

MPLS망을 이용한 인터넷 멀티캐스트 라우팅 프로토콜 모델

김 영 준[†]

요 약

본 논문은 MPLS(Multiprotocol Label Switching) 망에서 인터넷 멀티캐스트 라우팅 프로토콜을 지원하기 위한 방안을 제시한다. 인터넷 멀티캐스트 라우팅 프로토콜은 트리의 구성방식과 속성에 따라 최단거리트리 방식, 공유트리 방식 및 복합트리 방식 등으로 분류된다. 각 방식에 따라 MPLS 망에서 지원해야 하는 기능이 달라진다. 특히, 단/양방향성 트리 링크, Flooding/Prune 방식의 트리 구성, 하나의 멀티캐스트 그룹에 다중 트리의 존재 등 여러 가지 미해결되거나 해결하기 어려운 문제점들이 존재한다. 본 논문에서는 몇가지 가정을 세우고, 이를 바탕으로 각 멀티캐스트 라우팅 프로토콜을 MPLS 망에서 지원하기 위한 방안을 제시한다. Flooding/Prune 방식의 멀티캐스트 라우팅 프로토콜에 대해서는 새로이 브로드캐스팅용의 레이블을 정의하여 해결하고, 공유트리 방식에 대해서는 기 제안된 블록기반 멀티캐스트 전송기법을 도입하여 셀 인터리빙(Cell Interleaving) 문제를 해결하며, 마지막으로 복합트리 방식에 대해서는 기존 멀티캐스트 라우팅 프로토콜에 수정된 SPT 전환 기능을 추가하므로써 MPLS 지원문제를 해결한다. 제안하는 방법이 기존의 방법보다 비교해 보았을 때 전송지연에서 우수한 성능을 보임을 모의 실험을 통해서 확인하였다.

Internet Multicast Routing Protocol Model using MPLS Networks

Young Jun Kim[†]

ABSTRACT

This paper describes the new method for Internet multicast routing protocols using MPLS (Multiprotocol Label Switching) networks. Internet multicast routing protocols are divided into three categories in terms of tree types and tree characteristics : a shortest path tree, a shared tree and hybrid tree types. MPLS should support various multicast mechanisms because of extremely different IP multicast architectures, such as uni-/bi-directional link, Flooding/Prune tree maintenance mechanism, the existence of different tree types with the same group, etc. There are so many problems over MPLS multicast that the solutions can't be easily figured out. In this paper, we make a few assumptions on which the solutions of IP multicast routing protocols over MPLS networks are given. A broadcasting label is defined for the shortest path tree types. Cell interleaving problems of the shared tree types is solved by using block-based transmission mechanism. Finally, the existing hybrid-type multicast routing protocol is reasonably modified Shortest path tree type to support MPLS multicast. It has been shown that these modifications give better performance (transmission delay) than the original method.

키워드 : MPLS, DVMRP, CBT, PIM-SM, QOS

1. 서 론

현재 인터넷에서의 데이터 트래픽을 살펴보면 TCP/IP (Transport Control Protocol/Internet Protocol)를 근간으로 하는 서비스가 날로 급증하고 있으며, 앞으로의 데이터 통신은 인터넷을 기반으로 발전해 나갈 전망이다. 일반 인터넷 사용자 뿐만 아니라, 기존 ISP(Internet Service Provider) 들도 인터넷을 통해 실생활과 연계되는 E-Commerce, VOD (Video on Demand), 화상회의 등의 고품질 멀티미디어 서비스, 멀티캐스트 서비스와 같은 새로운 형태의 서비스를 요구하게 되었다[1, 2]. 이에 따라 네트워크상의 대역폭, 전송속

도, 확장성, 보안성 등 여러 방면에 걸쳐 기존 인터넷 서비스 보다 더 향상된 네트워크 품질이 필요로 되었다. 이런 경향들로 인해 공급 서비스의 양적인 팽창뿐만이 아닌 질적인 향상이 요구되자 이의 해결을 위하여 가입자망의 중설보다는 기간망의 품질을 높이기 위한 ATM(Asynchronous Transfer Mode)망, 멀티 Gigabit Ethernet망, 또는 DWDM(Dense Wavelength Division Multiplexing)망 등의 고속 고품질 기간정보통신망이 도입되고 있다. 이러한 기간망의 지원으로 새로운 통신서비스 요구를 충족시키며 너무나도 다양하게 진화되고 있는 네트워크 상에서의 문제점들을 해결하기 위한 기술로서 MPLS가 개발되고 있다[3].

MPLS는 IP를 기반으로 하는 인터넷을 좀 더 효율적으로 제공하기 위하여 계층 2에서의 스위칭 기술과 계층 3의

[†] 정 회 원 : 인하공업전문대학 정보통신과 교수
논문접수 : 2002년 7월 29일, 심사완료 : 2002년 10월 28일

라우팅 기술을 접목시키고자 제안되었다[4]. 현재 단방향전송(Unicast)에 대해서는 기본적인 논의를 끝내고 몇몇 사용자들을 통하여 실험적으로 운용되고 있다. 그러나, 이제 세계적인 MPLS의 연구 동향은 IP의 고질적인 문제인 QoS와 멀티캐스트에 대한 지원문제로 선회하게 되었고, 기본적인 내용부터 아직까지 미해결인 이슈까지 그 해결 방안을 고려하고 있다[5].

멀티캐스트란 네트워크에 분산되어 있는 모든 그룹 멤버들에게 같은 데이터를 효율적으로 전송하기 위한 네트워크 기능이다. 인터넷 멀티캐스트 그룹은 IP 주소 D 클래스 영역 중 하나의 주소로 식별되며, 그룹에 가입하려는 호스트는 IGMP(Internet Group Management Protocol)를 사용하여 원하는 그룹에 가입하거나 탈퇴한다[6]. 송신자는 모든 수신자에 대해 패킷의 복사본을 개별적으로 전송하지 않고 하나의 패킷만을 전송한다. 송신자가 전송한 패킷을 받은 멀티캐스트 라우터는 송신자에 의해 전송된 데이터를 멀티캐스트 라우팅 프로토콜에 의해 생성된 멀티캐스트 전송 트리를 따라 필요한 인터페이스로만 복사 및 전송하여 그룹의 모든 수신자에게 전송한다[7].

멀티캐스트 라우팅 프로토콜은 멀티캐스트 패킷이 전송될 수 있도록 멀티캐스트 라우터 간에 논리적 트리를 구성해 주는 프로토콜이다. 멀티캐스트 라우팅 프로토콜은 트리 구성 방식에 따라 최단거리트리(SPT: Shortest Path Tree) 방식과 공유 트리(ShT: Shared Tree) 방식으로 나눌 수 있다. 최단거리트리 방식은 각 그룹에 대해 각 송신자에서 시작하여 그룹 멤버들까지의 최단 경로를 갖는 트리를 구성하며, 대표적인 프로토콜로는 DVMRP(Distance Vector Multicast Routing Protocol)와 MOSPF(Multicast Open Shortest Path Tree) 및 PIM-DM(Protocol Independent Multicast-Dense Mode)이 있다[8-10]. 공유 트리 방식은 코어에서 시작되어 그룹 멤버들로의 최단 경로를 이루는 트리를 구성하며, 그룹 주소를 갖는 모든 데이터는 데이터의 송신자와 상관없이 트리를 공유하게 된다. 이 방식에는 현재 CBT(Core Based Tree)와 PIM-SM(Protocol Independent Multicast-Sparse Mode)가 포함된다[11, 12]. CBT와 PIM-SM은 공유트리방식이면서도 서로 다른 특징적 요소를 갖는다. CBT는 양방향 트리를 가지며, 이 트리를 통해 수신자들은 동시에 송신자의 역할을 수행할 수 있다. 그러나, PIM-SM의 경우에는 단방향 트리를 만들기 때문에 수신자는 오직 중앙의 RP(Rendezvous Point)로부터 멀티캐스트 패킷을 받기만 하며, 패킷을 전송하려면 반드시 RP까지 터널링을 해야 한다. 또한, CBT의 문제점으로 인식되는 코어와 공유트리에의 트래픽집중문제를 PIM-SM에서는 높은 비트율로 트래픽을 발생시키는 소스로부터 독립적인 트리를 구성해 낸다. 따라서, PIM-SM은 복합트리방식이라 할 수 있다.

MPLS 계층에서 효과적인 인터넷 멀티캐스트 지원 방안

을 새로이 제안한다. 인터넷 멀티캐스트 라우팅 프로토콜은 트리의 구성방식과 속성에 따라 최단거리트리 방식, 공유트리 방식 및 복합트리 방식 등으로 분류된다. 각 방식에 따라 MPLS 망에서 지원해야 하는 기능이 달라진다.

단/양방향성 트리 링크, Flooding/Prune 방식의 트리 구성, 하나의 멀티캐스트 그룹에 다중 트리의 존재 등 여러 가지 미해결이거나 해결하기 어려운 문제점들이 존재한다.

본 논문은 2장에서는 기존 인터넷 멀티캐스트 라우팅 프로토콜에 MPLS를 적용하는 방안을 제안하며, 3장에서는 성능평가를 하며, 4장에서는 결론을 제시한다.

2. MPLS기반의 고속 인터넷 멀티 캐스트 구조 설계

멀티캐스트에서의 MPLS를 적용하기에는 단/양방향성 트리 링크, Flooding/Prune 방식의 트리 구성, 하나의 멀티캐스트 그룹에 다중 트리의 존재 등 여러 가지 미해결이거나 해결하기 어려운 문제점들이 있다. 이에 본 논문에서는 멀티캐스트 프로토콜들 중 가장 일반적으로 이용되고 있는 DVMRP, PIM-DM, CBT, PIM-SM에 대한 해결책을 제시한다.

2.1 가 정

- ① MPLS 레이어에서 IP 레이어의 캐쉬를 업데이트 하는 컨트롤이 필요하다. 최종 LSR에서는 IP 레벨까지 올라가서 문제가 없으나 중간 LSR에서는 IP 계층의 MRC(Multicast Routing Cache)의 업데이트가 필요하게 된다.
- ② 레이블 할당은 Request driven방식을 취하도록 한다. 이 방식을 사용하게 되면 LSP 설정이 컨트롤 메시지를 통하여 제어되므로 Piggybacking이나 기존 프로토콜의 큰 수정없이 적용이 용이하다. 또한, 컨트롤 메시지를 통해 미리 설정이 된 레이블을 쓰므로 데이터 도착시 바로 레이블을 이용하는 것이 가능하게 된다. 그러나 레이블의 고갈이 우려되므로 타이머 같은 장치를 두도록 한다.
- ③ MPLS 스위칭 테이블은 양방향성이 가능하다. 단, 들어온 인터페이스로는 나가지 않도록 한다. 또한, 양방향성으로 설정할 프로토콜은 CBT, PIM-SM등이다.
- ④ MFC(Multicast Forwarding Cache)에 기록되는 내용은(GID, iif, oif, prune_timer)이다. GID는 그룹주소, iif는 입력인터페이스, oif는 출력인터페이스, prune_timer는 레이블 설정에 따른 시간 값을 기록하게 된다.

2.2 최단거리트리 방식의 멀티캐스트 라우팅 프로토콜

2.2.1 DVMRP/PIM-DM의 MPLS 지원시 문제점

이 방식에 MPLS를 적용함에 있어서의 문제점은 어떤 그룹 주소든지 기본적으로 주기적인 Flooding/Prune 방식을 사

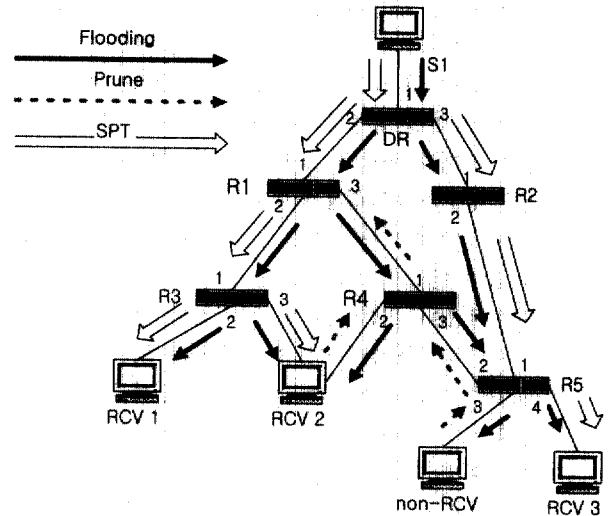
용하여 멀티캐스트 트리를 구성하게 되는데, 최초 Flooding 된 데이터를 보고 멀티캐스트 경로상에 레이블을 할당하게 된다면 Prune을 받은 인터페이스로는 레이블을 할당하지 않게 된다. 이런 설정이 끝난 이후에는 각 LSR에서 Mixed L2/L3 Forwarding이 지원되지 않는다면 레이어 2까지만 거치고 전송이 이루어지므로, Prune 타이머가 끝난 이후에 레이블이 할당되지 않은 인터페이스로 Flooding 되어야 할 데이터들이 전송되지 않으므로 레이블 할당도 되지 않는 문제점을 가지게 된다. 이를 해결하기 위해 본 제안에서는 브로드캐스팅용 레이블을 새로이 추가하고 이것을 DVMRP의 기본 레이블로 설정하여 데이터 도착과 동시에 이 기본 레이블을 할당하고 Prune을 받지않은 인터페이스에 대해서만 downstream에서의 레이블 분배 방식을 사용하여 LSP를 설정하도록 한다.

2.2.2 해결 방안

(그림 1)과 <표 1> 및 <표 2>에 해결방안으로 제시한 방안에 대한 네트워크 모델 및 상태정보가 도시되어 있다. 제안하는 방안은 다음과 같다.

- ① DVMRP는 Flooding/Prune의 동작으로 이루어지므로 첫 데이터가 발생시 DR에 도착하면 IP 레이어에서 그룹주소를 확인하고 브로드캐스팅용 레이블을 할당하여 인접한 각 LSR에 브로드캐스팅 되게 된다(레이블 값 중 4~15까지는 차후 기본적 레이블 사용을 위해 남겨져 있으므로 이 중 하나의 값을 브로드캐스팅용으로 할당하도록 한다[13]).
- ② 이 때, 브로드캐스팅용 레이블을 통해 패킷을 전달 받은 LSR에서는 IP 레이어에서 그룹 주소를 확인하고 하위 인접 LSR이 있다면 다시 다음 인접 LSR로 브로드캐스팅용 레이블을 통하여 브로드캐스팅을 하게 되며 모든 하위 LSR에서 일정시간 동안 Prune을 받지 않으면 수신측이 하나라도 있는 것으로 간주하여 LDP를 발생시키게 된다. 그러나 모든 인접 LSR이나 인접 수신측으로부터 Prune 메시지를 받게되면, 상위로 Prune 메시지를 보내게 되지만 레이블 할당용 컨트롤은 필요가 없게 된다.
- ③ 이와 같은 과정을 하위 LSR로 반복하며 최종 수신 측까지 도달하게 된다. 또한, 2의 과정에서 레이블 할당 메시지를 받은 상위 LSR은 그에 따른 레이블 설정을 하도록 하여 이후, LDP를 통하여 설정된 레이블을 이용하여 전송을 하게 되며, Prune 메시지를 받은 LSR은 MFC에 해당 인터페이스와 브로드캐스팅용 레이블, 그리고 Prune 설정시간을 Prune 타이머에 기록하도록 한다.
- ④ 일정 시간이 지난 후, Prune 타이머가 종료되면 다시 해당 브로드캐스팅 레이블을 통하여 브로드캐스팅을 하게 되며 위의 과정을 반복하게 된다. 또한, Prune 타이머

가 끝나지 않았지만 Prune을 받았던 인터페이스로부터 그룹에 가입하겠다는 Graft 메시지를 받는 경우는 레이블 할당 LDP가 동시에 이루어지도록 하여 앞에서 서술한 레이블 설정과정을 통하여 설정된 레이블을 사용한다.



(그림 1) DVMRP방에서의 MPLS 적용예

<표 1> Flooding시 브로드캐스팅 레이블 할당을 통한 LSP 설정

	(ilable, iif)	(olable, oif, Prune Timer)
R1	(4,1)	(4,2, 현재시간), (4,3, 현재시간)
R2	(4,1)	(4,2, 현재시간)
R3	(4,1)	(4,2, 현재시간), (4,3, 현재시간)
R4	(4,1)	(4,2, 현재시간), (4,3, 현재시간)
R5	(4,1), (4,2)	(4,3, 현재시간), (4,4, 현재시간)

<표 2> Prune을 받은 이후 레이블 할당을 통한 각 라우터의 LSP 설정

	(ilable, iif)	(olable, oif, Prune Timer)
R1	(10,1)	(11,2, N), (4,3, Prune 시간)
R2	(10,1)	(10,2, N)
R3	(11,1)	(N,2, N), (N,3, N)
R4	(4,1)	(N,2, Prune 시간), (4,3, Prune 시간)
R5	(4,1), (10,2)	(N,3, Prune 시간), (N,4, N)

주) N : None, ilable/olable : 입/출력 레이블, iif/oif : 입/출력 인터페이스

2.3 공유트리방식의 멀티캐스트 라우팅 프로토콜

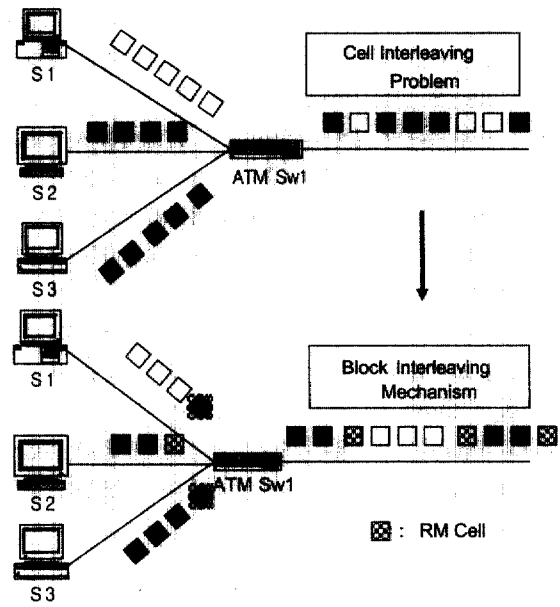
2.3.1 CBT의 MPLS 지원시 문제점

CBT에 MPLS를 적용하는 데에 있어서의 문제는 레이블 융합의 문제이다. 이는 MPLS의 단방향 전송과 멀티캐스트 내에서의 문제이기도 하므로 본 논문에서는 ATM의 경우 RM 셀을 이용한 블록 단위 전송에 의한 해결 방안만 언급하도록 하겠다. CBT의 기본동작은 그룹의 각 멤버들은 코어로의 Join 메시지를 통하여 그룹에 참여하게 되며 각 라우터에서 컨트롤 메시지에 따라 양방향성 멀티캐스트 경로를 설정하게 되므로 본 제안에서는 경로 설정을 위

한 컨트롤 메시지를 통하여 MPLS 경로를 설정하게 된다.

2.3.2 해결 방안

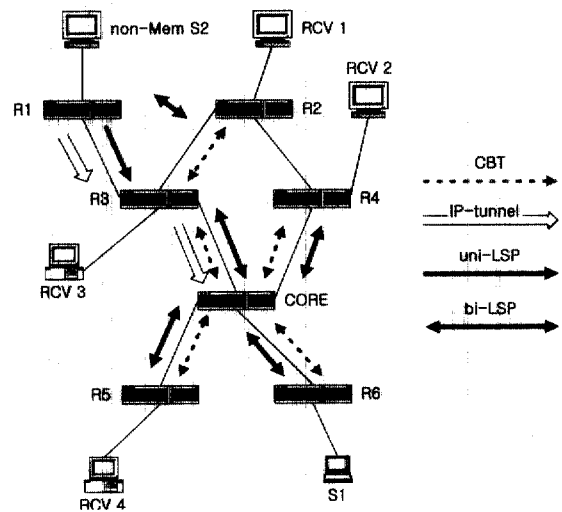
- ① 수신 측으로부터 Join 메시지를 발생시 이 메시지를 보고 LDP를 보내게 하는 Request Driven 방식을 사용하여 레이블을 할당하도록 한다. 이 때 MFC에 기록을 할 때 처음부터 Pruning 타이머영역에 설정시간 값을 넣도록 하며 이 영역에 기록된 값은 Prune 설정시간의 의미가 아닌 할당된 멀티캐스트 레이블 사용 유효시간의 의미로 쓰인다. 또한, 그룹의 멤버로의 경로를 설정할 경우는 양방향성 레이블을 설정하도록 한다. 즉, 상향과 하향에 모두 레이블을 설정하도록 하며 이 레이블 값이 반드시 같을 필요는 없다.
- ② 그룹의 비멤버 송신단인 경우 IP 상의 터널링을 통한 코어로의 경로를 설정하게 되는데 이를 MPLS 레이블 터널로써 설정하도록 한다. 즉, 1과 동일하게 설정되 양방향성을 제거한 형태가 되는 것이다. 이 문제에 대해서는 아직 단방향 전송쪽에서도 논의 되고 있으므로 본 논문에서는 논의하지 않기로 하겠다. 그러나 MPLS 단방향 경로로 설정을 하게 되면 반드시 코어를 거쳐서 돌아가야 하는 불합리성이 해결될 수 있기에 이 방식을 제안한다.
- ③ Join 메시지를 통하여 코어로의 경로를 최소경로로 찾아 가면서 각 LSR 단위로 양방향성 레이블을 할당하게 되며, 각 LSR에서는 어디서든 융합이 가능하다고 가정한다. MPLS 단방향이나 멀티캐스트에서 아직 FR이나 ATM에서의 융합에 따른 CIP(Cell Interleaving Problem)문제가 논의되고 있으나 ATM에서의 RM 셀을 이용한 블록단위의 전송방법을 쓰게 되면 이전에 논의되어 오던 VP(Virtual Path) Merging이나 VC(Virtual Channel) Merging에서 제기된 자원낭비의 문제를 효과적으로 해결할 수 있으므로 이를 참조한다 [14]. 즉, (그림 2)와 같이 각 경로를 통해 셀 단위로 나누어져 들어온 여러 개의 패킷들이 하나의 출력경로를 통해 융합되어 나가게 될 때 각 셀들이 뒤섞여 전송된 이후 다시 조합이 힘들어지는 문제를 CIP라고 하는데 이를 각 패킷 단위로 RM 셀을 붙여 블록단위로 구분지어서 전송함으로써 해결한다.
- ④ 이 후 데이터가 올 때 할당된 레이블을 쓰도록 하며 동시에 Prune 영역에 현재의 시간을 넣도록 한다. 할당된 레이블을 사용하여 전송할때에 기존 CBT에서는 반드시 코어까지 전송된 이후 각 수신 측에게로 전송되어 왔으나, 각 LSR의 MPLS 레이어에서 작성된 MFC를 통하여 중간 노드상에 수신 측으로 가는 경로가 있다면 그 출력 경로가 입력 인터페이스와 같지않다면 전송토록 한다. 그러나, 코어로 도달 전에 모든 수신 측들에게 전송되어진다 할지라도 반드시 코어까지는 데이터가 가도



(그림 2) CIP와 RM 셀을 이용한 블록 전송 방안

록 하여 그룹에 대한 정보가 유지되도록 한다. 예로 (그림 3)의 R3의 경우를 보면 IP 터널과 CBT 경로가 공존하는 상황이므로 코어로 경유하여 다시 전송되어져 오는 현상이 생기지만, 그림과 같이 설정된 LSP를 이용하게 되면 그러한 불합리한 점을 고칠 수 있게 된다.

⑤ 레이블의 고갈을 막기 위하여 flush_tree, quit_notification 등의 메시지가 오는 경우는 레이블을 제거하고, 레이블 고갈시에는 pruning 타이머 영역의 값을 현재의 시간과 비교하여 오래된 레이블을 제거하여 새로 설정하는 곳에 쓰도록 한다. 즉, 데이터가 들어 오는 동안은 pruning 타이머 영역에 데이터 입력시간을 적어 놓기 때문에 그 곳의 값이 오래 되었다는 것은 그 소스로부터의 데이터 전송이 거의 없다는 뜻이므로 레이블을 제거하여 사용하여도 지장이 없게 된다.



(그림 3) CBT 망에서의 MPLS 적용에

2.4 복합트리 방식의 멀티캐스트 라우팅 프로토콜

2.4.1 PIM-SM의 MPLS 지원시 문제점

PIM-SM에서는 ShT와 SPT를 동시에 공유하고 있는 프로토콜로서 최초 트리 구성시에는 공유 트리를 사용하다가 일정 데이터율을 넘게 되면은 해당 송신 측으로의 최단 경로 트리 전환을 일으키게 된다. 그러나 앞에서 고찰해 본 바와 같이 트리 전환에 MPLS를 쓰기에는 부적절한 부분이 많다. 즉, 레이블에는 송신측을 표기할 만한 공간이 없다는 점과 레이블로 연결된 MPLS라는 경로자체가 각 노드 상에서 일대 일로 이루어진다는 점이다. 따라서 PIM-SM에서 MPLS를 사용하기 위해서는 PIM-SM를 재 정의의 내리거나, MPLS를 PIM-SM 경로상에 부분적으로만 쓰도록 해야 한다. 만약 MPLS를 부분적으로 쓰게되면 다음과 같은 경우가 있을 수 있다.

- L2/L3 forwarding이 형성되는 경우.
- SPT인 경우만 MPLS를 쓰는 경우, 또는 SPT의 형태로 MPLS를 쓰는 경우.

위의 방법들은 이미 MPLS Working Group에서 제안한 대안들이며, 이것을 그대로 사용하기에는 불합리한 점이 많으므로 좀 더 현실적인 대안을 모색하고 있다. 이와는 별개로 멀티캐스트 워킹 그룹내에서 PIM-SM 자체를 좀 더 현실적으로 바꾸자는 의견들이 있으며 그 내용은 다음과 같다[15].

- 양방향성 트리를 구성 (CBT와 유사).
- 코어로부터의 소스 트리 전환 컨트롤 메시지 관리(CBT와 PIM-DM의 혼재 형태).

위의 대안들을 살펴보면 PIM-SM에서의 MPLS 적용 방안이라는 것은 이상적인 PIM-SM의 특성 자체를 그대로 유지하면서 지원을 하기 위한 방안이기에 PIM-SM의 제한적인 요소가 MPLS에도 그대로 적용되고 있으며 또한 비실용적이다. 반면, PIM-SM 자체를 바꾸자는 제안은 현실적으로 적용하기 복잡한 PIM-SM의 특성들을 보다 쉬운 형태로 전환하여 다른 새로운 기능들도 적용 가능토록 꾀하는 것이기에 PIM-SM을 수정한 것에 MPLS를 지원토록 하는 것도 검토해 볼만한 가치가 있으며 멀티캐스트 워킹 그룹내의 동향으로 볼 때 오히려, 더 적합한 접근방안이라고 생각된다. 따라서, 본 논문에서는 PIM-SM 자체를 수정하는 방안을 토대로 MPLS 지원방안에 대하여 다음과 같은 전제조건을 통해 대안책을 제시해보기로 한다.

- ① 양방향성 PIM-SM를 사용하는 경우, CBT와 같이 MPLS 경로를 설정토록 하며, PIM-SM에서 극복하고자 한 CBT의 단점은 QoS 라우팅을 통하여 해결하도록 한다. 즉, 최초 멀티캐스트 트리 설정시에 최소 대역

폭, 최소 전송률, 최대 지연시간 등을 요구하여 트리 설정 이후 SPT로의 전환이 일어날 요소를 제거한다.

- ② 그러나, 데이터율이 높아져서 어떤 송신 측에서 SPT (Shortest Path Tree)로의 전환을 피한다고 하면, 결국 그 송신 측으로부터 오는 데이터는 이미 코어쪽에서 많은 대역폭을 차지하고 있다는 말이 되므로 기존의 PIM-SM와 같이 수신 측에서 전환을 요구하는 것이 아니라 코어쪽에서 송신 측 트리로의 전환을 주도하여 하나의 수신 측에서만 SPT가 생성되는 것이 아니라 그룹내의 모든 수신 측들에게 동시에 새로운 그룹주소를 가진 SPT를 생성시켜 주도록 한다.

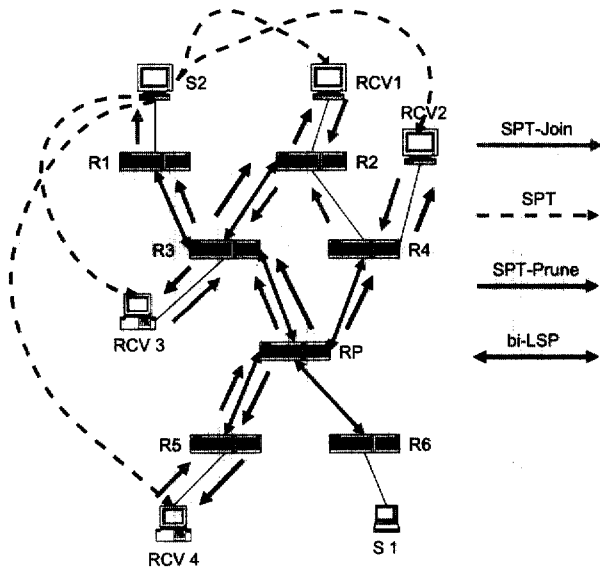
2.4.2 해결 방안

PIM-SM를 MPLS에서 지원하기 위한 기 제안된 방식들은 L2/L3 Forwarding을 요구하거나, 프로토콜 처리 과정이 매우 복잡한 단점들을 가지고 있다. 본 논문에서는 PIM-SM의 트리전환 프로토콜을 약간 수정하므로써, 보다 효율적인 MPLS 지원이 가능함을 보인다.

- ① 최초의 LSP 설정은 CBT의 경우와 동일하게 MPLS 상에서 양방향성 레이블을 설정하도록 한다. 단, 비멤버인 경우 생성되던 단방향 MPLS 터널링도 양방향성으로 대체된다. 즉, 송신 측에서 RP(Rendezvous Point)까지의 경로와 RP에서 수신측 상의 MPLS 경로가 생성된다.
- ② 문제가 되는 소스 트리로의 전환 메시지는 코어에서 트래픽을 모니터링을 하고 있다가 특정 송신측으로부터의 데이터가 일정 데이터율을 넘게되면 코어로부터 모든 PIM-SM내의 수신 측들에게 기존에 있던 SPT-Prune 메시지를 소스 트리 전환 컨트롤 메시지로 사용하여 보내지게 된다.
- ③ 이 메시지를 받은 각 LSR에서는 IP 레이어 상의 MRT에 SPT-bit을 설정하여 소스 트리 전환이 일어났음을 표시하도록 하며, 각 수신 단은 SPT-Join 메시지를 통해 수신 측으로부터 해당 송신측까지 최단 경로를 설정하며 각 노드상에 새로운 레이블을 할당시키도록 한다. SPT 경로가 설정된 이후에는 그 송신측으로부터의 데이터는 최초 설정된 그룹용 레이블이 아닌 새로이 설정된 레이블을 통하여 보내지게 된다. (그림 4)를 보면 이미 설정되어 있는 양방향성 LSP 멀티캐스트 트리에서 RCV 4로부터 SPT-Join 메시지가 오면서 S2로 소스트리가 구성이 되면 RP로부터 SPT-Prune 메시지가 각 노드상으로 멀티캐스팅되면서 다른 수신측도 S2로의 최단 경로를 설정해 나가게 된다.
- ④ ③과 같이 SPT가 생성이 되고 난 뒤 새로운 멤버가 가입하고자 할 때 Join 메시지를 통해 가지 노드로부터 레이블을 할당 받으면서 MRT에 SPT-bit가 표시되어

있다면 곧바로 그 송신측을 향해 SPT-Join 메시지를 보내어 SPT를 설정토록 한다.

- ⑤ 이상과 같이 MPLS 경로를 설정 할 경우 L2/L3 forwarding의 지원을 요구하거나 중복 전송의 문제는 발생하지 않지만 컨트롤 오버헤드나 label 의 고갈이 우려되므로 레이블 할당시 되도록 QoS 라우팅을 적용하여 경로를 설정하도록 하며, CBT와 같이 타이머를 이용하여 레이블을 관리토록 한다.



(그림 4) 수정된 PIM-SM 동작방식

3. 모의 실험 및 결과

3.1 모의 실험 환경

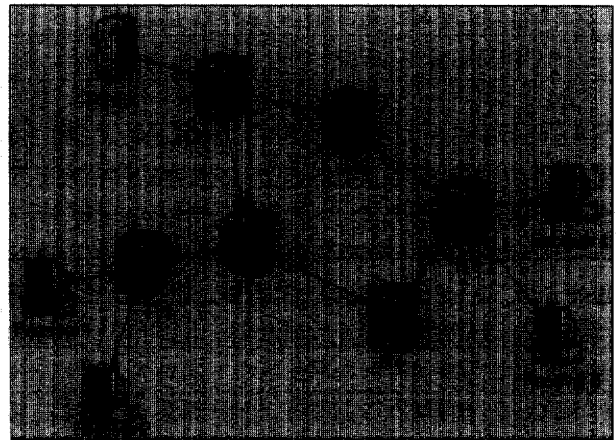
본 논문에서 제안한 방안 중에서 DVMRP, CBT 및 PIM-SM에서의 MPLS의 적용방안에 대해 시뮬레이션 툴인 MIL3 OPNET6.0(Radio Version)를 가지고 모의 실험을 수행한다[17].

3.1.1 네트워크 모델

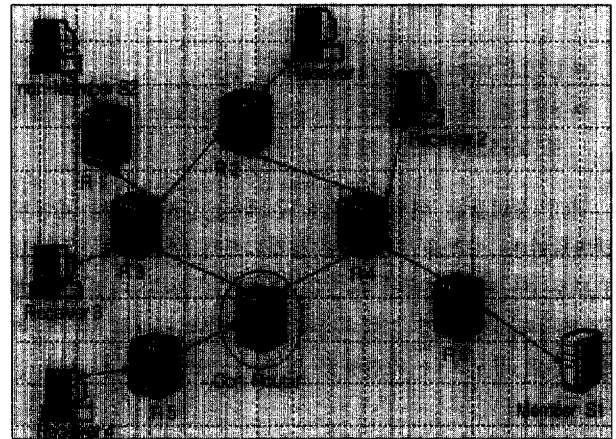
(그림 5), (그림 6) 그리고 (그림 7) 각각에 DVMRP, CBT, PIM-SM을 실험하기 위한 네트워크 모델의 OPNet 모델을 구성하였다.

(그림 6)의 CBT 네트워크 모델에는 CBT 라우팅 모듈이 탑재된 MPLS 스위치가 적용되었으며, DVMRP에서와 같은 스위치 구조를 갖도록 구성되었다. 그림 중앙의 원 안의 스위치는 코어 라우터를 의미하고 있다. 실험망 내에 구성되는 모든 멀티캐스트 트리는 코어 라우터를 중심으로 공유트리 형태로 구성된다. 소스는 그룹멤버인 S1과 비멤버인 S2가 있고 수신 전용 시스템으로는 4개가 할당되었다.

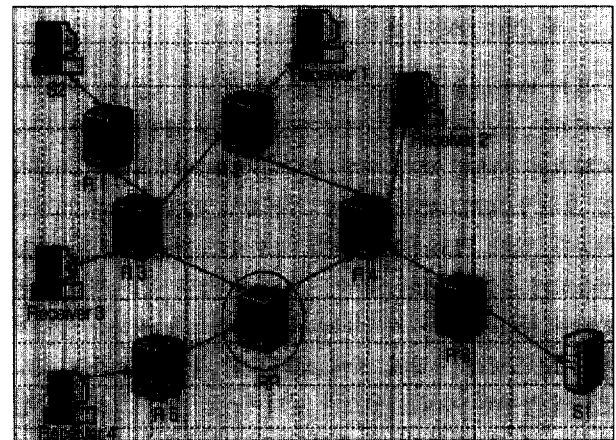
PIM-SM 네트워크 모델은 CBT 모델과 동일한 구조를 가지고 있으며, 코어 라우터 위치에 RP를 도입한다.



(그림 5) DVMRP 네트워크 모델



(그림 6) CBT 네트워크 모델

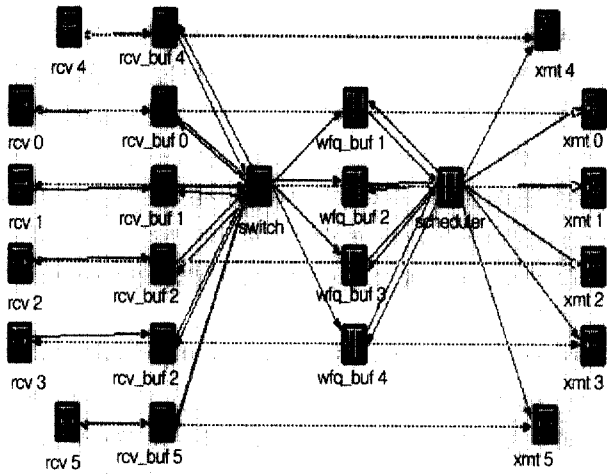


(그림 7) PIM-SM 네트워크 모델

3.1.2 노드 모델

(그림 8)과 같은 노드 모델을 구성하였다. 상단에 보이는 것이 라우터 모델로서 입력버퍼로부터 MPLS 레이어와 IP 레이어를 거쳐 4개의 클래스로 나누어진 버퍼를 WFQ(Weighted Fair Queuing)를 적용한 스케줄러가 통제하여 출력시 키도록 구성하였다. 그리고 하단의 왼쪽이 송신측 모델이며,

하단 오른쪽이 수신측 모델이 되겠다. WFQ 알고리즘은 트래픽을 QoS에 따라 몇 개의 클래스를 두고 각 클래스별로 가중치를 두어 공평하게 Queuing을 하는 알고리즘으로써 이 실험에서는 Shim Header내의 Exp비트를 사용하여 4개의 클래스로 나누어서 관리하도록 하였다.



(그림 8) 노드 모델

3.1.3 시뮬레이션 파라미터

시뮬레이션 파라미터로는 <표 3>에 나타낸다.

<표 3>에서 스위치 처리 능력은 초당 1Mbps급으로 비정상적으로 설정되어 있는데, 이는 버퍼 오버플로우를 유도하여 클래스별 서비스 제공성을 확인한다.

<표 3> 시뮬레이션 파라미터 1

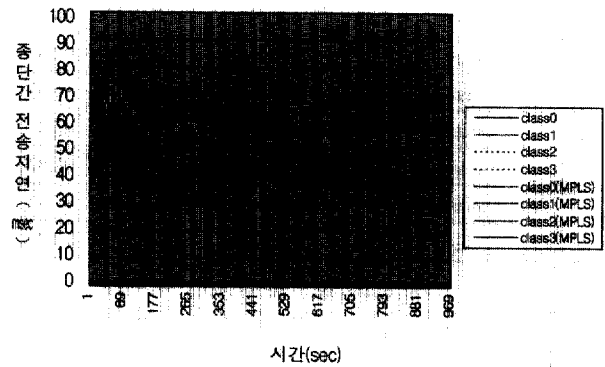
파라미터		값	
스위치 처리능력		1Mbps	
패킷 발생율		1000 packets/sec(Exp. PDF)	
버퍼 크기	일렉 버퍼	Infinity	
	클래스 버퍼	100 packets	200 packets
스위치간 평균거리		1000 Km	2000 Km
스위치 처리 지연시간		0.5 msec	1 msec

3.2 모의 실험 결과 분석

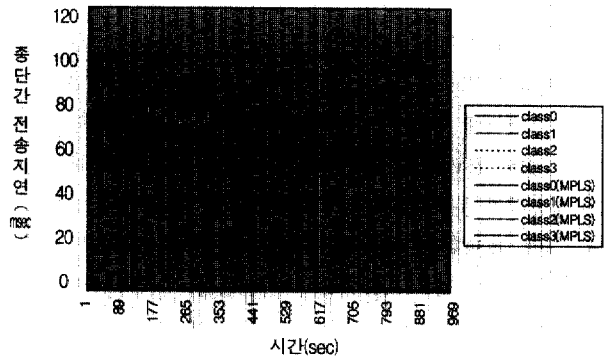
이 모의 실험에서는 제안하는 각 멀티캐스트 네트워크에 MPLS를 적용함으로써 Layer 3라우팅(IP 라우팅)으로 전송하는 경우와 비교하는데 그 목적이 있다. MPLS의 레이블을 사용한 경우와 사용하지 않고 IP 멀티캐스트 라우팅 테이블 정보를 통한 전송만을 사용한 경우에 있어서 (그림 5)~(그림 7)에서 보인 각 멀티캐스트 네트워크에 실험 파라미터를 적용한 경우 각 수신자가 경험하는 종단간 지연시간을 성능평가의 지표로 삼는다. 또한, 레이블 할당시 QoS에 따라 4개의 클래스로 나누어 Exp비트에 기록한 것을 WFQ를 구현한 버퍼를 가진 라우터에서 각 클래스별로 0.4, 0.3, 0.2, 0.1의 선형적인 가중치를 두어 관리함으로써 혼잡상황시에

QoS가 높은 상위 클래스의 데이터들이 우선적으로 전송될 수 있도록 보장 하였다.

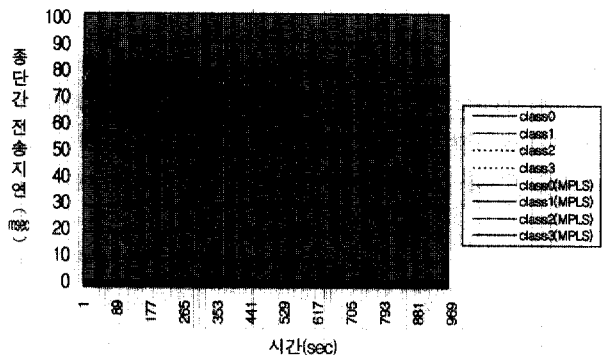
(그림 10)~(그림 12)는 각각 클래스별 평균전송지연 시간을 측정 한 것이다. 실험결과에서 나타나 있듯이 MPLS 멀티캐스트 방식을 적용할 경우에 상대적으로 작은 전송지연 시간을 보여주고 있으며 각각의 클래스별 전송지연 시간도 우선순위에 따라 높은 우선순위의 클래스가 좀 더 나은 서비스를 받고 있음을 알 수 있다.



(그림 9) DVMRP 실험결과



(그림 10) CBT 실험결과



(그림 11) PIM-SM 실험결과

<표 4>는 (그림 9)~(그림 11)을 정리한 것이며, 최종적으로 안정화 되었을 때 각 프로토콜당 종단간 전송지연시간을 나타낸다.

(그림 9)의 DVMRP 실험결과에서 보여진 결과는 스위

치 버퍼에 패킷들이 들어오기 전까지는 다시 말해서 Shim header를 이용한 블록 전송 메커니즘이므로 한 블록이 형성되기까지는 전송지연이 걸리고 그후에는 지속적인 입력으로 인해서 지속적인 블록이 형성되므로 안정을 찾게 된다.

<표 4> DVMRP, CBT, PIM-SM 실험결과 분석표

적용 계층	클래스 등급	프로토콜당 중단간 전송지연(msec)					
		DVMRP		CBT		PIM-SM	
IP	class 0	67.05	80.82	94.19	118.5	80.19	110.3
	class 1	59.36	72.24	83.75	103.7	73.06	99.45
	class 2	53.07	67.87	79.06	100.5	74.82	92.89
	class 3	53.09	65.56	76.59	95.84	64.06	84.90
MPLS	class 0	58.10	72.98	73.56	96.56	73.75	95.34
	class 1	46.56	60.79	71.59	90.23	70.59	91.87
	class 2	42.32	54.48	61.86	80.54	62.59	81.90
	class 3	39.19	49.89	58.10	75.98	56.56	73.89

MPLS를 적용한 클래스 3의 경우가 전송지연이 가장 작고, MPLS를 적용하지 않은 클래스 0이 전송지연이 가장 크다. 전체적으로 보았을 때 MPLS를 적용한 경우가 전송지연이 작다.

(그림 10)의 CBT 실험결과와 (그림 9)의 DVMRP 실험결과와 비교하면 평균적인 전송 지연이 크다. 이는 공유트리 방식에 의한 코아라우터의 집중으로 인한 결과이며, (그림 9)의 DVMRP 실험결과 경우와는 다르게 최적 경로로 전송되지 못하므로 전송지연이 크다. 예를 들어 (그림 6)의 CBT 네트워크 모델에서 비멤버 S2가 데이터를 송신하고자 할 경우 코아 라우터까지 터널링을 하여야 한다. 다시 코아 라우터에서 수신 전용시스템 3이 데이터를 받으려면, 코아 라우터를 경유하여야만 하므로 우회되는 경로가 크다. 이 경우 DVMRP가 적용되었다면, 코아 라우터까지 가지 않고 R1, R3 라우터만 경유하는 최적경로를 따라 데이터가 전송될 수 있다.

(그림 11)은 (그림 9)와 (그림 10)이 복합적으로 형성된 경우이므로 (그림 9)와 (그림 10)의 평균 전송지연 정도를 갖는다. (그림 11)의 PIM-SM 실험 결과치는 (그림 10)의

CBT 실험결과치와 같이 RP로 전송된후 포워딩되므로, 전송지연이 크다. 그러나 중간에 트리 스위칭 때문에 좀 더 전송지연을 줄일 수 있다. 예를 들어 (그림 7)의 PIM-SM 네트워크 모델의 경우에서 S1이 데이터를 송신하기 위해서는 RP까지 데이터를 보내야 하며, 이를 RP가 수신전용시스템 1이 수신되기 위해서는 R2로 조인 메시지를 보내며, 이러한 조인메시지가 RP까지 연결되어 경로를 형성하여 데이터를 수신하게 된다. 수신전용 시스템 2가 데이터를 수신할 경우 R6, R4를 경유하고 다시 RP까지 경유하여 다시 R4를 경유하여 수신전용 시스템 2로 데이터가 전송되므로, 이 경우 송신측이 R6, R4를 경유하여 수신전용 시스템 2로 가게되는 Shortest Path Tree를 형성한다. 이러한 트리 스위칭 때문에 (그림 10)의 CBT 실험결과치 보다 평균 전송지연이 작음을 알 수 있다.

IP 멀티캐스트 라우팅 정보를 이용한 전송과 레이블을 이용한 전송의 차이가 생기는 이유는 두 방법 모두 같은 경로를 따라 전송되어지지만, IP 라우팅에서는 IP Header의 많은 처리과정으로 인해서 즉 IP 패킷을 다음 홉으로 전송하기 위해서는 TTL, Checksum, longest prefix matching 방식을 이용한 라우팅 테이블 룩업시간이 소요되며, MPLS의 경우에는 exact matching 방식을 이용하여 전송되므로 전송지연이 작아진다.

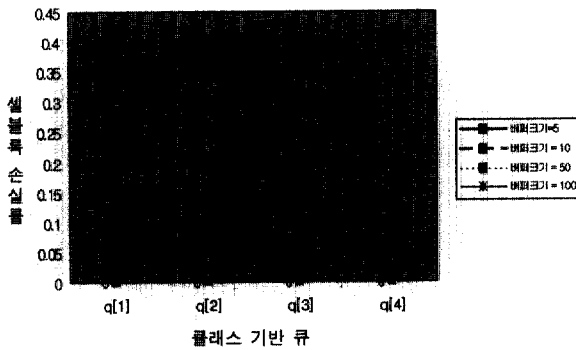
또한 각 클래스별 전송의 차이를 보이는것은 가중치를 주었을때의 각 클래스별 전송지연을 볼 수 있다. 이와 같은 결과가 나오는 이유는 모든 패킷의 크기는 일정하나 Shim Header내에 있는 Exp 비트에 표기된 클래스별로 출력버퍼에 나누어 스케줄러에서 가중치에 따라 각 클래스별 큐에 있는 셀들을 꺼내음으로써 일어난 결과이다. 이는 혼잡상황에서도 우선순위의 클래스에 있는 셀들을 먼저 전송시킴으로써 보장된 전송을 할 수 있는 것이다. 셀블록 손실율의 성능평가를 위해 미리 정한 모의실험 환경은 <표 5>와 같다. EPD 메커니즘에 의한 버퍼 관리를 위해 각 큐의 특정 한계값을 큐 버퍼용량의 80%로 정한다. 또한, 블록 스케줄링을 위해 각 큐의 서비스 가중치는 임의대로 0.4, 0.3, 0.2, 0.1로 정하여 트래픽 유형에 따른 차별화된 성능을 나타냄을 검증하고

< 표 5 > 시뮬레이션 파라미터 2

항 목	변 수 명	값 1	값 2
셀블록의 도착시간 간격	arrival_time_interval	0.001(secs/cell_block)	0.002(secs/cell_block)
출력 링크 대역폭	link_capacity	100(cell_blocks/sec)	200(cell_blocks/sec)
각 큐의 특정 한계값	q[i].threshold	80% × q[i].buffer_size	90% × q[i].buffer_size
각 큐의 서비스 가중치	q[0].weight	0.4	0.5
	q[1].weight	0.3	0.4
	q[2].weight	0.2	0.3
	q[3].weight	0.1	0.2

자 하였다.

트래픽 유형별로 셀블록 손실률을 구하였다(그림 12 참조). 셀블록의 손실률은 도착한 전체 셀블록의 수와 서브스된 셀블록 수와의 차를 전체 셀블록 수로 나누어 구한다. 전체적으로 클래스 1 유형의 트래픽이 낮은 셀블록 손실률을 보임을 알 수 있다. 유의할 점은 버퍼 크기가 증가할수록, 전체적인 셀블록 손실률이 매우 낮아진다는 사실이다. 실제로, 버퍼 크기를 100으로 하였을 때, 모든 큐의 셀블록 손실률이 '0'을 나타내었다.



(그림 12) 버퍼크기에 따른 트래픽 유형별 셀블록 손실률

이상의 결과로부터 제안한 알고리즘이 전송측면에서 각 프로토콜당 전송지연 시간이 정량적으로 개선됨을 실험결과로써 보여주었다.

4. 결 론

논문에서는 기존 인터넷 멀티캐스트 라우팅 프로토콜에 대한 개념정립과 분류를 하였고, MPLS의 개략적 내용을 제시하였다. 기존 멀티캐스트 라우팅 프로토콜은 크게 최단경로 트리 방식, 공유트리 방식 및 복합트리 방식으로 분류되며, 각각은 독특한 트리 링크 속성과 트리 유지 방식을 가지고 있다. 이는 MPLS에서 멀티캐스트를 지원하는데 있어 큰 어려움을 갖도록 한다. 이러한 문제점을 해결하는데, 본 논문에서는 Flooding/Prune 방식의 멀티캐스트 라우팅 프로토콜에 대해서는 브로드캐스팅용의 임시 레이블값을 정의하여 해결하고, 공유트리방식에 대해서는 기제안된 블록기반 멀티캐스트 전송기법을 도입하여 셀인터리빙 문제를 해결하며, 마지막으로 복합트리방식에 대해서는 기존 멀티캐스트 라우팅 프로토콜에 수정된 SPT 전환기능을 추가함으로써 MPLS 지원문제를 해결한다. 이상의 결과로부터 제안한 알고리즘이 전송측면에서 각 프로토콜당 전송지연시간이 정량적으로 개선됨을 앞서 분류한 멀티캐스트 라우팅 프로토콜에 적용하여 Opnet을 이용하여 시뮬레이션을 한 후 결과를 분석한 결과, 정량적으로 전송지연시간이 개선됨을 알수 있

었다. 향후 연구과제로는 MPLS를 통한 멀티캐스트 Qos 지원방안을 연구개발하여야 할 것이다.

참 고 문 헌

- [1] Bob Quinn and Kevin Almeroth, "IP Multicast Applications : Challenges and Solutions," <draft-ietf-mboned-mcast-apps-00.txt>, February, 1999.
- [2] Sanjoy Paul, "Multicasting on The Internet and its Applications," Kluwer Academic Publisher, May, 1998.
- [3] Eric C. Rosen, Arun Viswanathan, and Ross Callon, "Multi-protocol Label Switching Architecture," <draft-ietf-mpls-arch-06.txt>, August, 1999.
- [4] R. Callon, P. Doolan, N. Feldman, A. Fredette, G. Swallow, and A. Viswanathan, "A Framework for Multiprotocol Label Switching," <draft-ietf-mpls-framework-04.txt>, July, 1999.
- [5] D. Ooms, W. Livens, B. Sales, M. Ramalho, A. Acharya, F. Griffoul, and F. Ansari, "Framework for IP Multicast in MPLS," <draft-ietf-mpls-multicast-00.txt>, June, 1999.
- [6] W. Fenner, "Internet Group management Protocol, Version 2," IETF 2236, Nov., 1997.
- [7] K. C. Almeroth and M. H. Ammar, "Multicast Group Behavior in the Internet's Multicast Backbone (Mbone)," *IEEE Network Magazine*, pp.124-129, June, 1997.
- [8] D. Waizman, S. Deering and C. Patridge, "Distance Vector Multicast Routing Protocol," RFC1075, Nov., 1988.
- [9] J. Moy, "Multicast Extensions to OSPF," RFC1584, March, 1994.
- [10] S. E. Deering, D. Estrin, D. Farinacci, V. Jacobson, C-G Liu and L. Wei, "Protocol Independent Multicast(PIM), Dense Mode Protocol Specification," March, 1994.
- [11] A. Ballardie, "Core Based Trees (CBT) Multicast Routing Architecture," IETF RFC2201, Sept., 1997.
- [12] D. Estrin, D. Farinacci, A. Helmy, D. Thaler, S. Deering, M. Handley, V. Jacobson, and L. Wei, "Protocol Independent Multicast-Sparse Mode (PIM-SM) : Protocol Specification," RFC2362, June, 1998.
- [13] WonTae Kim and YongJin Park, "SAMA : Scalable wM-ATM Multicast Architecture for Dynamic Group Service," the proceedings of the 4th CDMA International Conference 1999 (CIC '99), Sept., 1999.
- [14] Deborah Estrin, Mark Handley, Isidor Kouvelas, and Lorenzo Vicisano, "A New Proposal for Bi-directional PIM," <draft-kouvelas-pim-bidir-new-00.txt>, October, 1999.
- [15] K. Nagami, H. Esaki, Y. Katsube, and O. Nakamura, "Flow

Aggregated, Traffic Driven Label Mapping in Label-Switching Networks," *IEEE JSAC*, Vol.17, No.6, June, 1999.

- [16] 김영준, 박용진, "MPLS망에서의 인터넷 멀티캐스트 라우팅 프로토콜 지원방안", 전자공학회논문지 TE편, 제37권 제3호, pp.93-103, 2000.
- [17] 김영준, "MPLS에서의 인터넷 멀티캐스트 라우팅 프로토콜", 한양대학교 박사학위논문, 2001.



김 영 준

e-mail : yjkim@inhac.ac.kr

1986년 한양대학교 전자공학과(학사)

1991년 한양대학교 대학원 전자공학과(석사)

1996년~2001년 해천대학 정보시스템 계열
조교수

2001년 한양대학교 대학원 전자공학과(박사)

2001년~현재 인하공업전문대학 정보통신과 조교수