

# 확장성 있는 실시간 스트리밍 콘텐츠 전송을 위한 오버레이 멀티캐스트 프로토콜 비교 분석

박진홍<sup>0</sup>, 윤미연, 신용태  
 숭실대학교 컴퓨터학과  
 (elzk<sup>0</sup>, myyoon, shin)<sup>0</sup>@cherry.ssu.ac.kr

## A Comparison of Overlay Multicast Protocol for Scalable Real-Time Streaming Contents Transmission

Jinhong Park<sup>0</sup>, Miyouon Yoon, Yongtae Shin  
 Dept. of Computing, Soongsil University

### 요 약

IP 멀티캐스트의 도입이 기존 인프라의 수정을 수반함으로써 그 실현 가능성이 희박한 가운데, 멀티캐스트의 기능을 응용계층에 둔 응용계층 멀티캐스트에 대한 연구가 활발히 진행되고 있다. 그러나 이러한 연구는 현재 표준을 거치지 못한 상태로 다양한 응용프로그램의 요구를 완벽하게 충족시키지는 못하고 있다. 본 논문에서는 확장성 있는 높은 대역폭을 요구하는 실시간 스트리밍 전송에 적합한 프로토콜을 개발하기 위해 고려해야 할 요소를 도출하여, 각 유형별 대표 프로토콜들을 비교 분석한다.

### 1. 서 론

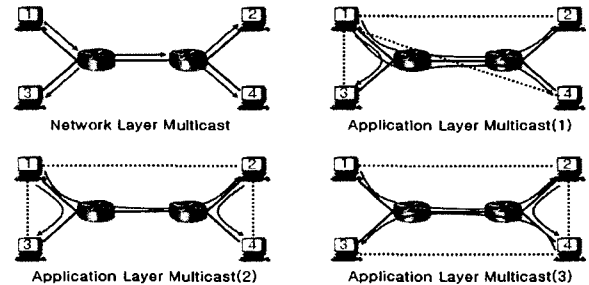
최근 IP 멀티캐스트 도입의 많은 문제로 인해 오버레이 멀티캐스트라 불리는 응용계층 멀티캐스트에 대한 연구가 활발히 진행되고 있다. 기존의 IP 멀티캐스트는 모든 라우터가 그룹의 상태정보를 관리해야 하기 때문에 하드웨어적인 수정이 불가피하고, 신뢰성 및 흐름제어 등을 보장하기 위해 추가의 응용기술이 필요하다. 또한 멀티캐스트 트래픽에 대한 실질적인 과금 모델이 정의되지 못하는 등 여러 가지 문제로 인해 멀티캐스트의 적용이 Mbone상으로 국한되었다. 응용계층 멀티캐스트는 멀티캐스트 기능을 응용계층에 두어 어떠한 하드웨어적인 수정 없이 현존하는 인터넷 인프라 상에서 멀티캐스트를 수행할 수 있도록 설계되었다. 이러한 응용계층 멀티캐스트는 전송 효율에 있어서 순수 IP 멀티캐스트보다 못한 성능을 보이지만, IP 멀티캐스트를 대체할만한 기법으로 급부상하고 있다.

응용계층 멀티캐스트 상에서 콘텐츠 전송을 위한 여러 가지 프로토콜들이 활발히 연구되고 있다[2,3,4]. 이러한 프로토콜들은 그룹관리를 위한 메시지를 먼저 생성하는 트리 우선 접근방식, 데이터 전송을 위한 트리를 먼저 생성하는 트리 우선 접근방식으로 나뉘며, 그 외에 정규화 되지 않은 여러 프로토콜들이 연구 개발 중에 있다. 현재 이러한 응용계층 멀티캐스트 프로토콜들은 표준화 단계를 거치지 못한 상태로 지속적인 연구 활동이 이루어지고 있다. 본 논문에서는 여러 가지 응용계층 멀티캐스트 프로토콜을 분석하여 각 프로토콜의 특징과 제약사항에 대해 분석하고, 분석된 결과를 바탕으로 높은 대역폭을 요구하는 실시간 스트리밍 전송에 타당한 프로토콜 개선 연구 방향을 도출하도록 한다.

본 논문의 2장에서는 여러 가지 응용계층 멀티캐스트 프로토콜의 전송관리 및 그룹관리 측면에 대해 살펴보고, 3장에서는 각 프로토콜의 특징, 장단점 및 제약사항에 대해 분석하여 확장성 있는 높은 대역폭의 실시간 스트리밍 전송 프로토콜 발전 방향을 도출하도록 한다. 마지막으로 결론 및 향후 연구방안에서는 본 연구의 타당성과 향후 연구 방향에 대해 기술하였다.

### 2. 응용계층 멀티캐스트 프로토콜

[그림 1]은 IP 계층 멀티캐스트와 여러 형태의 응용계층 멀티캐스트의 기본 동작을 나타내고 있다. 순수 IP 멀티캐스트의 경우 데이터 패킷이 라우터에서 복사되어 목적지까지 전달되는 것에 반해 응용계층 멀티캐스트는 데이터 패킷의 복제가 종단 호스트들에서 이루어진다. 논리적으로 종단 호스트들은 오버레이 네트워크를 형성하게 되며, 궁극적인 응용계층 멀티캐스트의 목적은 데이터 전송을 위한 효율적인 오버레이를 구축하고 유지하는데 있다. [그림 1]의 응용계층 멀티캐스트 형태에서 점선으로 표시된 부분은 논리적으로 오버레이 상에 노드가 위치하고 있음을 나타내고 있다.



[그림 1] IP 멀티캐스트 및 응용계층 멀티캐스트

응용계층 멀티캐스트 프로토콜은 토폴로지 구성에 따라 트리 우선방식, 트리우선기반 방식, 혼합방식 등으로 구분된다. 트리우선방식의 대표적인 프로토콜로는 NARADA를 들 수 있으며, 트리우선방식의 프로토콜에는 YOIO와 HMTP등이 있다. 혼합방식 및 그 외 규정되지 않은 형태의 프로토콜 들에는 NICE, CAN, SCRIBE, PASTRY등 여러 프로토콜이 존재하지만 각각 특정환경에 종속적이며, 그중 NICE가 가장 보편적으로 알려져 있다. 본 논문에서는 세 방식의 대표적인 프로토콜인 NARADA, YOIO, NICE 프로토콜을 여러 가지 관점의 오버레이 전송 특성을 파악하여 보다 효율적인 프로토콜 제안을 위해 활용하도록 하였다.

\* 본 논문은 한국과학재단 특정기초(과제번호 : R01-2001-000-00362-0(2002) 연구비지원에 의해 수행되었음.

**2.1. Narada[2]**

Narada는 데이터 전송 트리를 구성하기 전에 각 노드 간 메시지를 구성하는 메시지 우선 접근방식의 응용계층 멀티캐스트 프로토콜이다. 멀티캐스트 그룹에 참여하는 모든 노드들에 대해 모든 노드들이 연결되어 있는 메시지를 형성하며, 생성된 메시지는 가입 탈퇴 절차를 통해 형성되고 세분화된다.

**2.1.1. 데이터 전송 트리**

Narada의 트리 생성은 메시지 생성 후 메시지를 이용하여 RSP(Reverse Shortest Path) 트리를 생성하는 2단계 절차를 갖는다. 이때 메시지는 그룹관리 기능을 담당하고, 트리는 데이터 전송경로를 담당하게 된다. 멤버들은 메시지 상에서 거리벡터 프로토콜을 이용하여 다른 멤버들의 경로와 라우팅 값(거리 벡터)을 함께 유지한다. 이러한 정보를 바탕으로 RSP 경로를 측정하여 트리를 구축한다. 라우팅 측정에 사용되는 거리 벡터 프로토콜은 두 노드간의 지연시간을 메트릭으로 사용한다.

**2.1.2. 그룹 관리**

새로운 멤버가 멀티캐스트 그룹에 가입하고자 할 때 Narada 프로토콜은 다음과 같은 절차를 따른다.

1. RP로부터 메시지에 이미 가입되어 있는 그룹 멤버들의 리스트를 얻는다(RP는 멀티캐스트 그룹에 가입하고 있는 모든 멤버들의 상태정보 파악 및 유지)
2. 새로운 멤버는 임의로 멤버들의 일부 집단을 선택하고 이 멤버들의 이웃으로 메시지에 가입을 시도한다(이들 멤버들 중 하나라도 새 멤버를 자신의 메시지 이웃으로 수락한 경우 가입 절차는 성공)
3. 메시지 가입 후 자신의 이웃들과 주기적 갱신 메시지를 주고받아 메시지 상태정보를 갱신한다. 기존의 노드가 해당 그룹에서 탈퇴하고자 할 경우에는 이웃 노드들에게 탈퇴 메시지를 보냄으로써 노드정보를 변경하도록 한다. 메시지 상의 어떤 노드가 결함 발생 시 인접 이웃 노드들은 결함이 발생한 노드의 주기적인 갱신 메시지를 받지 못하게 된다. 이러한 경우 인접 노드들은 결함 발생노드에게 상태 확인 신호를 보내고, 상태 확인 신호에도 결함 노드가 응답하지 않을 경우 결함노드가 더 이상 동작하지 않는 것으로 간주한다. 즉시 다른 이웃노드들에게 이러한 메시지 변화 정보를 전달하여 메시지의 상태를 변경하도록 한다.

**2.2. Yoid[3]**

Yoid는 직접적으로 데이터 전송 트리를 먼저 구축하는 트리 우선 접근방식의 대표적인 프로토콜이다. Yoid는 곧바로 데이터 전송 트리를 생성하므로 트리 생성에 대한 여러 조건(트리 이웃 선택, 자식노드 허용 임계값)들을 직접 제어한다.

**2.2.1. 데이터 전송 트리**

Yoid를 포함한 모든 트리 우선 기반의 프로토콜은 공유트리를 생성하고 각 멤버들은 트리 상에서 자신의 부모를 찾을 의무를 가지며, 만족하는 부모를 찾지 못할 경우, 자신을 새로운 트리의 루트로 선언하고 이러한 사실을 RP에 통보한다. 트리 상의 각 멤버들은 트리 상에서 지원하는 자식 노드의 수를 제한한다.

**2.2.1 그룹 관리**

새로운 멤버가 그룹에 가입하고자 할 경우, 1. 새 멤버가입 희망 호스트는 새로운 부모를 찾기 위해 RP에게 질의를 전송한다 2. 통상적으로 RP는 수신된 질의에 대해 이미 멀티캐스트 그룹에 가입되어있는 멤버 리스트를 제공한다 3. 새 멤버는 이 리스트 상에서 자신의 부모가 될 가능성 있는 호스트를 검색한다. 이때 어떤 멤버 y가 다른 멤버 z의 부모가 되기 위해서는 y가 x의 부모가 됨으로 인해 트리에 루프가 발생하지 않아야 하며, y는 새로운 자식 노드를 가질 수 있도록 자식 노드 허용 임계값에 여유가 있어야 한다. 만약 x가 이러한 조건에 만족하는 부모를 찾지 못할 경우 자신을 공유 트리의 루트로 선언하고 이 사실을 RP에 통보한다. 탈퇴로 인해 트리가 분리될 경우 각 분리된 트리에서 자신을 루트로 선언하는 2개의 호스트가 발생할 수 있다. 이런 경우 RP는 두 파티션의 트리를 하나로 통합하여(파티션 복구) 이를 중재한다. 각 멤버들은 공유트리의

더 좋은 거점 마편을 위해 주기적으로 더 좋은 조건의 부모를 검색하며, 멤버가 그룹 내에서 부모를 변경할 경우 루프 탐지 및 회피 메커니즘을 자체적으로 적용시킨다.

**2.3 NICE[4]**

응용계층 멀티캐스트 프로토콜중의 하나인 NICE 프로토콜은 메시지 우선 접근방식과 트리 우선 접근방식 두 가지 범주에 속하지 않는 혼합형 접근 방식의 응용계층 멀티캐스트 프로토콜이다. NICE는 평면상의 모든 호스트들을 일정 규모로 클러스터링 하여 각 클러스터의 리더들끼리 다시 클러스터링 하는 방식의 계층적 구조를 가진다.

**2.3.1. 데이터 전송 트리**

Nice 프로토콜은 호스트들이 논리적인 계층 구조를 이루도록 트리가 구성된다. 계층 구조는 응용계층 멀티캐스트 전송경로를 의미한다. 대부분의 멤버들은 최하위 계층에 위치하고 단지 일정 상수 수준의 다른 멤버들에 대한 정보를 유지한다. 최상위 계층의 멤버들은 단지  $O(\log N)$  수준의 다른 멤버들에 대한 정보를 유지한다. 각 멤버들은 계층구조 내에서 자신과 가까운 위치에 있는 멤버들의 상태 정보를 가지고 있으며, 멤버간의 거리 측정으로 지연시간 메트릭을 사용한다.

**2.3.2. 그룹 관리**

NICE 프로토콜은 1. 새로운 호스트 가입을 위한 클러스터 할당, 2. 주기적인 클러스터 유지 및 세분화, 3. 호스트 이탈 및 RP 실패로 인한 클러스터 갱신, 이렇게 3가지 주요 컴포넌트를 가진다. 응용계층 멀티캐스트 그룹에 참여하길 원하는 호스트는 가입 절차를 초기화하기 위해 RP로의 접근을 시도한다(RP 호스트는 모든 멤버들을 알고 있다고 가정한다). 새로운 호스트가 멀티캐스트 그룹에 가입하고자 하는 경우

1. 호스트는 RP에게 자신의 가입 요청 질의를 전송한다
2. RP는 해당 호스트에게 최상위 계층의 호스트들을 알려주는 것으로 응답한다
3. 호스트는 최상위 계층의 모든 멤버들 중 자신과 가장 가까운 호스트를 찾기 위해 모든 멤버들에게 질의를 전송한다
4. 계층의 가장 가까운 호스트는 새로운 호스트에게 자신의 하위 계층의 멤버들을 알려주는 것으로 응답한다
5. 새로운 호스트는 하위계층의 멤버들에게 다시 질의를 전송하게 되며 이러한 일련의 행동은 최하위 계층에 이를 때까지 반복된다. 이렇게 하여 새로운 호스트가 속하게 될 클러스터를 지정 해주게 된다.

**3. 프로토콜 분석**

**3.1 고려요소 정의**

각 프로토콜의 특성을 비교 분석하기 위한 척도로 다음과 같은 요소들을 사용한다. 이러한 요소들을 통한 분석으로, 각 프로토콜이 사용될 수 있는 환경을 파악하고, 나아가 확장성 있는 높은 대역폭의 실시간 스트리밍 전송시 제약사항 도출과 개선방향 제시가 가능하다.

- 구성방식 : 그룹관리를 위한 메시지, 전송관리를 위한 트리 형성 우선순위를 나타냄
- 트리형태 : 데이터 전송트리의 형태를 나타내는 것으로 소스 기반 역 최단거리 트리 구축과 공유트리 구축을 구분함
- 지연시간 보장 : 프로토콜의 실시간 스트리밍 전송 유무를 나타내는 지표로 지연시간을 보장하지 않는 경우 실시간 스트리밍 전송이 불가능함을 나타냄
- 확장성 : 형성되는 그룹 크기에 대해 그룹의 규모가 확장될 경우 프로토콜의 성능에 어느 정도 영향을 미치는지 나타냄
- 멤버유지정보 : 그룹내의 멤버들이 유지하고 있어야 하는 정보로 멤버유지정보가 과다할 경우 높은 제어 오버헤드 발생
- 멤버 상태유지 제어 오버헤드 : 멤버 유지정보에 따른 각 멤버상의 오버헤드를 나타냄

[표 1] 프로토콜 별 특성 분석

	NARADA	VOID	NICE
구성 방식	메시 우선 방식	트리 우선 방식	혼합형
트리 형태	역 경로 최단거리 트리	공유 트리	메시형 계층 트리
대역폭 보장 여부	보장하지 못함	보장함	보장하지 못함
지연시간 보장	보장함(실시간 데이터 전송 적합)	보장하지 못함	보장함(실시간 데이터 전송 적합)
확장성	낮음	낮음	높음
그룹멤버가 유지해야 할 정보	그룹에 속한 멤버정보 멤버들간 유니캐스트 라우팅경로	자식노드 정보	자신이 속한 클러스터 내의 이웃 정보
상태유지 제어 오버헤드	$O(N^2)$	$O(\max \text{ degree})$	$O(1)$

3.2 확장성 있는 실시간 스트리밍 멀티캐스팅을 위한 기존 프로토콜 분석

[표 1]은 각 프로토콜의 여러 가지 특성을 분석한 것으로, 세 가지 프로토콜의 여러 가지 특성을 잘 나타내고 있다. Narada는 그룹에 속한 모든 멤버의 정보를 주기적으로 갱신하므로 그에 따르는 상태유지 제어 오버헤드가  $O(N^2)$ 으로 나타난다. 이것은 그룹의 크기(N)가 커짐에 따라 제어 오버헤드가 지수 증가함을 뜻한다. 또한 데이터 전송트리 구축 시 대역폭을 고려하지 않은 최단거리 지연시간만을 메트릭으로 사용하기 때문에 고용량의 데이터 전송에 부적합하다. VOID는 유일하게 어느 정도의 전송데이터에 의한 대역폭을 보장할 수 있으나, 데이터 전송 트리 구축 및 유지에 따르는 관리 오버헤드가 크기 때문에 확장성을 보장할 수 없다. 또한 멤버 가입 시 어떠한 노드의 자식노드으로도 가입할 수 없는 상황에서 발생하는 오버헤드로 인해 실시간성을 지닌 데이터 전송에는 부적합하다. NICE는 몇몇 상위 클러스터 리더를 제외한 모든 노드들이 자신의 클러스터 내 멤버들의 정보만을 유지하기 때문에 멤버간 상태유지 제어 오버헤드가 일반적으로 낮은 편이다. 또한, 계층 구조의 도입으로 그룹의 확장성 문제를 해결하였으며, 지연시간을 보장하기 때문에 실시간성 데이터 전송에도 문제가 없다. 그러나, 대역폭을 보장하지 못하기 때문에 고용량의 실시간 스트리밍 전송에는 적합하지 않다.

능한 소규모 그룹통신 프로토콜이다. Narada에서는 대역폭 문제와 멤버간의 제어 오버헤드로 인해 양방향 TV와 같은 많은 양의 데이터를 전송하고, 높은 대역폭을 필요로 하며, 실시간으로 많은 수의 멤버들에게 데이터를 전송하는 응용프로그램에는 적합하지 않다.

VOID 프로토콜은 트리 우선 기반 프로토콜로 곧바로 전송트리를 생성한다. 기본적으로 메시 구축에 따르는 오버헤드는 없으나, 데이터 전송 트리의 구축 및 유지에 따르는 관리 오버헤드가 매우 큰 편에 속한다. 따라서 VOID는 실시간 요소를 지닌 데이터 전송에는 부적절하지만 지연에 민감하지 않는 높은 대역폭을 필요로 하는 데이터 전송에 적합한 프로토콜이다.

NICE 프로토콜은 멤버들 간의 계층 구조를 형성하여 확장성에 중점을 둔 멀티캐스트 프로토콜이다. 이것은 멀티캐스트 그룹의 멤버가 더 넓은 지역에 분포되어 있을 경우에도 좋은 성능을 보인다. 그러나 기존의 Narada 프로토콜에 비해 그 확장성은 효과적이지만 계층의 수가 제한적이므로 계층이 확장되어 감에 따라 하위그룹이 비대해지는 형태를 이룰 수 있으며, 이러한 현상은 그룹의 가입 및 탈퇴 시 부가적인 오버헤드를 발생시킬 수 있다. 또한 Narada 프로토콜과 마찬가지로 대역폭을 고려하고 있지 않기 때문에 높은 대역폭을 요하는 응용에는 적합하지 않다.

[표 2] 프로토콜 별 장단점 비교

프로토콜	장점	단점
NARADA	메시구축으로 인한 강건한 동적 그룹관리 그룹의 가입 및 탈퇴 간단	실시간 수신자 수가 많은 경우 사용 불가능 지연시간만을 고려한 라우팅 대역폭 소모가 많은 경우 사용 불가능
VOID	메시구축 오버헤드 감소 고용량 대역폭의 보장	트리생성시 제어 오버헤드 멤버 가입절차 복잡 모든 멤버의 자식노드 임계값에 여유가 없을 경우 최악의 지연 초래 실시간 데이터 전송 부적합
NICE	계층적 구조로 확장성 제고 넓은 지역 분산 그룹 형성 $O(\log N)$ 의 짧은 데이터 전송길이	대역폭 소모가 많은 경우 사용 불가능

[표 2]는 지금까지 분석한 결과를 토대로 각 프로토콜의 장단점에 대해 도식화 한 것이다. 이러한 분석결과를 통해 각 프로토콜이 높은 대역폭의 실시간 스트리밍 콘텐츠 전송을 위해 개선해야 할 요구사항을 파악할 수 있다. Narada 프로토콜은 멀티미디어 회의, 가상 교실, 네트워크 게임 등의 소규모 분산 그룹을 위한 프로토콜이다. 소규모 그룹간의 효율적인 멀티캐스트를 제공할 수 있으며, 메시지를 통해 강건한 그룹관리가 가

4. 결론 및 향후 발전방향

본 논문에서는 응용계층 멀티캐스트 프로토콜의 각 유형을 대표하는 세 가지 프로토콜의 동작 절차에 대한 고찰과, 각 프로토콜의 특성 및 장단점 분석을 통해 대용량 실시간 스트리밍 전송을 위한 발전방향을 모색하였다. 각 프로토콜은 현재 가장 활발히 연구되고 있는 대표적인 프로토콜이지만 아직까지 높은 대역폭을 요하며, 널리 분포된 그룹의 확장성을 모두 갖춘 프로토콜에 대한 연구는 미진한 상태이다. 본 연구에서는 그러한 요구조건을 만족할 수 있도록 연구방향을 제시하였으며, 현재 NICE 프로토콜에 확장성을 보장할 수 있는 방안에 대해 지속 연구중에 있다. 향후 본 논문을 바탕으로 더욱 효율적인 프로토콜 개발을 위해 지속적인 연구를 수행 할 예정이다.

참고문헌

- [1] S. Banerjee, B. Bhattacharjee. A Comparative Study of Application Layer Multicast Protocols. In *Submitted for review*, 2002.
- [2] Y.-H. Chu, S. G. Rao, and H. Zhang. A Case for End System Multicast. In *Proceedings of ACM SIGMETRICS*, June 2000.
- [3] Paul Francis. Yoid: Extending the Internet Multicast Architecture, <http://www.icir.org/yoid/docs/ycHtml/ht-mlRoot.html>, April, 2000.
- [4] S. Banerjee, B. Bhattacharjee, C. Commareddy. Scalable Application Layer Multicast. In *Proceedings of ACM SIGCOMM*, Pittsburgh, PA August 2002.