

論文2000-37TE-3-14

MPLS망에서의 인터넷 멀티캐스트 라우팅 프로토콜 지원 방안

(Internet Multicast Routing Protocol Supporting Method over MPLS Networks)

金榮俊*, 朴容震**

(Young-Jun Kim and Yong-Jin Park)

요약

본 논문은 MPLS (Multiprotocol Label Switching) 망에서 인터넷 멀티캐스트 라우팅 프로토콜을 지원하기 위한 방안을 제시한다. 인터넷 멀티캐스트 라우팅 프로토콜은 트리의 구성방식과 속성에 따라 최단거리 트리 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 미해결이거나 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 기능을 추가하므로써 MPLS 지원문제를 간단히 해결한다.

Abstract

This paper describes Internet multicast routing protocols over MPLS (Multiprotocol Label Switching) networks. Internet multicast routing protocols are divided into 3 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 to support MPLS multicast.

* 正會員, 惠泉大學 정보시스템 계열

(Dept. of Schools of Information Systems Heychon College)

** 正會員, 漢陽大學校 電氣 電子 工學部

(Division of Electrical and Computer Engineering, Hanyang University)

接受日字:2000年7月5日, 수정완료일:2000年9月18日

I. 서론

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), 화상 회의 등의 고품질 멀티미디어 서비스, 멀티캐스트 서비스와 같은 새로운 형태의 서비스를 요구하게 되었다^[1,2]. 이에 따라 네트워크상의 대역폭, 전송속도, 확장성, 보안성 등 여러 방면에 걸쳐 기존 인터넷 서비스보다 더 향상된 네트워크 품질이 필요로 되었다. 이런 경향들로 인해 공급 서비스의 양적인 팽창뿐만이 아닌 질적인 향상이 요구되자 이의 해결을 위하여 기업자망의 증설보다는 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 통신 서비스 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 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 지원문제도 선회하게 되었고, 기본적인 내용부터 아직까지 미

해결인 이슈까지 그 해결 방안을 고려하고 있다^[5].

본 논문은 II장에서 기존 인터넷 멀티캐스트 라우팅 프로토콜에 대해 알아보고, III장에서는 MPLS의 기본적인 내용과 MPLS 멀티캐스트의 문제점들을 살펴본다. IV장에서는 기존 인터넷 멀티캐스트 라우팅 프로토콜에 MPLS를 적용하는 방안을 제안하며, V장에서는 결론을 제시한다.

II. 인터넷 멀티캐스트 라우팅 프로토콜

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 트

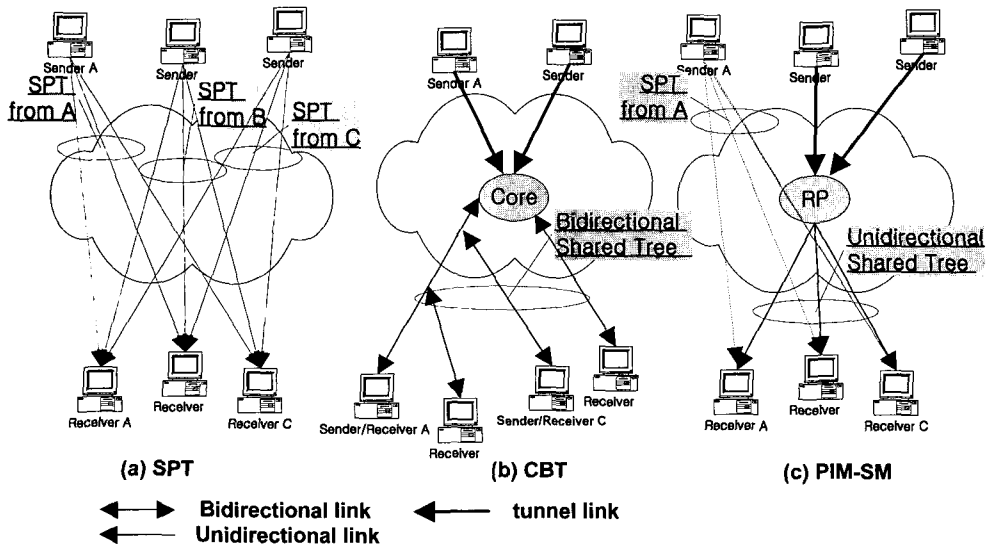


그림 1. (a) SPT (b) CBT (c) PIM-SM
 Fig. 1. (a) SPT (b) CBT (c) PIM-SM.

리를 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][9][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]}. 최단거리트리 방식과 공유트리방식의 개념적 특징을 그림 1에 도시한다. CBT와 PIM-SM는 공유트리방식이면서도 서로 다른 특징적 요소를 갖는다. CBT는 양방향성 트리를 가지며, 이 트리를 통해 수신자들은 동시에 송신자의 역할을 수행할 수 있다. 그러나, PIM-SM의 경우에는 단방향성 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 그림1에서 보이는 바와같이 높은 비트율로 트래픽을 발생시키는 소스로부터 독립적인 트리를 구성해 낸다. 따라서, PIM-SM는 복합트리방식이라 할 수 있다. 보다 자세한 내용은 III장 2절에서 다룬다.

III. MPLS 기본 구조

1. MPLS 개요

MPLS란 IP 주소의 긴 Prefix 대신 20 비트의 레이블 (Label)을 식별자로 사용함으로써 IP 계층의 라우팅 기능과 데이터 링크 레이어의 스위칭 기능을 접목시키는 효과를 가져와 빠른 전송, 보장된 전송을 할 수 있다. 즉, MPLS 도메인 내로 데이터가 들어 오게 되면 Ingress LSR (Label Switched Router)에서 기존의 데이터링크 계층의 헤더와 PDU (Payload Data Unit)사이에 1바이트의 Shim Header를 삽입하여 각 LSR의 IP 레이어와 데이터 링크 레이어 사이의 MPLS 레이어에

서 이를 통하여 경로를 설정, 전송하게 된다. 이 후 MPLS 도메인의 Egress LSR에서 빠져나가 외부로 전송하게 될 때는 패킷에서 레이블을 제거하고 전송하게 된다. Shim Header는 ATM이나 FR (Frame Relay)에서는 여유공간이 없으므로 다른 형태로 붙여지게 되는 데, ATM에서는 VPI(Virtual Path Identifier)/VCI (Virtual Channel Identifier)가, FR에서는 DLCI (Data Link Connection Identifier)가 각각 사용된다^[13].

레이블은 그림 2와 같이 20비트의 레이블과 1비트의 스택 (Stack)비트, 8비트의 TTL (Time To Live), 그리고 3비트의 Exp (Experimental) 비트로 구성되어 있다. FEC (Forwarding Equivalency Classes)란 MPLS 도메인을 지나가는 데이터 플로우를 같은 목적지 또는, 같은 서비스 등급 등의 몇 가지 기준에 의하여 몇 개의 집합으로 나누어 놓은 것이라고 볼 수 있으며 이렇게 나누어 놓은 FEC에 각각 레이블이 붙여지게 된다. 이때 레이블 값은 MPLS 도메인을 지나가는 동안 거치는 각 LSR과 LSR 간에 서로 다른 값을 가지고 설정이 되어 질 수도 있으며 이에 따라 각 LSR을 거치면서 입력 레이블과 출력 레이블 간의 레이블 교체(Label Swapping)가 이루어지게 된다.

20	3	1	8
Label	Exp	S	TTL

그림 2. Shim Header의 Frame Format
Fig. 2. Frame Format of Shim Header.

2 MPLS와 IP 멀티캐스트

본 절에서는 MPLS망에서 멀티캐스트를 제공하기 위해 기본적으로 고려해야 하는 설계규칙을 정의한다. 이는 IETF의 MPLS 멀티캐스트를 다룬 초안에 기반을 두었다.

1) 트리 구성방식에 따른 LSP설정 문제

① 최단거리트리 방식

DVMRP, PIM-DM의 경우 데이터가 도착해야지만 멀티캐스트 경로가 구성되므로 Request Driven 방식은 사용하기 힘들게 되어 Traffic Driven 방식을 사용하는 것이 제안되고 있다. 그러나 이 방식 또한, 주기적으로 Flooding/Prune과정을 거쳐야 하므로 레이블을 주기적으로 설정해야 하는 문제를 안고 있다.

② 공유트리 방식

CBT의 경우 LDP 설정에 있어서는 Request Driven 방식이 선호되지만, 트리 구성에 있어서의 문제점은 노드상의 경로가 양방향성이므로 데이터의 융합 문제가 발생하게 된다. 레이블 융합으로 인해 ATM에서 문제가 발생하게 되는데, ATM에서는 순서 번호가 없으므로 레이블 융합 이후 셀들이 뒤섞이게 되면 수신자가 수신된 셀들을 각 송신자별로 구분할 수 없게 된다. 이를 CIP(Cell InterLeaving Problem)라고 하며 이의 해결 방안으로 VP Merging, VC Merging등이 제안되고 있다.

③ 복합트리 방식 (최단거리트리+공유티리)

그림 3에서, ShT가 IJ와 PJ를 통해 만들어진 상태에서 RCV1에서 RP까지의 최단거리가 R1-R2-R5라고 할 때, RCV1의 DR (Designated Router)인 R1이 S1으로부터 일정한 이상의 데이터를 받게 되면 S1에 대한 SPT를 설정하기 위해 PJS를 보내게 된다. R2에 새로 설정된 SPT로부터 데이터를 받은 이후, R2는 R5에 S1에 대한 prune을 보내게 된다. 이렇게 해서 R2에서는 SPT와 ShT의 정보가 공존하게 되며, prune을 받아 ShT상으로 S1으로부터 오는 데이터를 전송할 필요가 없어진 R5에서도 SPT와 ShT의 정보가 마찬가지로 공존하게 된다. 만일 R5에서 SPT정보를 가지고 있지 않다면, S1으로부터의 트래픽이 다시 R2로 유입되어 중복 패킷이 RCV1에 수신되게 된다.

(S,G) RPT-bit의 역할은 SPT를 통하여 보내진 데이터가 ShT를 통하여 재전송되지 않도록 하는 기능을 가지고 있으므로 레이블 스위칭을 하더라도 이것을 구분하여 전송해야 중복 전송하는 일이 없을 것이다. 그러나, 레이블에는 송신자를 구별할 수 있는 정보공간이 없으므로 한 그룹 내에 하나의 레이블만 할당하는 방식으로서는 ShT와 SPT의 공유형태를 지원할 수 없게 되므로 다음과 같은 해결방안이 제시되고 있지만, 각각 문제점을 가진다.

첫째, SPT로의 전환 발생시 LSP(Label Switched Path)를 모두 없애고 모든 트래픽을 레이어 3에서 전송시킨다. 이 방법은 결국 PIM-SM에서 MPLS 전송을 하지 않음을 의미한다.

둘째, LSP 설정시 ShT내의 각 송신측별로 레이블을 할당한다. 즉, 공유트리용이 아닌 모든 개별적 소스를 구별하는 LSP를 생성시킨다. 따라서, 이 방법은 송신단이 많을 경우 발생될 수 있는 레이블의 고갈과 컨트롤 시그널링 오버헤드가 우려된다.

셋째, SPT만 레이블 스위칭을 시키고 ShT는 계층 3에서 전송시킨다. 이 방법은 Mixed L2/L3 Forwarding 이 지원되어야만 가능한 방법이다.

넷째, RP부터 공유정보를 필요로 하는 LSR까지 (S, G) SPT-bit state를 가진 LSP를 생성한다. 두번째 방안 에 비해 레이블의 소모는 적더라도 컨트롤 오버헤드

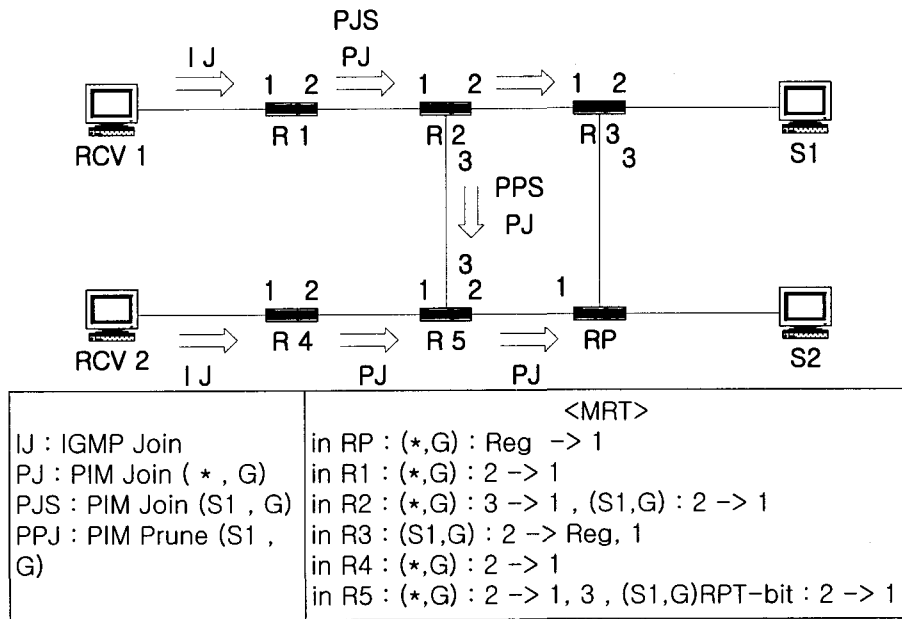


그림 3. PIM-SM망에서의 공유트리와 최단거리트리의 공존
 Fig. 3. Coexistence of Shared tree and Shortest path tree over PIM-SM Networks.

나 적용에 있어서의 어려움을 들 수 있다.
 이와 같이 제시된 대안책 모두 MPLS를 사용하면서
 얻게 될 이점보다는 프로토콜의 복잡성이 크게 증가되
 어 새로운 대안이 요구된다.

- 2) 인터넷 멀티캐스트 라우팅 프로토콜 특징 정리
- 표 1에 지금까지 기술한 내용을 요약하였다. 매우 다
 양한 트리속성과 트리 유지방식을 볼 수 있다.

IV. MPLS망에서의 멀티캐스트 라우팅 방안

멀티캐스트에서의 MPLS를 적용하기에는 위에서 언
 급한 바와 같은 여러 문제점들이 있다. 이에 본 논문에서
 는 1절에서 제시하는 가정을 기반으로 기존 멀티캐
 스트 프로토콜들 중 가장 일반적으로 이용되고 있는
 DVMRP, PIM-DM, CBT, PIM-SM에 대한 해결책을
 제시한다.

1 가정

- ① MPLS 레이어에서 IP 레이어의 캐쉬를 갱신 하는
 제어가 필요하다. 최종 LSR에서는 IP 레벨까지 올
 라가서 문제가 없으나 중간 LSR에서는 IP 계층의
 MRC(Multicast Routing Cache)의 업데이트가 필요
 하게 된다.
- ② 레이블 할당은 Request driven 방식을 취하도록 한
 다. 이 방식을 사용하게 되면 LSP 설정이 컨트롤
 메시지를 통하여 제어되므로 Piggybacking이나 기
 존 프로토콜의 큰 수정없이 적용이 용이하다. 또한,
 컨트롤 메시지를 통해 미리 설정된 레이블을 쓰므
 로 데이터 도착 시 바로 레이블을 이용하는 것이
 가능하게 된다. 그러나 레이블의 고갈이 우려되므
 로 타이머 같은 장치를 두도록 한다.
- ③ MPLS 스위칭 레이블은 양방향성이 가능하다. 단,

들어온 인터페이스로는 나가지 않도록 한다. 또한,
 양방향성으로 설정할 프로토콜은 CBT, PIM-SM등
 이다.

- ④ MFC(Multicast Forwarding Cache)에 기록되는 내
 용은 (GID, iif, oif, prune_timer)이다. GID는 그룹
 주소, iif는 입력 인터페이스, oif는 출력 인터페이스,
 prune_timer는 레이블 설정에 따른 시간값을 기록
 하게 된다.

2 최단거리트리 방식의 멀티캐스트 라우팅 프로토콜

1) DVMRP/PIM-DM의 MPLS 지원시 문제점

이 방식에 MPLS를 적용함에 있어서의 문제점은 어
 떤 그룹 주소든지 기본적으로 주기적인 Flooding/Prune
 방식을 사용하여 멀티캐스트 트리를 구성하게 되는데,
 최초 Flooding된 데이터를 보고 멀티캐스트 경로상에
 레이블을 할당하게 된다면 Prune을 받은 인터페이스로
 는 레이블을 할당하지 않게 된다. 이런 설정이 끝난 이
 후에는 각 LSR에서 Mixed L2/L3 Forwarding이 지원
 되지 않는다면 레이어2까지만 거치고 전송이 이루어지
 므로, Prune 타이머가 끝난 이후에 레이블이 할당되지
 않은 인터페이스로 Flooding 되어야 할 데이터들이 전
 송되지 않으므로 레이블 할당이 되지 않는 문제점을
 가지게 된다. 이를 해결하기 위해 본 제안에서는 브로
 드캐스팅용 레이블을 새로이 추가하고 이것을 DVMRP
 의 기본 레이블로 설정하여 데이터 도착과 동시에 이
 기본 레이블을 할당하고 Prune을 받지 않은 인터페이
 스에 대해서만 downstream에서의 레이블 분배 방식을
 사용하여 LSP를 설정하도록 한다.

2) 해결 방안

그림 4와 표 2 및 표 3에서 해결방안으로 제시한 방
 안에 대한 네트워크 모델 및 상태정보가 도시되어 있
 다. 제안하는 방안은 다음과 같다.

표 1. 멀티캐스트 프로토콜 비교
 Table 1. Comparison of Multicast Protocol.

	DVMRP	PIM-DM	CBT	PIM-SM
Tree Type	SPT	SPT	ShT	ShT
Flood & Prune	O	O	X	X
State Co-existence	X	X	X	O
Uni/Bi-directional	Unidirectional	Unidirectional	Bidirectional	Unidirectional
Encapsulation	X	X	O	O
Loop	X	X	X	X
RPF Check	O	O	X	O

- ① DVMRP는 Flooding/Prune의 동작으로 이루어지므로 첫 데이터가 발생시 DR에 도착하면 IP레이어에서 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^{[15])}).
- ② 이 때, 브로드캐스팅용 레이블을 통해 패킷을 전달 받은 LSR에서는 IP레이어에서 그룹 주소를 확인하고 하위 인접 LSR이 있다면 다시 다음 인접 LSR로 브로드캐스팅용 레이블을 통하여 브로드캐스팅을 하게 되며 모든 하위 LSR에서 일정 시간 동안 Prune을 받지 않으면 수신 측이 하나라도 있는 것으로 간주하여 LDP를 발생시키게 된다. 그러나 모든 인접 LSR이나 인접 수신측으로부터 Prune 메시지를 받게 되면, 상위로 Prune 메시지를 보내게 되지만 레이블 할당용 제어는 필요없게 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 타이머가 끝나지 않았지만 Prune을 받았던 인터페이스로부터 그룹에 가입하겠다는 Graft 메시지를 받는 경우는 레이블할당 LDP가 동시에 이루어지도록 하여 앞에서 서술한 레이블 설정과정을 통하여 설정된 레이블을 사용한다.

표 2. Flooding시 브로드캐스팅 레이블 할당을 통한 LSP 설정

Table 2. Configuration of LSP through broadcasting label assignment

	(ltable,if)	(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)

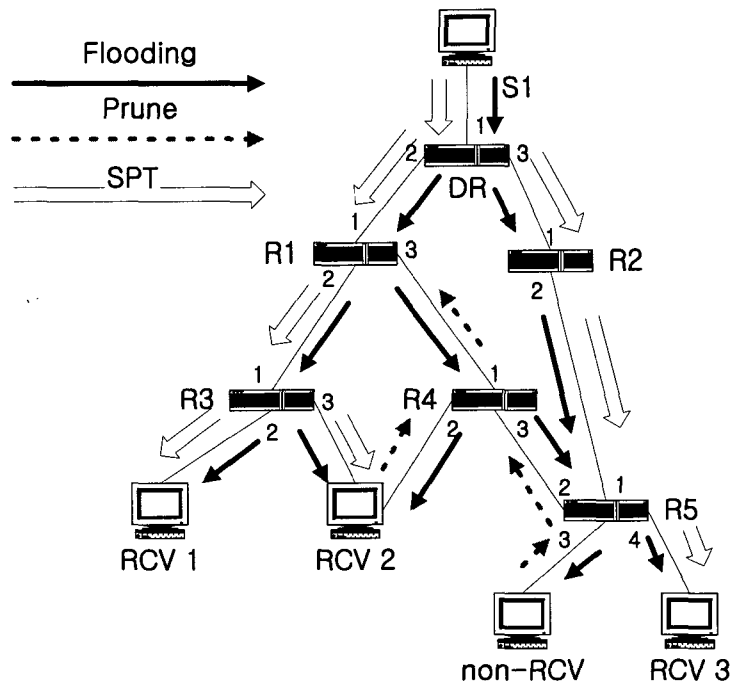


그림 4. DVMRP망에서의 MPLS 적용예
Fig. 4. MPLS Application example over DVMRP Network.

표 3. Prune을 받은 이후 레이블 할당을 통한 각 라우터의 LSP 설정

Table 3. Configuration of each Router's LSP through label assignment After Prune.

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

<N:None, ilable/olable:입/출력 레이블, iif/oif:입/출력 인터페이스>

3. 공유트리방식의 멀티캐스트 라우팅 프로토콜

1) CBT의 MPLS 지원시 문제점

CBT에 MPLS를 적용하는 데 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 기본 동작은 그룹의 각 멤버들이 코어로의 Join 메시지를 통하여 그룹에 참여하게 되며 각 라우터에서 컨트롤 메시지에 따라 양방향성 멀티캐스트 경로를 설정하게 되므로 본 제안에서는 경로 설정을 위한 컨트롤 메시지를 통하여 MPLS 경로를 설정하게 된다.

2) 해결 방안

① 수신 측으로부터 Join메시지 발생시 이 메시지를 보고 LDP를 보내게 하는 Request Driven 방식을 사용하여 레이블을 할당하도록 한다. 이 때 MFC에 기록을 할 때 처음부터 Prune 타이머영역에 설정 시간 값을 넣도록 하며 이 영역에 기록된 값은 Prune 설정시간의 의미가 아닌 할당된 멀티캐스트 레이블 사용 유효 시간의 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 IP상의 터널링을 통한 코어로의 경로를 설정하게 되는데 이를 MPLS 레이블 터널로써 설정하도록 한다. 즉, ①과 동일하게 설정하되 양방향성을 제거한 형태가 되는 것이다. 이 문제에 대해서는 아직 단방향 전송쪽에서도 논

의 되고 있으므로 본 논문에서는 논의하지 않기로 하겠다. 그러나 MPLS 단방향 경로로 설정하게 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 자원 낭비의 문제를 효과적으로 해결할 수 있으므로 이를 참조한다^[6]. 즉, 그림 5와 같이 각 경로를 통해 셀 단위로 나누어져 들어 온 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 RM셀을 붙여 블록단위로 구분지어 전송함으로써 해결한다.

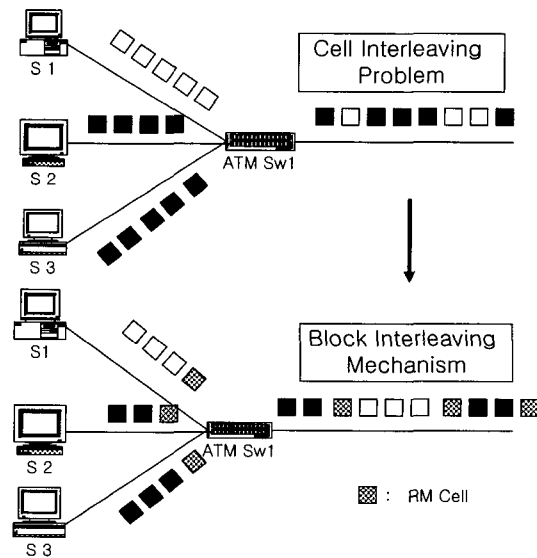


그림 5. CIP와 RM 셀을 이용한 블록 전송 방안
Fig. 5. Block transmission Method using CIP and Rm Cell.

④ 이 후 데이터가 올 때 할당된 레이블을 쓰도록 하며 동시에 Prune 영역에 현재의 시간을 넣도록 한

다. 할당된 레이블을 사용하여 전송할 때에 기존 CBT에서는 반드시 코어까지 전송된 이후 각 수신 측에게로 전송되어 졌으나, 각 LSR의 MPLS 레이어에서 작성된 MFC를 통하여 중간 노드상에 수신 측으로 가는 경로가 있다면 그 출력 경로가 입력 인터페이스와 같지 않다면 전송토록 한다. 그러나, 코어로 도달 전에 모든 수신 측들에게 전송되어 진다 할 지라도 반드시 코어까지는 데이터가 가도록 하여 그룹에 대한 정보가 유지되도록 한다. 예로 그림 6의 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.

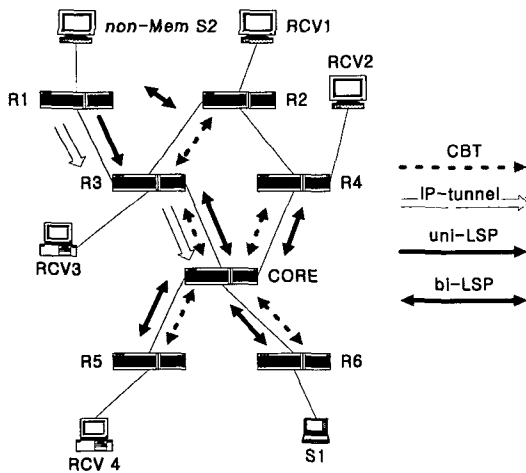


그림 6. CBT 망에서의 MPLS 적용예
Fig. 6. MPLS Application Example over CBT.

4. 복합트리 방식의 멀티캐스트 라우팅 프로토콜
1) PIM-SM의 MPLS 지원시 문제점

PIM-SM에서는 ShT와 SPT를 동시에 공유하고 있는 프로토콜로써 최초 트리 구성 시에는 공유 트리를 사용하다가 일정 데이터율을 넘게 되면 해당 송신 측으로의 최단 경로 트리 전환을 일으키게 된다. 그러나 앞에서 고찰해 본 바와 같이 트리 전환에 MPLS를 쓰기에 부적합한 부분이 많이 있다. 즉, 레이블에는 송신 측을 표기할 만한 공간이 없다는 점과 레이블로 연결된 MPLS라는 경로 자체가 각 노드 상에서 일대일로 이루어진다는 점이다. 따라서 PIM-SM에서 MPLS를 사용하기 위해서는 PIM-SM를 재정의하거나, MPLS를 PIM-SM 경로상에 부분적으로만 쓰도록 해야 한다. 만약 MPLS를 부분적으로 쓰게 되면 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 그룹 내에서 PIM-SM 자체를 좀 더 현실적으로 바꾸자는 의견들이 있으며 그 내용은 다음과 같다^[17].

- 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 특성을 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 피하는 것이기에 PIM-SM을 수정한 것에 MPLS를 지원토록 하는 것도 검토해 볼만한 가치가 있으며 멀티캐스트 워킹 그룹 내의 동향으로 볼 때 오히려, 더 적합한 접근방안이라고 생각된다. 따라서, 본 논문에서는 PIM-SM 자체를 수정하는 방안은 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 해결

하도록 한다. 즉, 최초 멀티캐스트 트리 설정 시에 최소 대역폭, 최소전송률, 최대 지연 시간 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 코어측에서 많은 대역폭을 차지하고 있다는 말이 되므로 기존의 PIM-SM과 같이 수신 측에서 전환을 요구하는 것이 아니라 코어측에서 송신 측 트리로의 전환을 주도하여 하나의 수신 측에서만 SPT가 생성되는 것이 아니라 그룹내의 모든 수신 측들에게 동시에 새로운 그룹주소를 가진 SPT를 생성시켜 주도하도록 한다.

2) 해결 방안

PIM-SM을 MPLS에서 지원하기 위해 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 수신 측 상의 MPLS 경로가 생성된다.
- ② 문제가 되는 소스 트리로의 전환 메시지는 코어에서 트래픽을 모니터링 하고 있다가 특정 송신측으로부터의 데이터가 일정 데이터율을 넘게 되면 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 IP레이어 상의 MRT에 SPT-bit을 설정하여 소스 트리 전환이 일어났음을 표시하도록 하며, 각 수신단은 SPT-Join 메시지를 통해 수신 측으로부터 해당 송신 측까지 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 7을 보면 이미 설정되어

있는 양방향성 LSP 멀티캐스트 트리에서 RCV 4로부터 SPT-Join 메시지가 오면서 S2로 소스트리가 구성되면 RP로부터 SPT-Prune 메시지가 각 노드상으로 멀티캐스팅되면서 다른 수신측도 S2로의 최단 경로를 설정해 나가게 된다.

- ④ ③과 같이 SPT가 생성되고 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 label의 고갈이 우려되므로 레이블 할당 시 되도록 QoS 라우팅을 적용하여 경로를 설정하도록 하며, CBT와 같이 타이머를 이용하여 레이블을 관리토록 한다.

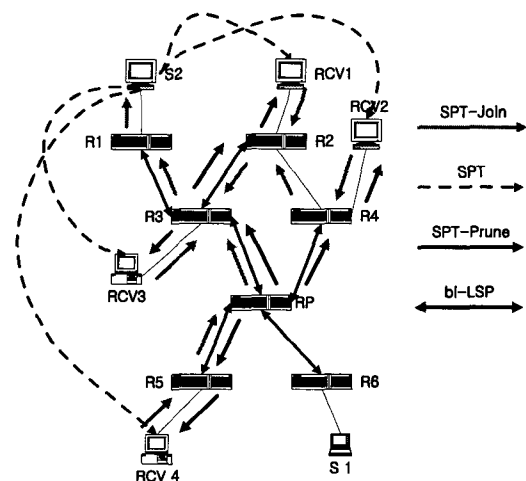


그림 7. 수정된 PIM-SM 동작방식
Fig. 7. Action Method of modified PIM-Sm.

V. 결론

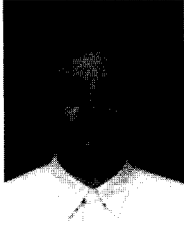
본 논문에서는 기존 인터넷 멀티캐스트 라우팅 프로토콜에 대한 개념정립과 분류를 하였고, MPLS의 개략적 내용을 제시하였다. 기존 멀티캐스트 라우팅 프로토콜은 크게 최단경로트리 방식, 공유트리 방식 및 복합트리 방식으로 분류되며, 각각은 독특한 트리 링크 속성과 트리 유지 방식을 가지고 있다. 이는 MPLS에서 멀티캐스트를 지원하는데 있어 큰 어려움을 갖도록 한

다. 논문에서는 MPLS에서 IP 멀티캐스트를 지원하는 데 있어서 발생가능한 문제점들을 도출하여 이를 앞서 분류한 멀티캐스트 라우팅 프로토콜에 적용하여 분석하였다. 마지막으로 MPLS 망에서 각 멀티캐스트 라우팅 프로토콜을 제공하기 위한 방안을 합리적으로 제시하였고, 구체적인 예를 들어 설명하였다. PIM-SM에 MPLS 제공을 위해 약간의 수정을 가하였고, 이는 MPLS 제공측면에서 매우 합리적인 방안이다. 따라서, IETF에 본 제안을 제출하여 PIM-SM 그룹과 MPLS 그룹에서 동시에 채택할 수 있는 방안을 모색하고자 한다.

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, "Multiprotocol 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] Loa Andersson, Paul Doolan, Nancy Feldman, Andre Fredette, and Bob Thomas, "LDP Specification," <draft-ietf-mpls-ldp-05.txt>, June 1999.
- [14] Bruce Davie, Yakov Rekhter, Eric Rosen, Arun Viswanathan, Vijay Srinivasan, and Steven Blake, "Use of Label Switching With RSVP," <draft-ietf-mpls-rsvp-00.txt>, March 1998.
- [15] Eric C. Rosen, Yakov Rekhter, Daniel Tappan, Dino Farinacci, Guy Fedorkow, Tony Li, and Alex Conta "MPLS Label Stack Encoding," <draft-ietf-mpls-label-encaps-04.txt>, April 1999.
- [16] 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.
- [17] 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.

저 자 소 개



金 榮 俊 (正會員)

1986년 한양대학교 전자공학과 졸업. 1990년 한양대학교 대학원 전자공학과 공학석사. 1994년 한양대학교 대학원 전자공학과 박사과정 수료. 1996년~현재 혜천대학 정보시스템 계열 조교수. 주관심 분야: 컴

퓨터 통신, 분산 시스템, 초고속 통신, 이동데이터 통신



朴 容 震 (平生會員)

1969년 와세다 대학교 전자통신공학과 졸업. 1971년 와세다 대학원 공학석사. 1978년 와세다 대학원 공학박사. 1979년~현재 한양대학교 전기전자공학부 교수. 1983년~1984년 Univ. of Illinois, Urban 전산학과

방문 교수. 1991년~1992년 영국 Kent 대학 방문교수. 주관심분야: 컴퓨터 통신, 분산시스템, 초고속 통신, 이동데이터 통신