

# 오버레이 멀티캐스트 기술 분석

박주영\* · 정옥조\* · 강신각\*

\*한국전자통신연구원 표준연구센터

## Study on Overlay Multicast Mechanisms

Juyoung Park\* · Ok Jo Jung\* · Shin Gak Kang\*

\*ETRI Protocol Engineering Center

E-mail : jypark@etri.re.kr

### 요 약

최근 대두되는 네트워크 게임이나 인터넷 생방송 등의 각종 그룹 통신응용을 지원하기 위해선 IP 멀티캐스트가 지원되어야 한다. 하지만 IP 멀티캐스트를 적용하기 위해선 해결되어야 할 문제점들로 인하여 현재의 인터넷에서는 적용되지 못하고 있다. 오버레이 멀티캐스트 기술은 IP 멀티캐스트가 도입되지 않은 현재의 유니캐스트 기반 인터넷 환경에서 그룹 통신을 보다 효율적으로 제공해 줄 수 있는 메커니즘으로, 최근 많은 연구가 진행되고 있다. 본 고에서는 이들 오버레이 멀티캐스트 기술을 보다 체계적으로 분석한다.

### ABSTRACT

To support emerging group applications such as network games and Internet live-casting efficiently, IP multicast mechanism is highly needed. But IP multicast still has not been deployed in the current Internet because of its difficulty to manage. Overlay multicast mechanisms are proposed to serve group communications more effectively in the current unicast oriented Internet. In this paper we categorize the proposed ones into topological view point.

### 키워드

IP Multicast, overlay multicast

## 1. 서론

IP 멀티캐스트의 가장 큰 특징은 송신자가 다수의 특정 수신자들에게 데이터를 반복적으로 전달하지 않고, 대신에 라우터가 설정한 데이터 1:N 전달 경로를 통하여 단 한번의 데이터만 보낸다는 것이다. 그렇기 때문에 송신자의 입장에서는 데이터를 중복적으로 보내기 위해 소모되는 시스템 자원을 절약할 수 있으며, 아울러 중복된 데이터로 인한 망 자원의 낭비를 방지할 수 있다. 그러나 IP 멀티캐스트는 인터넷 그룹 주소의 할당문제, 그룹 상태 관리를 위해 라우터에 수반되는 오버헤드, 현재 인터넷에 설치된 라우터를 멀티캐스트가 가능하도록 변경하기 위한 비용 문제 등의 이유로 인해, 실제 인터넷 망에 널리 적용하지 못하는 실정이다. 이러한 문제점들로 인하여 하부 인터넷 라우터를 수정하기 보다는, CP(Contents Provider)단위로 중

단 호스트 혹은 서버들을 이용한 오버레이 멀티캐스트 기술 방안이 논문 혹은 업계의 솔루션 형태로 많이 제안되고 있다.

오버레이 멀티캐스트의 특징은 IP 라우터가 멀티캐스트를 지원하는지에 관계없이, 종단 호스트 간에 데이터 전달 트리를 구성하여 그룹 통신 서비스를 제공해 주는 것이다. 오버레이 멀티캐스트의 동작 방식은 크게 데이터 전달 경로를 설정해 주는 부분과 데이터를 전달해 주는 부분으로 나눌 수 있는데, 근본적인 기술은 IP 멀티캐스트의 그것과 매우 유사하다. 즉 데이터 전달 경로를 설정해 주는 메커니즘은 멀티캐스트 라우팅 프로토콜과, 데이터 전달 메커니즘은 IP 멀티캐스트의 패킷 복사-후-전달 방식이 오버레이 멀티캐스트 엔터티에서도 구현되어 있다. 그러나 오버레이 멀티캐스트 환경에서 데이터 전달 경로를 설정하기 위한 이웃 노드

들은 IP 멀티캐스트의 경우와는 달리 비록 논리적으로는 이웃하지만 물리적으로는 서로 떨어져 있을 수 있기 때문에, 데이터 전달 경로를 형성해야 하기 위해 필요한 이웃 노드들에 관한 위치를 얻을 수 있는 메커니즘이 필요하다. 또한 데이터 전송 경로에 있는 노드들이 일반 PC나 워크스테이션이 될 수 있기 때문에 쉽게 다운되거나 세션에서 탈퇴할 수 있기 때문에 보다 복잡한 트리 관리 메커니즘이 필요하다. 그러나 IP 멀티캐스트와 상반되는 개념으로써, 획일화된 데이터 전송 경로가 아니라 응용 서비스에 따라 구성하기 때문에 신뢰적인 데이터의 전송이나 보안등의 기능을 보다 쉽게 제공할 수 있다. 본 고에서는 현재의 유니캐스트 인터넷 기반 환경에서 그룹 통신을 가능토록 제안된 각종 오버레이 멀티캐스트 기술들이 해결하고자 하는 기술적 특징 및 발전 방향을 기술한다.

## II. 탄생배경

멀티캐스트는 불특정 다수의 수신자들에게 동일한 데이터를 전송하는 브로드캐스트 방식과는 달리 어느 특정한 수신자 그룹으로만 데이터를 전송하는 기능을 말한다. 이러한 기능을 IP 망에서 지원하는 기술을 IP 멀티캐스트라고 말하는데, IP 멀티캐스트 기능을 구현하기 위해서는 모든 IP 망 장비(라우터)들이 멀티캐스트 라우터 기능을 지원해야 한다. 멀티캐스트 라우터는 멀티캐스트 라우팅 프로토콜(MRP)를 통하여 송수신자간 데이터 전송 트리를 구성하고, 이 트리를 따라 데이터를 복사 및 전달 하는 과정을 거쳐 수신자들에게 데이터를 전달한다.

IP 멀티캐스트의 가장 큰 장점으로는 망 대역폭을 절약할 수 있다는 것이다. 또한 부수적으로 다수의 수신자들에게 동일 데이터를 반복적으로 송신하지 않아도 되기 때문에 송신자의 자원을 절약할 수 있다는 것이다. 그러나 단점으로써는 다음과 같은 것들을 들 수 있다.

- 1) IP 라우터(멀티캐스트 라우터)가 반드시 [송신자, 멀티캐스트 그룹] 정보를 갖는 데이터 전달 상태를 관리하기 위한 부하가 든다.
- 2) 라우터는 각각의 멀티캐스트 데이터들을 그룹으로 전달하기 위해 데이터 전달 테이블을 검색하여야 한다. 그런데 멀티캐스트 포워딩 테이블은 주소에 따라 머지 되지 못하기 때문에, 모든 주소 쌍을 검색해야 한다.
- 3) IP 멀티캐스트는 open-ended 방식을 기본으로 구현하였기 때문에 보안에 대한 취약점을 안고 있다. 이러한 약점 때문에 악의를 갖는 수신자들이 그룹에 가입/탈퇴를 임의로 할 수 있다.
- 4) 가장 큰 문제점중의 하나로써 IP 멀티캐스트의 주소는 사용자가 임의로 사용할 수 있다는 것이다. 이런 문제로 Class D의 범위를 사용하는 IP 멀티캐스트 주소의 할당을 고려할 수 있다.
- 5) IP 멀티캐스트는 L3 계층의 전송 메커니즘으

로써 비신뢰적 best effort 전송만을 가능하다. 따라서 신뢰적인 데이터 전송이 필요할 경우, 신뢰적 멀티캐스트 전송 프로토콜을 구현해야 하며, best effort 전송만을 고려하기 때문에 흐름 및 체증 제어가 필요하다.

6) 앞의 이유 때문에 IP 멀티캐스트의 채용이 늦어지고 있으며, 이에 따른 ISP나 CP(Contents Provider)들의 수익 모델이 결여되고 있다.

이러한 문제점의 대안으로써 2000년경부터 Application Layer Multicast (혹은 Overlay Multicast)이라 불리는 기술이 많은 연구단체에서 제안되고 있다.

ALM 기술은 IP 멀티캐스트 라우터의 L3계층에서 데이터를 복사 전달하는 과정을 중단 호스트의 L5계층에서 처리하도록 한다.

이러한 이유로 다음과 같은 단점을 보인다.

- 1) L3계층이 아닌 L7 계층에서 데이터 전달을 수행함에 따른 delay
- 2) 망 장비가 아닌 중단 호스트에서 패킷 전달이 이루어짐에 따른 대역폭 사용 증가
- 3) 망 장비가 아닌 중단 호스트이기 때문에, 세션에 가입 혹은 탈퇴에 따른 세션 트리의 안정성 그러나 ALM의 장점으로써 다음의 사항을 고려할 수 있다.
  - 1) 라우터가 아닌 중단 호스트 혹은 부분적인 하드웨어의 설치만으로도 쉽게 그룹통신을 할 수 있기에 인터넷에서 빠르게 수용될 수 있다.
  - 2) 범용성을 갖는 망 장비가 아닌 중단 노드의 응용으로 동작함으로써, 목적하는 응용서비스에 따라 효율적인 멀티캐스트 데이터 트리를 관리할 수 있다.
  - 3) 획일적인 트리보다는 응용 서비스의 방식에 따른 효율적인 트리 구성이 가능하다.

## III. Overlay Multicast

오버레이 멀티캐스트는 일반 IP 멀티캐스트와는 달리 응용 서비스와 매우 밀접한 관계를 가지고 있기 때문에 응용 서비스의 요구사항에 따라 효율적으로 설계될 수 있다. 이들 오버레이 멀티캐스트는 수율, 구성원 및 오류 복구에 사용되는 메커니즘에 따라 자기 다른 성능을 보인다.

또한 오버레이 멀티캐스트에서 가장 중요한 부분은 바로 노드 간에 전송트리를 어떻게 구성할 것인가라고 할 수 있다. 다양한 메커니즘들 중에서 오버레이 멀티캐스트를 특성지을 수 있는 몇 가지 특징은 다음과 같다. 첫 번째 오버레이 멀티캐스트에서 트리를 구성하기 위해서 특정 노드를 지속적으로 할당하는가의 여부에 따라 Fixed Node-based Overlay와 Dynamic Nodes-based Overlay로 구분된다. 두 번째로는 데이터 전송 트리와는 별도의 제어 트리가 어떤 형태로 구성되는가에 따라 Mesh-first based Approach와 Tree-first based Approach로 구분된다. 세 번째로 데이터 전송 트

리가 응용 서비스의 성격에 따라 구성되는가에 따라 Structured 방식과 Un-structured 방식으로 구분된다. 즉, Unstructured 방식은 일반적인 IP 멀티캐스트 기능을 제공하기 위한 방식인 반면, structured 방식은 응용 서비스에 최적인 트리를 구성하는 방식이다. 끝으로 오버레이 멀티캐스트는 응용 계층에 존재하는 기술로써, 하부 네트워크의 특징을 알 수 없기 때문에 하부 망의 토폴로지에 대한 추정을 하는지의 여부에 따라 Topology-Aware Overlay 방식 여부를 결정한다. 오버레이 멀티캐스트에서 데이터를 전달하는 방법은 홉-바이-홉 유니캐스트 방식으로 데이터를 전송하며, 전송 채널은 TCP, UDP 혹은 IP-in-IP 방식 등 응용 서비스에 따라 적합한 방식을 사용한다.

이러한 특성을 갖는 오버레이 멀티캐스트 메커니즘은 다음과 같이 오버레이 트리를 구성하는 방식에 따라 차이를 보이며, 아래에 대표적인 프로토콜을 통해 소개하도록 한다.

● Fixed Node Based Multicast

이 방식은 오버레이 멀티캐스트 세션을 제공할 때, 특정 고정 노드들을 사용하는 방법이다. 기본적인 개념은 하나 이상의 고정된 노드들을 통해 트래픽 전송을 위한 오버레이 데이터 전송 트리를 구성한다. 현재 이 기술을 사용하는 프로토콜들로는 RON, Overcast, ScatterCast, AMCast 등이 제안되고 있다. 이 중 ScatterCast는 버클리 대학에서 제안한 프로토콜로써, 주로 비교적 작은 규모의 그룹에 신뢰적인 파일 전송에 초점을 맞추고 있다. 신뢰적 전송을 위하여 TCP 데이터 채널을 사용하고 트리를 구성하기 위해선 RMX (Reliable Multicast proXies) 방식을 사용한다.

● Dynamic Node Based Multicast

고정 노드 기반과는 달리 동적인 노드 기반의 멀티캐스트는 어떤 특정한 노드를 가지는 대신, 세션에 가입한 모든 노드들로 하여금 멀티캐스트 포워딩 기능을 제공해 주도록 하는 방식이다. 하지만 고정 노드가 아닌 일반 노드이기 때문에 세션의 탈퇴 등에 따라 멀티캐스트 트리가 흔들릴 수 있다.

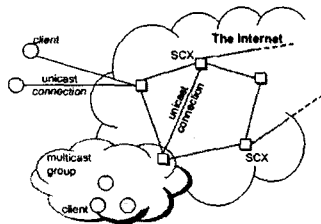


그림 1 ScatterCast

동적 노드 기반의 오버레이 멀티캐스트에서는 응용 계층의 서비스에 따라 데이터 전송 트리를

구성하는 방식을 정의할 수 있는데, 특정한 응용에 상관없이 일반적인 그룹 데이터를 전송용 오버레이 멀티캐스트 트리를 구성하는지에 따라 unstructured 방식과 structured 방식으로 구분된다. 또한 이웃 노드들과의 제어 트리를 구성하는 방법에 따라 mesh 방식과 tree 방식으로 구분한다.

1) unstructured & Mesh-first 방식

이 방식은 다이나믹 노드 기반의 한 방식으로써, 주 목적은 "컨텐츠와 무관한 오버레이 멀티캐스트 트리의 효과적인 관리"를 하는 것으로써, 멀티캐스트 데이터 트리가 아닌 제어 트리를 그룹 멤버들간 mesh 형태로 구성하여 타 그룹 멤버들에 대한 정보 관리를 한다.

이런 제어용 mesh를 통하여 멀티캐스트 트리의 partitioning을 방지 혹은 검출할 수 있는데, 트리를 구성하는 방식은 우선 이웃 노드들 간에 mesh를 구성하고, 오버레이 멀티캐스트 라우팅 프로토콜을 이용하여 멀티캐스트 데이터 전송 트리 구성한다.

이 메커니즘은 빈번한 가입과 탈퇴 이벤트에 있어서 망의 상태에 보다 빠르게 적용할 수 있다는 장점이 있다.

이 메커니즘을 사용하는 프로토콜들로는 End System Multicast 이나 Banana Tree Protocol 등이 있다. End system multicast 프로토콜은 특히 "small & sparse" 그룹에서 멀티캐스트 서비스를 제공하는 것을 목표로 한다.

멀티캐스트 트리를 생성하는 방법은 mesh를 통한 shortest path tree(SPT)를 구성하는 것이다. 만일 새로운 멤버 가입하였을 경우, well-known 노드들에 대한 부분적인 멤버 리스트 획득한다. 또한 모든 멤버들은 다른 멤버들의 멤버십을 주기적 갱신하면서 동작하지 않는 링크 목록을 별도 관리함으로써 partitioning 검사한다.

트리의 품질을 개선하기 위해선 자신의 부모 노드가 아닌 다른 멤버들에 대한 주기적으로 탐색하면서 만일 더 좋은 링크 발견시, 해당 링크로 연결한다.

2) unstructured & Tree-first 방식

이 방식도 mesh 방식과 동일한 목적을 갖지만, 제어 트리로써 오버레이 멀티캐스트 데이터 트리를 직접 사용하는 방식이다. 이 방식 또한 생성된 멀티캐스트 데이터 트리가 동적인 환경에서도 트리를 구조를 강건하게 유지하도록 지속적인 관리를 수행한다. 그런데 트리 구조에서는 각각의 노드들이 전체 노드들에 대한 정보를 가지고 있지 않기 때문에, SPT 알고리즘 이외에도 "loop detection & avoidance"에 관한 알고리즘이 요구된다. 이 메커니즘을 사용하는 프로토콜들은 YOID, Host Multicast, TBCP, Scalable Self-organizing overlays, HBM 등이 있다.

Yoid는 Yallcast로부터 호칭 변경된 것으로써 호스트 기반으로 컨텐츠를 분배하는데 사용되는 프

로토콜이다. Yoid는 응용에게 "lower middle-ware" 로써의 서비스를 제공한다. 또한 데이터 전달 트리를 TCP 혹은 UDP 채널로 구성함으로써, 신뢰적 혹은 비신뢰적 데이터를 전달할 수 있다. 이 프로토콜은 3개의 코어 프로토콜들로 구성되어 있는데, 첫번째는 데이터 분배 트리를 구성하는 YMTP(Yoid Tree Management Protocol)이다. 두번째는 데이터를 분배하기 위한 트리를 관리하는 YDP(Yoid Distribution Protocol) 및 마지막으로 각각의 데이터를 명칭하기 위한 YIDP (Yoid Identification Protocol)로 구성된다.

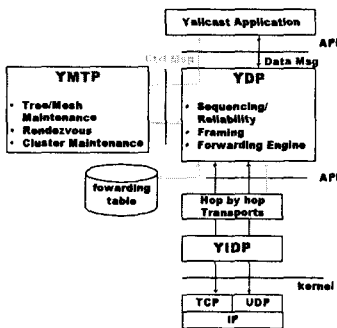


그림 3 Yoid Protocol의 구성

3) structured 방식

Unstructured 방식의 단점은 응용 서비스가 요구하는 사항이 아닌 일반적인 멀티캐스트 데이터 트리를 형성해야 하므로, 모든 노드들은 주기적으로 타 노드들로의 거리를 공고해야 한다. 그렇기 때문에 각각의 노드들은 모든 노드들에 대한 상태를 관리해야 한다. 또한 토폴로지가 변경될 때 빠른 위치 거리 정보를 습득 기능 필요한데 structured 방식은 unstructured방식의 단점을 보완한 메커니즘이다. 이 메커니즘을 사용하는 대표적인 프로토콜들은, CAN, Bayeus, Deaunay Triangulation Overlays, AMLI등이 있는데, 이 중 ALMI는 session controller 와 multiple session member (< 10s) 들로 구성되는 트리를 주 대상으로 삼는 오버레이 기법이다. 트리 구성을 위해서는 중앙집중식 방법을 사용하는데, 세션에 대한 정보를 관리하는 서버를 session controller라고 하며, 세션이 진행되는 도중 멤버들은 응용에 특화된 정보를 이용하여 주기적인 MST를 계산함으로써 효율적인 트리 구성한다. 현재 JAVA로 구현되어 있으며, TCP를 사용하여 신뢰적인 전송 서비스를 제공한다.

● Topology-Aware 방식

앞서 언급한 메커니즘들은 주로 멀티캐스트 전달 트리를 구성하거나 새로운 부모노드로의 천이 등에 초점을 맞추었다. 그런데 오버레이 멀티캐스트는 응용 계층에서 동작하는 메커니즘으로써 IP 계층의 실제 망 경로에 대해서 모르기 때문에 최적

화된 트리를 구성하기 어렵다. 따라서 응용 계층에 하부 네트워크 정보를 얻을 수 있는 방법으로써 제안된 메커니즘으로써 목표는 오버레이 멀티캐스트 트리를 IP multicast 트리에 가장 유사하도록 구성하는 것이다. 이 메커니즘은 노드들간에 거리(RTT)를 측정하고, 측정된 거리에 따라 range를 결정하여 그룹 멤버들에게 보다 효율적인 overlay tree 제공한다. 이 메커니즘을 사용하는 프로토콜들은 Topology-Aware node selection, Topology Aware Grouping들이 있다.

Topology-Aware node selection은 simple, scalable, completely distributed 된 양상을 띄는데, 망의토폴로지를 측정하기 위해서 "well-known Landmark machine" 사용한다. 이 랜드마크들은 거리 측정을 위해 사용되는데, 노드로부터 "well-known Landmark machine"와의 거리(RTT) 측정 후 이의 결과인 레벨 정의 및 특정 그룹(bin)을 선택하여 그루핑을 하는 방법이다.

아울러 이 메커니즘은 타 overlay multicast 기술에도 적용 가능하며 특히 메쉬 형태를 갖는 프로토콜들, 예를 들어 CAN, end-system multicast 등에 쉽게 적용할 수 있다는 장점을 갖는다.

IV. 결 론

이상 오버레이 멀티캐스트에 대한 기술 조사를 하였다. 인터넷 방송, 증권정보, real-time broadcasting 등의 멀티캐스트 응용 서비스가 제공되기 위해서는 빠른 멀티캐스트 기능이 도입되어야 하지만, 앞서 언급한 바와 같이 IP multicast은 그 도입에 따라 해결해야 할 문제점들이 쉽지않은 않은 상태이다. 따라서 망이 멀티캐스트를 지원하지 않을지라도 응용 서비스에 따른 실질적인 멀티캐스트 서비스를 제공할 수 있는 메커니즘이다. 또한 IP multicast가 도입된 이후에 있어서도 Inter-Domain multicast 이슈가 완전 해결될때까지, 지속적으로 사용될 것으로 예측된다.

오버레이 멀티캐스트 기술은 "궁극적인 멀티캐스트 서비스의 촉매제"역할과 "서비스에따른 임시적 솔루션 성격"을 띄고는 있지만, 라우터들이 하드웨어 스위치로 진화하면서 멀티캐스트를 위한 최후의 솔루션으로 제공될 수 있다.

참고 문헌

- [1] Y. Chu, S. G. Rao, and H. Zhang, "A Case for End System Multicast," Jun. 2000.
- [2] D. Pendarakis, S. Shi, D. Verma and M. Waldvogel, "ALMI: An Application Level Multicast Infrastructure," Mar. 2001.
- [3] Y. Chawathe, S. McCanne, and Eric A. Brewer, "RMX: Reliable Multicast for Heterogeneous Networks," Jun. 2000.