

MPLS망에서 IP Multicast 적용

김갑동* · 박창민** · 김상하*

*충남대학교 컴퓨터학과 · **한국전자통신연구원

IP Multicast deployment in an MPLS Environment

Kap-Dong Kim* · Chang-Min Park** · Shang-Ha Kim*

*Dept. of Computer Science, Chungnam National University

**Electronics and Telecommunications Research Institute

E-mail : kdkim@cclab.chungnam.ac.kr

요 약

MPLS (Multiprotocol Label Switching)는 인터넷에서 속도, 확장성 그리고 서비스 제공 능력을 향상시키기 위한 백본 네트워크로 이용하기 위하여 Internet Engineering Task Force (IETF) standard로 급속하게 성장하고 있다. MPLS는 기존의 인터넷 라우팅에서 사용하는 longest prefix match 방식을 이용하는 라우팅 방법 대신에 short label exact match 방식과 L3 forwarding 방법을 사용함으로써 고속 인터넷 서비스 기술을 제공한다. 한편, 기존 인터넷의 IP format을 보면 Host Id, (Address Prefix, Host Id), Broadcasting, Multicasting의 네 가지 종류의 구조를 가지고 있다. 그러나 아직 MPLS에서의 멀티캐스트 서비스에 대한 방식이 아직 표준화되고 있지 않고 있는 상황에서, MPLS 서비스 도메인 내에서 멀티캐스트 패킷은 기존의 멀티캐스트 라우팅 프로토콜과 연계되어 MPLS상에서 멀티캐스트 서비스가 이루어져야 한다. 본 논문에서는 기존의 인터넷상에서 제공되어지는 멀티캐스트 IP 플로우들을 MPLS 상에서 수용하기 위한 방식을 제안한다.

I. 서 론

MPLS (Multiprotocol Label Switching)는 기존의 인터넷 라우팅에서 사용하는 L3 포워딩 방식을 이용하는 라우팅 방법 대신에 L2 포워딩 방식을 사용함으로써 고속의 인터넷 서비스를 제공한다. 또한, 동일한 FEC (Forwarding Equivalency Classes)에 속하는 패킷들을 merging하여 하나의 레이블을 사용하는 방법 즉, MPT (Multipoint to Point Tree)를 형성함으로써 확장성을 향상시킨다. 여기에서 FEC (Forwarding Equivalency Classes)는 동일한 경로를 따르는 패킷들의 집합을 말한다. 현재까지 MPLS 스펙에서 제안되고 있는 FEC 타입에는 IP Address Prefix, Router ID, Application flow가 있다. 기존 인터넷의 IP format을 보면 Host Id, (Address Prefix, Host Id), Broadcasting, Multicasting의 네 가지 종류의 구조를 가지고 있다. 본 논문에서는 MPLS에서의 멀티캐스트 서비스에 대한 방식이 아직 표준화되고 있지 않고 있는 상황에서, MPLS 서비스 도메인 내에서 멀티캐스트 패킷을 기존의 멀티캐스트 라우팅 프로토콜과 연계하여 MPLS상에서 멀티캐스트 서비스가 이루어지기 위한 방안으로 멀티캐스트 라우팅 프로토콜로 PIM(Protocol Independence Multicast)의 Sparse Mode와

Dense Mode를 제안 하고자 한다.

II. MPLS

1. 개 요

MPLS는 기존의 인터넷 라우팅에서 사용하는 longest prefix match 방식 대신 short label exact match 방식을 사용함으로써 단순하고 고속의 포워딩 기술을 제공한다. 기존 방식에서 매 홉마다 패킷에 대한 포워딩 결정은 헤더의 내용과 라우팅 알고리즘에 기반을 두고 스트림을 할당한다. MPLS에서는 특정 패킷을 스트림으로 할당하는 것을 패킷이 망에 들어 올 때 한번만 수행을 한다. 패킷에 할당된 스트림은 레이블이라는 값으로 인코드 되며 레이블이 할당된 다음부터 패킷은 레이블에 의해 다음 홉으로 포워딩 된다. 즉 첫 번째 패킷에 대해서만 L3 포워딩을 하고 다음에 오는 패킷에 대해서는 L2 포워딩을 함으로써 라우터의 성능을 향상시킨다.

2. 구성요소

MPLS에서 사용되는 레이블은 스트림을 식별하기 위해 사용하는 작고 고정된 크기의 식별자로 패킷이 할당된 스트림을 기반으로 하며, 레이블의

선택은 IP 헤더의 내용과 FEC에 따라 선택된다. MPLS의 가장 기본적인 요소는 Upstream LSR, Downstream LSR, outgoing Label, incoming Label, Stream 등이 있다. Upstream LSR는 Downstream LSR로 패킷을 전송하는데 Stream에 대해 Label을 협상하게 된다. 협상이 되면 Label은 Upstream LSR에게는 outgoing Label이 되고 Downstream LSR에게는 incoming Label이 된다. 패킷은 네트워크 계층 또는 데이터 링크 계층의 적절한 위치에서 레이블 인코딩 절차에 따라 전달된다.[1]

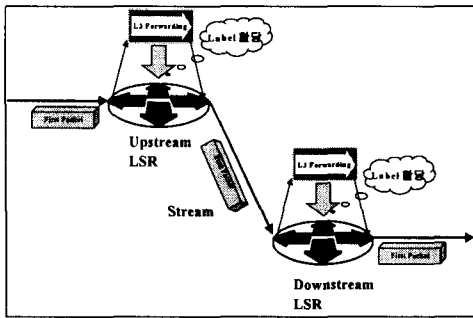


그림 1 L3 Forwarding

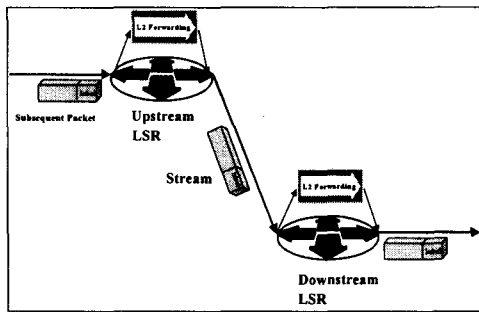


그림 2 L2 Forwarding

3. 레이블 할당 방식

가. Topology Driven 방식

라우팅 프로토콜의 제어 트래픽의 정상적인 처리에 응답하여 레이블이 할당된다. 이 방식의 특성은 레이블 할당과 분배에 걸리는 계산 부하와 레이블을 분배하는데 소비되는 대역폭은 네트워크의 크기에 의해 정해지며, 트래픽이 도착하기 전에 레이블들이 미리 할당되어 분배된 상태로 트래픽이 도착했을 때는 레이블 설정을 위한 지연이 발생하지 않고 바로 레이블 스위칭으로 데이터가 전송된다.[1]

나. Request Driven 방식

Request 기반 제어 트래픽의 정상적인 처리에 응답하여 레이블을 할당한다. 레이블 스위칭 라우

터가 제어 메시지를 처리할 때 레이블을 할당한다. 이 방식의 특성은 레이블 할당과 분배에 걸리는 계산 부하와 레이블을 분배하는데 소비되는 대역폭은 시스템내의 제어 트래픽의 양에 의해 정해지며 트래픽이 도착하기 전에 레이블들이 미리 할당되어 분배된 상태로 트래픽이 도착했을 때는 레이블 설정을 위한 지연이 발생하지 않고 바로 레이블 스위칭으로 데이터가 전송된다.[1]

다. Traffic Driven 방식

LSR에 데이터가 도착했을 때 레이블을 할당하고 분배한다. 이 방식의 특성은 레이블 할당과 분배 비용이 트래픽 패턴의 함수로 플로우의 수와 이 플로우의 지속성에 비례하여 오버헤드가 증가하며, 하나의 플로우에 레이블을 할당하는데 지연이 발생한다.[1]

4. LSP(Label Switched Path) 제어 방식

LSP 설정방법은 각 노드에서 시작하거나 출구 노드에서 실행하는 방법으로 구분할 수 있다. 각 노드에서 수행하는 방법은 각 노드에서 새로운 레이블 할당을 시작하여 Upstream 노드에게 전달하는 방법으로 각 노드에서 수행하는 관계로 할당하기 위한 레이블 granularity의 불일치가 발생할 수 있으며 이를 독립적인 방식이라고 한다. 반면, 출구 노드에서 레이블 할당을 시작하여 upstream 노드에게 차례로 레이블이 바인딩 되는 방법은 하나의 FEC에 대한 LSP 설정에 대해 일치가 유지되기 때문에 granularity 불일치가 일어나지 않으며 이를 순서화된 방식이라 한다.[2]

5. FEC 타입

MPLS에서는 IP패킷들이 각 LSP에 매핑 되는 것을 정확히 정의할 필요가 있으며, 이러한 정의는 각 LSP에 대해 FEC 스펙을 제공함으로써 행해지며 FEC 타입은 다음과 같다.

- ◎ IP Address Prefix : 하나 이상의 IP address prefix의 리스트를 제공하는 것으로 목적지 주소가 하나 이상의 특정 Prefix에 일치하는 모든 IP 패킷은 관련된 LSP를 사용하여 포워딩 된다.
- ◎ Router Id : 라우터의 ID를 제공하는 것으로 목적지로 가는 경로가 특정 라우터를 통과하는 것으로 알려진 모든 IP 패킷은 관련된 LSP를 사용하여 포워딩 된다.
- ◎ Flow : 포트, 목적지 주소, 발신지 주소와 같은 데이터그램 정보를 특징짓는 것으로 집합이 없는 MPLS 플로우를 제공하기 위한 Label Distribution Protocol을 제공한다.

III. 멀티캐스트

현재 제안되고 있는 네트워크 계층에서의 멀티

캐스트 프로토콜들을 분류하면 공유 트리 방식 (Shared Tree)으로 PIM-SM, CBT가 있고, 소스 기반 트리 방식(Source-specific free)으로 PIM-DM, DVMRP, MOSPF등이 있다. 각 멀티캐스트 라우팅 프로토콜들은 다음과 같다.[7]

1. DVMRP와 MOSPF

DVMRP 프로토콜은 멀티캐스트 분산 트리를 구성하는데 broadcasting과 pruning을 기반으로 하는 data-driven 방식을 취한다. MOSPF는 진정한 의미의 단말까지의 Dijkstra 알고리즘을 기반으로 한 최단 경로를 구성한다. 두 프로토콜의 문제점으로 멀티캐스트 범위내의 모든 라우터는 멀티캐스트 그룹에 대한 상태 정보를 유지해야 하므로 주기적인 데이터의 flooding이 발생하며 프로토콜 자체의 오버헤드가 크기 때문에 확장성에 문제가 있다.[7]

2. CBT (Core Based Tree)

CBT는 전통적인 코어 중심의 멀티캐스트를 기반으로 한다. 코어 집합을 정의하여 그 중 하나를 Primary 코어로 나머지는 Secondary 코어로서 지정한다. CBT방식은 하나의 중앙 코어를 정점으로 하는 공유 트리를 구성하기 때문에 sub-optimal 경로문제, 모든 트래픽이 Primary 코어를 경유하여 공유 트리를 통해 흐름으로 발생되는 통신 병목현상 및 Primary 코어의 동작 불량 시 전체 그룹 통신이 마비되는 불안정성이 문제점으로 제시된다.[7]

3. PIM(Protocol Independent Multicast)

PIM은 Sparse Mode와 Dense Mode가 있는데, SM는 공유 트리 방식을, DM는 소스 기반 트리 방식을 지원한다. CBT의 단점을 극복하기 위한 방법으로 PIM-SM에서는 특정 데이터 전송을 임계치를 넘는 소스에 대해 공유 트리에서 소스 기반 트리로 전환할 수 있는 트리 전환 메커니즘을 제안하고 있다. 트리 전환 메커니즘을 이용하면 전송율이 높은 데이터가 공유 트리로부터 분리되어 독립적인 최단경로를 구성하므로 공유 트리에서의 병목현상을 없애고 sub-optimal 경로로 인한 전송지연을 줄일 수 있다는 장점이 있다.[4,5,7]

IV. MPLS망에서의 멀티캐스트

1. PIM Dense Mode의 매핑

PIM-DM는 LSR에 멀티캐스트 플로우에 대한 포워딩 캐쉬 엔트리가 없으면 이 플로우에 대한 첫 번째 데이터 패킷이 이 플로우의 포워딩 캐쉬를 인스톨한다. 새로 생성된 멀티캐스트 포워딩 캐쉬의 각 출력 인터페이스에 대하여 LSR은 downstream LSR에 패킷을 전송하기 위하여 레

이블을 할당한다. 이때 LSR은 플로우 테이블에 플로우 ID <source IP address, destination IP address, TTL>를 LSP { <input port, input label>, list of <output port, output label>}로 매핑 시키는 엔트리를 추가한다. 그리고 L2 스위칭 테이블에 { <input port, input label>, list of <output port, output label>}를 추가한다. 이 플로우의 다음에 오는 패킷들은 L2 스위칭으로 전송된다. IP 멀티캐스트 라우팅 프로토콜은 LSR에서 멀티캐스트 포워딩 상태를 설정하는데 사용되고 LSR은 이 상태를 점 대 다중점 레이블로 매핑한다.

PIM-DM에서 플로우의 포워딩 상태는 종결 타이머(expiration timer)와 연관이 있다. 이 타이머는 데이터 패킷이 포워딩 될 때마다 다시 작동한다. 그러나 LSR에서 레이블이 할당된 플로우의 데이터 패킷은 L2 스위칭을 하므로 종결 타이머를 다루는 규칙은 다음과 같이 바뀐다. 타이머가 종결되었을 때 최근에 타이머가 종결된 후 이 플로우의 LSP에 포워딩 동작이 있었는지 체크한다. 만약 이 기간동안 L2 스위칭이 있었다면 타이머는 다시 작동하고 그렇지 않았으면 포워딩 상태와 LSP는 삭제된다.

포워딩 캐쉬 엔트리와 연관된 출력 인터페이스의 리스트는 Prune 메시지와 Graft 메시지에 의해 수정된다. Prune 메시지를 받은 인터페이스의 <output port, output label>은 L2 스위칭 테이블에서 지워진다. 즉, 점 대 다중점의 레이블의 branch가 삭제된다. 만약 branch의 삭제 후 이 엔트리의 branch가 하나도 남아 있지 않으면 LSP를 삭제한다. 그리고, 입력 인터페이스의 레이블은 반환된다. 그룹 G에 대한 Graft 메시지를 받았을 때 PIM-DM는 모든 (S, G) 포워딩 엔트리에 대한 출력 인터페이스 리스트에 레이블을 할당하고 <output port, output label>을 추가시킨다. LSR은 L2 스위칭 테이블에 Graft 메시지에 의해 수정된 각각의 (S, G)에 상응하는 LSP에 <output port, output label> branch를 추가한다. (S, G) 엔트리에 LSP가 없으면 새로운 LSP는 첫 번째 패킷이 도착했을 때 생성된다. MPLS는 PIM 제어 메시지와 관련된 PIM 프로토콜 동작을 바꾸지 않는다. MPLS는 멀티캐스트 포워딩 상태를 L2 스위칭 테이블로 어떻게 매핑 할 것인가 결정한다.

2. PIM Sparse Mode의 매핑

PIM-SM는 리프로 수신자를 루트로 RP (Rendezvous Point)를 갖는 공유 포워딩 트리를 생성한다. PIM-SM는 PIM-DM와는 다르게 중간 라우터에서의 포워딩 엔트리는 last-hop 라우터에서 공유 트리의 RP 또는 소스 기반 트리의 소스 방향으로 explicit Join 메시지를 보냄으로써 생성된다. MPLS에서 중간 라우터는 RP 트리의 공유 포워딩 엔트리를 per-source LSP로 매핑한다. 패킷 레벨의 라우터는 공용의 (*, G) 패킷 레벨

터캐스트 포워딩 상태를 유지하지만 MPLS에서는 L3 포워딩 상태는 모든 (*, G) 플로우에 의해 공유되지만 L2 포워딩은 각 (S, G) 플로우로 분리된다. LSP를 분리하는 이유는 PIM-SM는 디폴트로 모든 수신자에 도착하려고 공유 트리를 사용하면서 동시에 특정한 송신자에 대하여 수신자의 서브셋에 도착하려고 공유 트리의 부분을 이용한다. 그러므로 트리에 있는 중간노드는 (*, G)와 (S, G)의 포워딩 엔트리를 포함한다. (*, G) 엔트리 또는 (S, G) 엔트리를 사용하는 결정은 각 패킷에 기반을 두고 있다. 그러나 이런 결정은 L2에서 데이터가 포워딩 될 때는 결정이 내려질 수 없다. 그래서, MPLS는 공유의 점 대 다중점 LSP가 사용되는 대신 구분된 LSP가 각 멀티캐스트 플로우에 대하여 사용된다. MPLS에서 멀티캐스트 플로우의 첫 번째 데이터 패킷은 PIM-SM와 PIM-DM 양쪽이 비슷하게 다루어진다.

Prune 이나 Join 메시지가 멀티캐스트 포워딩 엔트리에 영향을 미칠 때 플로우 테이블에 있는 엔트리는 수정된다. Prune 메시지를 받으면 포워딩 엔트리와 연관된 각 LSP는 삭제된다. Join 메시지를 받으면 레이블이 선택되고 포워딩 엔트리와 연관된 LSP가 추가된다.

RP는 Register 패킷을 받으면 플로우에 대한 LSP를 생성하지 않고 플로우에 대한 <input port, input label>은 없는 <output port, output label> 리스트를 포함한다. 소스로부터 멀티캐스트 패킷이 도착하면 LSP가 생성된다.

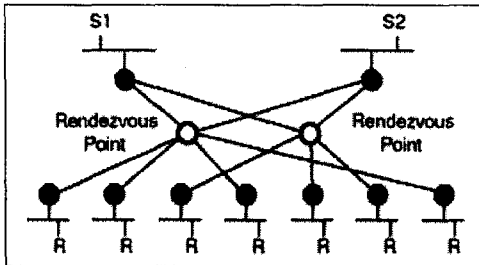


그림 3 Rendezvous Point

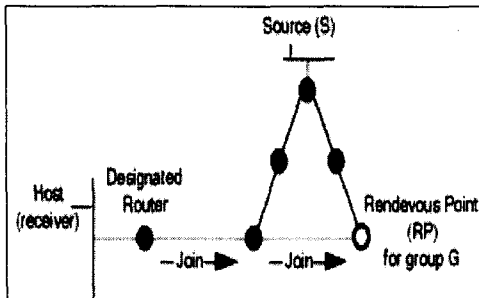


그림 4 Host Joins a Multicast Group

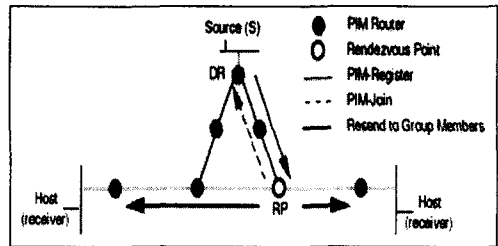


그림 5 Source Sends to a Multicast Group

V. 결론

MPLS 환경에서 IP 멀티캐스트의 적용은 기존의 인터넷 서비스에서 멀티캐스트 서비스를 위한 멀티캐스트 라우팅 방안들과 함께 연계하여 이루어져야 한다. 본 논문에서는 여러 멀티캐스트 라우팅 프로토콜들 중 PIM-DM와 PIM-SM를 MPLS에 적용하여 보았다. 현재 인터넷에서 속도, 확장성, 서비스 제공 능력의 향상을 위하여 MPLS가 급속히 성장하고 있는 추세에 발 맞추어 멀티캐스트 적용에 대한 연구가 활발히 진행되어야 한다.

VI. 참고문헌

- [1] R. Callon, et al. "A Framework for Multi-protocol Label Switching," Internet Draft, draft-ietf-mpls-framework-02.txt November 1997
- [2] E. Rosen, A. Viswanathan, and R.Callon, "A Proposed Architecture for MPLS", Internet Draft, July 1997.
- [3] D. Farinacci and Y. Rekhter, "Multicasting Tag binding and Distribution using PIM" IETF draft, draft-farinacci-multicast-tagsw-00.txt, December 1996
- [4] S. Deering and et. al. "Protocol Independent Multicast Version 2, Dense Mode Specification. Internet draft draft-ietf-idmr-pim-dm-05.txt, May 1997.
- [5] D.Estrin and et. al. "Protocol Independent Multicast Sparse Mode(PIM-SM)" Protocol specification. RFC 2117.
- [6] D.Ooms and et. al. "Framework for IP Multicasting in MPLS" Internet Draft, August, 1998.
- [7] Dave Kosiur. "IP Multicasting", Wiley Computer Publishing, 1998.