미 EBM으로 선출된 호스트에서 생성하는 *IAmEBM* 메시지나 EBM이 되고자 선출 작업에 참여하기 원하는 호스트가 생성하는 *EBM_willing* 메시지를 기다린다. 만일 이 시간 동안 *IAmEBM* 메시지가 수신되면 이미 EBM이 존재한다는 것을 의미하므로 현재 EBM의 IP 주소와 우선 순위 값을 저장 한 후 *Idle* 상태로 전이한다. 만일 *EBM_willing* 메시지가 수신되는 경우에 이 메시지를 보낸 호스트의 우선 순위와 자신의 우선 순위를 비교해서 자신이 높으면 *EBM_willing* 메시지를 전송한 후 *Elect* 상태로 전이하고 자신의 우선 순위가 낮은 경우에는 그대로 *Elect* 상태로 전이한다. *Elect* 상태에서는 일정 시간 동안 *EBM_willing* 메시지에 실려온 우선 순위를 비교하여 가장 높은 우선 순위를 갖고 있는 호스트가 EBM이 되어 *IAmEBM* 메시지를 멀티캐스트하고 *EBM* 상태로 전이한다. *EBM* 상태에 있는 호스트는 주기적으로 *IAmEBM* 메시지를 전송함으로써 자신이 EBM이라는 것을 서브넷에 알린다.

EBM은 자신이 관리하는 서브넷 안의 대역폭을 관리하며 클라이언트에 의한 대역폭 할당 요청에 대해서 승인 여부를 결정한다. EBM이 선출되면 주기적으로 *IAmEBM* 메시지를 멀티캐스트한다. 이때, 이 메시지는

자기가 EBM임을 알리는 목적이외에도 모니터링을 통해서 얻어진 정보를 근거로 현재 링크의 상태를 알려주는 역할도 한다. *IAmEBM* 메시지 내의 flag 파라미터는 현재 서브넷의 트래픽 양이 대역폭 임계치에 대한 초과 여부를 나타내는 파라미터이다. 만약 이 flag 파라미터가 세팅되어 각 클라이언트들에게 전달되면 각 클라이언트들은 부하제어형 트래픽들을 위해서 최선형 트래픽의 생성률을 조절하는 토큰의 생성률을 반으로 감소시킨다. 그리고 flag 파라미터가 다시 세팅되어 올 때까지 토큰 생성률을 지수적으로(exponential) 증가시킨다. 이를 "Slow Increase" 메커니즘이라 한다.

만일 현재의 EBM에 문제가 발생해서 EBM이 새로 선출되는 경우에는 기존의 대역폭 할당에 관련된 정보를 모두 잃어버리는 문제가 발생한다. 이런 문제의 해결을 위해서 각 클라이언트들은 *IAmEBM* 메시지를 수신할 때마다 현재 EBM의 주소를 보고 이 EBM이 새로운 EBM인 경우에는 기존의 대역폭 할당에 관련된 정보를 새로운 EBM에게 알려줌으로써 지속적으로 대역폭 할당 서비스를 받을 수 있다.

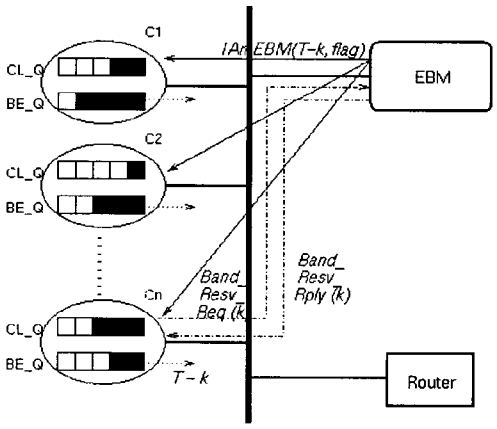


그림 11 EBM의 대역폭 관리 메커니즘

그림 11에서 클라이언트 *Cn*이 EBM에게 *k*만큼의 대역폭을 시그널링을 통해 할당받아서 사용하고 있고 각 클라이언트들은 최선형 트래픽을 내보낸다고 할 때, 최선형 트래픽이 사용하는 총 대역폭의 합과 *Cn*이 할당받은 대역폭 *k*의 합이 대역폭 임계값 *T*를 초과한다면 EBM은 모니터링을 통해 임계값 *T*를 초과하였음을 감지하고 flag 파라미터를 세팅한 후, best_rate 파라미터를 *T-k*로 해서 *IAmEBM* 메시지를 멀티캐스트 한다. 이제 최선형 트래픽을 생성할 수 있는 최대값은 *T-k*가 되고 "Slow Increase" 메커니즘에 따라 flag 파라미터가 또 다시 세팅되어 올 때까지 즉, *T-k*까지 최선형 트래픽의 생성률을 증가시킨다.

4. EBM 구현

EBM으로 동작이 가능한 한 호스트 내에서 구현한 모듈들은 그림 12에 나타나 있으며 이 호스트가 EBM으로 선출되어 동작하는 경우에는 점선으로 표시된 모듈들만이 동작하고 클라이언트로 동작하는 경우에는 실선으로 표시된 모듈들이 동작하도록 구현하였다.이더넷의 모니터링 기능을 위해서는 데이타 링크 계층을 접근할 필요가 있으므로 네트워크 인터페이스를 프로미스큐어스(promiscuous) 모드로 세팅하고 패킷 포착 라이브러리 *libpcap*을 이용하였다.

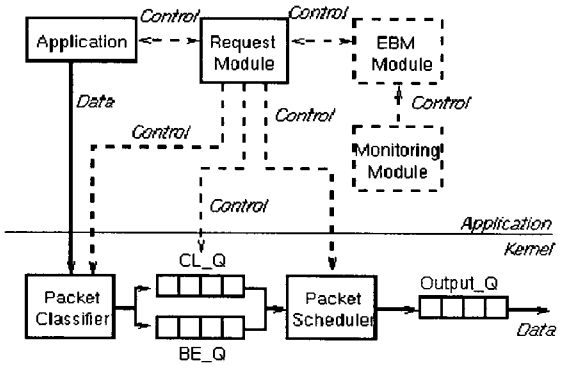


그림 12 EBM으로 동작하는 호스트의 모듈들

각 호스트에서 서비스 큐는 그림 12에서 보는 바와 같이 최선형 서비스용 큐(BE_Q)와 부하제어형 서비스용 큐(CL_Q)로 나뉘어진다. 각각의 큐는 토큰 버켓으로 구현하였다. 호스트가 시작되면 최선형 서비스큐가 생성되어 모든 패킷들은 최선형 큐에서 큐잉되어 전송되고, 요청 모듈(Request Module)과 EBM에 시그널링을 통해 대역폭을 할당 받은 플로우는 부하제어형 큐를 생성한다. 요청 모듈은 큐 생성에 필요한 파라미터(토큰 버켓 파라미터 *r*, *b*)를 커널에 세팅하고 패킷 분류자(Packet Classifier)에게 플로우를 구별하기 위한 정보, 즉 송신자의 주소/포트, 수신자의 주소/포트, 프로토콜 ID를 세팅한다. 패킷들이 응용으로부터 생성되어 커널에 내려오면 패킷 분류자는 플로우에 대한 테이블에서 정보를 얻어 패킷이 부하제어형 플로우의 패킷인지 최선형 플로우의 패킷인지를 구별하여 적당한 큐에 분류하여 넣는다. 분류된 패킷은 큐에 세팅된 전송률에 맞추어져 패킷 스케줄러(Packet Scheduler)에 의해 전송된다.

부하제어형 큐를 생성하고 각 플로우를 구별하는 패킷 분류자들은 커널 내의 프로그램이기 때문에 응용 프로그램과 커널 사이의 인터페이스가 필요하다. 따라서 이런 정보를 커널에 넘겨주기 위한 시스템 콜 *setsockopt()*를 사용하여 구현하였다.

5. 실험 결과

본 논문에서 제시하는 EBM의 동작 검증을 위해서 FreeBSD2.2.8 PC들과 Solaris2.5.1 UNIX 워크스테이션을 10Mbps 허브로 연결하여 실험하였다. UNIX 워크스테이션을 우선 순위가 가장 높게 해서 EBM으로 선출되도록 하였으며 트래픽의 생성 및 수신을 위해서 MGEN/DREC3.1[17] 툴을 사용하였다. 그림 13에서 예약형 트래픽을 3Mbps로 고정시키고 최선형 트래픽을 4Mbps로 생성시키는 경우에 이 두 트래픽의 합이 6Mbps를 초과하므로 최선형 트래픽의 양을 반으로 줄이고 다음 flag 파라미터가 세팅되어 올 때까지 최선형 트래픽의 양을 지수적으로 증가시키는 slow increase 메커니즘의 구현 결과를 볼 수 있다.

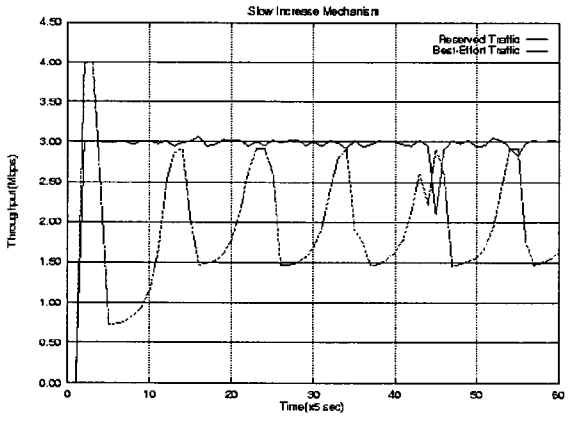


그림 13 Slow Increase 메커니즘

일반적인 이더넷 환경인 비제어형 트래픽들 즉, 호스트내의 큐에 대한 제어나 EBM을 사용하지 않는 경우에 최대 처리율은 그림 14와 같다. 이 결과는 각 호스트에서 1,024 바이트의 패킷을 1Mbps로 생성하는 호스트의 수를 하나씩 증가시키면서 측정한 결과로 최대 처리율이 9Mbps정도로 나타났다.그림 15에서는 EBM이 서브넷 내에서 최대 2Mbps만으로 전송하도록 제어하는 경우로 각 호스트에서 MGEN을 이용하여 트래픽을 5Mbps까지 생성하도록 한 결과 서브넷의 처리율이 2Mbps까지로 제어됨을 볼 수 있다. 그림 16에서는

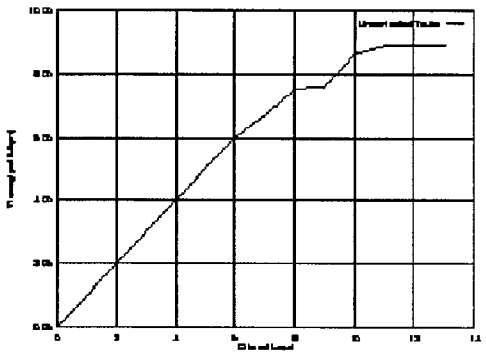


그림 14 비제어형 트래픽의 최대 처리율

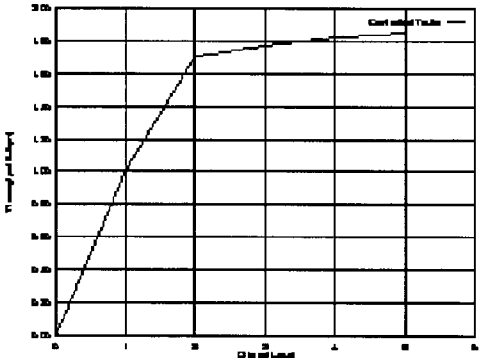


그림 15 최대 2Mbps로 제어되는 제어형 트래픽의 최대 처리율

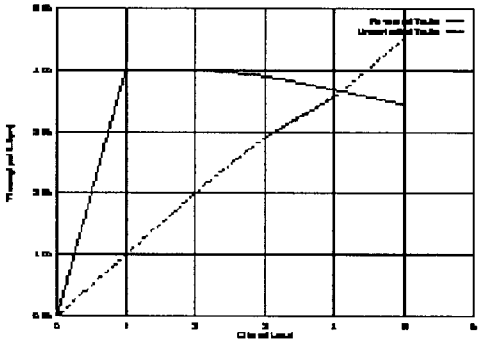


그림 16 예약형 트래픽과 비제어형 트래픽의 최대 처리율

EBM에 의해서 4Mbps로 대역폭을 고정시킨 예약형 트래픽과 비제어형 트래픽이 같이 존재하는 경우로 비제어형 트래픽이 2Mbps를 초과할 때 예약형 트래픽과 비제어형 트래픽의 대역폭 합이 6Mbps를 초과하므로 예약형 트래픽의 대역폭이 4Mbps를 유지하지 못하고 감소하는 현상을 볼 수 있다. 그림 17에서는 이 논문에서

제안하는 대역폭 관리 메커니즘에 의한 결과로써, 예약형 트래픽이 4Mbps의 대역폭을 할당받아서 사용하는 경우에 다른 트래픽들의 최대 처리량을 2Mbps로 제한함으로써 예약형 트래픽의 대역폭을 4Mbps로 유지할 수 있음을 보여주는 결과이다.

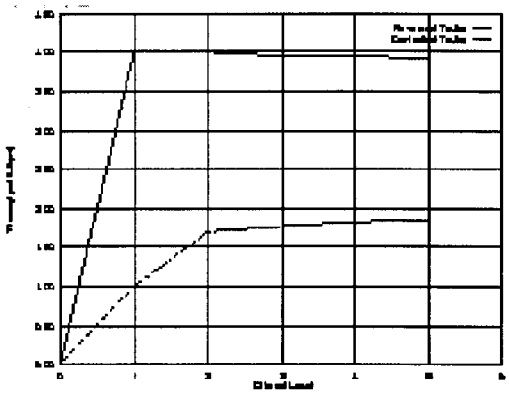


그림 17 예약형 트래픽과 제어형 트래픽의 최대 처리율

6. 결론 및 향후 연구 과제

인터넷에서 플로우에게 QoS를 제공하기 위해서는 IP 계층에서의 시그널링도 중요하지만 데이타 링크 계층에서의 QoS 제공 방법도 중요하다. 이를 위해서 IETF에서는 자원 예약을 위한 시그널링 프로토콜로 RSVP를 제시하였으며, 새로운 서비스 모델로 통합서비스 모델을 제시하였다. 그러나 RSVP와 통합서비스 모델은 여러 가지 기술로 구현될 수 있는 링크 계층과는 독립적으로 정의되어 있으므로 실제 종단간의 통합서비스 제공을 위해서는 실제 링크 계층으로 매핑 될 수 있는 자원 예약 방법이 부가적으로 필요하다.

IETF에서는 RSVP를 IEEE802 형태의 망으로 RSVP를 매핑 시키기 위해서 SBM을 제시하였으나 RSVP와 항상 같이 사용되어야 한다는 제한점과 메커니즘상의 여러 가지 문제점을 가지고 있다. 이를 해결하기 위해서 본 논문에서는 새로운 대역폭 관리 메커니즘 및 이더넷상에서 이를 담당하는 EBM을 제안하였으며 EBM의 구현내용 및 동작결과에 대해서 소개하였다.

EBM의 구현을 위해서 이더넷의 특성을 시뮬레이션 하고 이 결과를 이용해서 기존의 최선형 및 부하제어형 서비스의 제공 가능성을 보였고, 이를 이용한 EBM의 대역폭 관리 방법을 개발하였다. EBM의 동작을 검증하기 위해서 실제 이더넷 환경에서의 실험결과를 통해서 특정 플로우에게 통합서비스 모델 중에서 부하제어형 서비스의 제공 가능성을 보였다.

현재 EBM은 한 서브넷내에서 대역폭을 관리하는 동작 방식을 취하고 있기 때문에 여러개의 서브넷을 거치는 플로우에 대해서 특정 서비스를 제공하기 위해서는 여러 서브넷에 존재하는 EBM들간의 시그널링 프로토콜로의 확장이 필요하며, 인접해 있는 두 서브넷 사이에 사용할 수 있는 대역폭에 차이가 있을 때, 이 둘 간의 대역폭 매핑 방법도 추후 진행되어야 할 연구 과제이다.

참 고 문 헌

[1] R. Braden, D. Clark, and S. Shenker, "Integrated Services in the Internet Architecture: an Overview," RFC1633

[2] R. Braden, L. Zhang, S. Berson, S. Herzog, and S. Jamin, "Resource ReServation Protocol(RSVP) - Version 1 Functional Specification," RFC2205

[3] R. Yabatkar, D. Hoffman, Y. Bernet, and F. Baker, "SBM(Subnet Bandwidth Manager) : Protocol for RSVP-based Admission Control over IEEE 802-style networks," RFC2814

[4] A. Ghanwani, J. W. Pace, V. Srinivasan, A. Smith, and M. Seaman, "A Framework for Integrated Services Over Shared and Switched IEEE802 LAN Technologies," RFC2816

[5] M. Seaman, A. Smith, E. Crawley, and J. Wroclawski, "Integrated Service Mappings on IEEE 802 networks," RFC2815

[6] D. Feldmeier, "A Framework of Architectural Concepts for High Speed Communication Systems," Computer Communication Research Group, Bellcore, 1993

[7] R. Govindan and D. Anderson, "Scheduling and IPC Mechanisms for Continuous Media," 13th ACM Symposium on Operating Systems Principles (SIGOPS), 1991

[8] G. Coulson, A. Campbell, and P. Robin, "Design of QoS Controlled ATM Based Communication System in Chorus," IEEE JSAC, 1995

[9] G. Chesson, "XTP/PE Overview," 13th Conference on Local Computer Networks, 1988

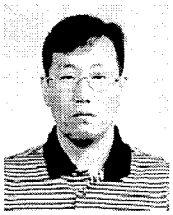
[10] D. Clark, M. Lambert, and L. Zhang, "NETBELT: A High Throughput Transport Protocol," Computer Communications Review, 17(5), 1987

[11] A. Danthine, Y. Baguette, G. Leduc, and L. Leonard, "The OSI 95 Connection-Mode Transport Service - Enhanced QoS," 4th IFIP Conference on High Performance Networking, 1992

[12] A. Campbell, "A Quality of Service Architecture," PhD thesis, Lancaster University, 1996

[13] S. Shenker, C. Partridge, and R. Guerin, "Specification of Guaranteed Quality of Service," RFC2212

[14] J. Wroclawski, "Specification of the Controlled-Load Network Element Service," RFC2211
[15] ANSI/IEEE Standard 802.3, "Carrier Sense Multiple Access with Collision Detection (CSMA/CD) Access Method and Physical Layer Specifications"
[16] UCB/LBNL/VINT Network Simulator - ns (Version 2)
[17] MGEN User's Guide



김 병 식

1993년 2월 충남대학교 컴퓨터공학과(공학사). 1995년 2월 충남대학교 컴퓨터공학과(공학석사). 1997년 8월 ~ 1998년 7월 미국 National Institute of Standards and Technology, Guest Researcher. 1999년 2월 충남대학교 박사과정 수료. 관심분야는 프로토콜 시험, 인터넷 QoS, MPLS망.

이 재 국

1997년 2월 배재대학교 전자계산학과(공학사). 2000년 2월 충남대학교 컴퓨터공학과(공학석사). 현재 (주)디지털 라이프 근무. 관심분야는 인터넷 QoS, MPLS망

전 우 직

1982년 2월 서울대학교 컴퓨터공학과(공학사). 1984년 2월 서울대학교 컴퓨터공학과(공학석사). 1984년 ~ 1987년 한국전자통신연구원 데이타통신연구실 1992년 5월 Univ. of Delaware 전산학과(공학박사). 1992년 ~ 1993년 한국전자통신연구원 정보통신표준센터. 1997년 8월 ~ 1998년 7월 미국 National Institute of Standards and Technology, Guest Researcher. 1993년 ~ 현재 충남대학교 컴퓨터공학과 부교수. 관심분야는 프로토콜 시험, 인터넷 QoS, MPLS망