

# 신뢰성을 보장하는 멀티캐스트 서비스를 위한 효율적인 전송 계층 트리 구축 방안

김 은숙\*, 최 중원  
숙명여자대학교 컴퓨터학과

eunsook@cs.sookmyung.ac.kr, choejn@sookmyung.edu

## A Mechanism for Constructing Efficient Transport Layer Trees for Reliable Multicast Service

Eunsook Kim\*, Jong Won Choe  
Dept. of Computer Science, Sookmyung Women's Univ.

### 요 약

인터넷을 통한 그룹 통신이 활발해지면서 신뢰성을 보장하는 멀티캐스트 서비스를 요구하는 응용들이 점점 다양해지고 증가하고 있다. 이에 따라 멀티캐스트에서의 신뢰성을 보장하기 위하여 다양한 기술들이 발표되었으며 특히 트리를 기반으로 오류 및 혼잡 제어를 수행하여 신뢰성 뿐만 아니라 확장성을 보장하는 효율적인 방법에 대한 연구가 지속되고 있다. 본 논문에서는 이러한 트리 기반의 프로토콜에서 효율적이고 동적인 트리 형성 및 관리를 위한 방안을 제안한다. 하나의 멀티캐스트 그룹을 여러 개의 지역 그룹으로 나누어 트리로 구성할 때 지역 그룹에 속할 수 있는 그룹 수신자 수에 임계값을 두고 지역 대표자는 이 임계값이 초과되지 않을 때까지 새로운 참여자를 받아들일 수 있도록 하였다. 또한 새로운 노드가 그룹에 참여할 때 지역 대표자 뿐 아니라 일반 수신자들도 지역 대표자에 대한 정보를 전송해 줌으로써 참여시간의 감소와 그룹 내의 메시지 오버헤드를 줄일 수 있도록 하였다. 동적인 트리 형성 및 관리 과정에서 지역 그룹 대표자에 연결된 노드수가 적고 깊이가 깊어질수록 트리 관리에 관한 오버헤드가 증가하는 문제를 해결하기 위해 간단한 트리 재조정을 수행함으로써 트리에서 발생하는 오버헤드를 감소시킬 수 있도록 하였다.

### 1. 서론

인터넷을 통한 그룹 통신에 대한 수요가 증가함에 따라 다자간 통신 기법인 멀티캐스트 서비스에 대한 수요 및 중요성이 증가하고 있다. 특히 원격 진료나 주식 거래 및 소프트웨어의 분배 등 다양한 분야의 응용들이 신뢰성을 보장하는 멀티캐스트 서비스를 요구하고 있다. 그러나 현재 인터넷에서의 멀티캐스트는 UDP를 기반으로 하는 best-effort 서비스에 의해 수행되기 때문에 전송 계층에서 오류 및 혼잡 제어를 수행함으로써 신뢰성을 보장하는 연구가 진행되었다[2, 3, 5, 6].

이들 기술들은 현재 IETF의 RMT-WG을 통해 수행되고 있는 표준화 작업을 통해 몇 가지 Building Block으로 나뉘어졌으며[9], 이 중 트리를 기반으로 오류 및 혼잡 제어를 수행하여 신뢰성뿐만 아니라 확장성을 보장하는 효율적인 방법에 대한 연구가 진행되고 있다[7, 8].

전송 계층의 트리는 네트워크 계층의 트리와 최대한 부합하면서 별도의 트리를 생성하는 부하를 최소화할 수 있고 동적으로 변화하는 그룹 수신자의 상황을 반영할 수 있도록 그

룹 관리를 보장할 수 있어야 한다.

본 논문은 트리 기반의 신뢰성을 보장하는 멀티캐스트 프로토콜을 위한 효과적인 트리 생성 및 관리 방법에 대하여 기술한다. 현재 대부분의 트리 기반 프로토콜에서 사용하는 ERS(Expanding Ring Search)[3] 방법은 그룹 생성 및 참여 과정에서 광역에 산재한 그룹의 경우 메시지 오버헤드가 커지는 단점이 있다.

그러므로 본 논문에서는 새로운 노드가 그룹에 참여할 때 지역 대표자 뿐 아니라 가장 가까운 위치의 일반 수신자들도 참여 요청에 대해 지역 대표자 정보를 전송해줌으로써 새로운 노드의 그룹 참여를 신속하게 처리하고 메시지 오버헤드를 줄일 수 있도록 하였다. 또한, 지역 대표자 각각에 오직 소수의 노드만이 연결되는 형태가 이어지면서 트리 깊이가 깊어지는 불균형 트리가 생성될 경우 트리에서 발생하는 오버헤드가 증가하는 문제를 해결하기 위해 불균형 트리를 억제할 수 있도록 간단한 트리 재조정 메커니즘을 제안한다.

본 논문은 2장에서 관련 연구를 살펴보고, 3장에서 제안된 트리 구축 방법을 기술하며, 4장에서 본 연구의 결론 및 향후 연구 과제를 살펴본다.

\* 이 연구는 한국과학기술원(KAIST)의 여자대학교 연구기반 확충지원 사업 (1999년도)의 지원을 받아 수행하였습니다.

2. 관련연구

전체 그룹을 트리로 구성하여 여러 개의 지역 그룹 기반으로 멀티캐스트에서의 신뢰성을 보장하는 방법은 한정된 지역의 패킷 손실이 전체 그룹의 네트워크 상태에 영향을 미치지 않으므로 효율적인 오류 처리를 할 수 있다. 또한, 가장 가까운 지역 내의 그룹 관리자에게서 손실 패킷을 재전송 받음으로서 재전송 요청이 송신자까지 전송되는 시간의 손실을 감소시킬 수 있으므로 우수한 확장성과 효율성을 가진다[4].

현재 전송 계층의 트리 생성에는 ERS(Expanding Ring Search) 기법이 대표적으로 사용되고 있다[3, 7]. ERS 방법은 그룹에 참여하고자 하는 노드가 참여 요청을 그룹 내로 멀티캐스트 할 때 네트워크의 부하를 감소시키며 자신과 가장 가까운 그룹 수신자를 찾기 위하여 TTL(Time To Live) 값을 1로 초기화하여 전송한다. 이 범위 내에서 하나의 노드가 참여 요청 메시지를 수신하면 이에 대한 응답을 보내고 그룹에 참여를 하게된다. 만일 요청 패킷에 대해 정해진 시간 안에 응답 메시지를 받지 못하면 TTL 값을 2로 증가시킨 후 다시 참여 요청 메시지를 그룹 내로 멀티캐스트 한다. 이 과정은 응답 메시지를 받거나 TTL 값이 255에 도달할 때까지 계속된다.

이 방법은 IP Multicast의 TTL 정보를 기반으로 하기 때문에 간단하면서도 견고하게 그룹 트리를 생성한다는 장점이 있지만 넓게 산재된 그룹 내에서 수신자 수가 증가함에 따라 네트워크에 멀티캐스트 되는 메시지 오버헤드가 급진적으로 증가한다는 문제가 있다.

3. 전송 계층 트리 구축 및 관리 메카니즘

트리 기반의 신뢰성있는 멀티캐스트 프로토콜은 효율성과 확장성 측면에서 우수한 성능을 보이는 것이 증명되었지만[4, 5] 이를 위한 전송 계층의 트리를 설계하기 위해서는 네트워크 계층의 멀티캐스트 라우팅 트리와 최대한 부합하면서 동적으로 변화하는 그룹 멤버십을 관리하여야 하는 어려움이 있다. 또한 지역 그룹의 대표자의 선정 방법 또한 효율성에 중대한 영향을 미치는 문제이다.

3.1 트리 생성

멀티캐스트 그룹에 참여한 수신자들은 Beacon 메시지를 받거나 데이터 메시지를 수신하면 TTL 값을 변화시켜 자신과 가장 가까운 지역 그룹 대표자를 찾기 트리에 연결된다. 이 때 지역 그룹 대표자에 연결될 수 있는 노드 수는 지정된 임계값<sup>1)</sup>을 넘지 못한다.

각 지역 그룹 대표자는 송신자로부터 받은 메시지의 시간 정보를 통하여 자신의 부모 노드를 선택한다. 각 그룹 대표자들은 자신의 직계 부모와 자식 대표자에 대한 정보만을 기록한다. 송신자의 경우 자신의 직계 자식 대표자 정보만 기록한

다. 이것은 송신자와 각 지역 그룹 대표자가 모든 대표 노드들에 대한 정보를 관리하는데 요구되는 많은 양의 정보 처리와 복잡성을 감소시킨다.

트리 생성의 기본 연산은 아래와 같다.

- **CreatGroup**(Group-id, Root, M<sub>thresh</sub>) : 송신자를 Root로 하고 지역 그룹 참여자 수에 대한 임계값 M<sub>thresh</sub>를 갖는 그룹 생성
- **DeleteGroup**(Group-id) 해당 그룹 id를 갖는 그룹 삭제. 송신자만 수행 가능
- **JoinRequest**(Group-id, NewMember-id, TTL) 자신의 id정보(IP Address, Port No.)와 TTL 정보를 이용하여 그룹에 참여
- **JoinAccept**(Group-id, LGroupHeader-id, Receiver-id, TTL, (Parent-id, TTL<sub>p</sub>)) : 지역 그룹 대표자 (LGroupHeader) 정보와 TTL 정보를 포함하며 대표자의 부모 노드의 정보 (Parent-id, TTL<sub>p</sub>)<sup>2)</sup> 등을 제공하여 지역 트리에 참여
- **LeaveGroup**(Group-id, Receiver-id) : 일반 수신자의 경우 그룹 탈퇴가 자유로우나 그룹 대표자 및 송신자의 경우 임의적인 그룹 탈퇴 불가

3.2 동적 트리 유지 및 관리

새로운 노드가 그룹에 참여하고자 할 때는 자신의 지역 그룹 대표자를 찾기 위하여 TTL 값을 1부터 증가시키며 참여 요청 메시지를 멀티캐스트 한다. 이 때 그룹 대표자는 좀 더 멀리 떨어져 있고 이미 그룹에 참여한 일반 수신자가 해당 메시지를 수신했다면 해당 영역의 지역 그룹 대표자 정보(ip address, TTL<sub>lg</sub>)<sup>3)</sup>를 송신한다. 새로운 노드는 송신받은 지역 그룹 대표자 정보를 이용해 곧바로 유니캐스트 참여 요청을 한다. 만일 요청에 대해 여러 개의 지역 그룹 대표자에 대한 정보를 수신한다면 TTL 값이 가장 작은 대표자를 선택한다. 그림 1에서 이 과정을 알고리즘으로 나타내었다.

```

A member want to join a active group
TTL= 1
while TTL <= 255 or receive JoinAccept
multicast JoinRequest
if a LGH receives it
    if N(LG) < Mthresh // N(LG) :지역그룹 수신자수
        send JoinAccept
if a Receiver receives it
    send LGH_id, TTLlg // LGH :지역그룹대표자
    그림 1 새로운 수신자의 그룹 참여 과정
    
```

일반 수신자는 자신의 지역 그룹 대표자에 연결하면 자신의 대표자 이 외에 그의 부모 노드에 대한 정보(IP주소 TTL<sub>p</sub>)를 저장한다. TTL<sub>p</sub>는 지역 대표자와 부모 대표자와의 TTL 값을 나타낸다.

1) 임계값은 TRAM[8]에서는 32로 지정하였고 [5]에서는 시뮬레이션 결과 50 이내에서 좋은 성능을 보이는 것으로 나타났다. 그러므로 본 연구에서도 50 이내의 범위의 값을 추천한다

2) 추후 자신의 지역그룹대표자(LGH)의 failure 발생 경우 사용. 그림 2 알고리즘에서 설명

3) TTL<sub>lg</sub> : 전송 노드 자신과 지역 대표자와의 TTL 값

이 정보를 이용하여 자신의 대표자가 오류로 인하여 응답 불가 상태가 되었을 경우 새로운 대표자를 신속하게 선정할 수 있다. 그림 2는 이 과정을 알고리즘으로 나타낸 것이다. 일단 자신의 대표자가 예고없이 그룹에서 이탈한 것을 감지하면 수신자는 새로운 대표자를 찾기 위해 TTL값을 다시 1부터 ERS를 실행하는 것이 아니고 부모 노드에게 참여 요청을 하여 세션의 진행을 유지한다. 이 후 수신한 TTL<sub>lg</sub>값을 이용하여 최소 TTL 값을 갖는 새로운 대표자를 선정한다.

$$TTL_{lg} = TTL_{lg} + TTL_p / 2$$

(TTL<sub>lg</sub> : 이전 지역 그룹 대표자와의 TTL 정보)

이것은 초기 트리 형성 과정에서 자신과 최단 거리의 대표자를 선택하였기 때문에 네트워크 상황이 변화했을 지라도 일단 TTL<sub>lg</sub> 범위보다 작은 범위에서 지역 그룹 대표자가 존재할 가능성이 적기 때문이다.

```
A receiver detect failure of LGH
Report LGH_fail into its local group
send JoinRequest to parent_of_LGH
while TTL < TTLp
    TTL = TTLlg+TTLp / 2
// find if there is closer LGH than parent_of_LGH
multicast JoinRequest with new TTL
if receive any JoinAccept
    make new connection with the new LGH
```

그림 2 그룹 수신자가 LGH fail 감지 후 새로운 LGH 설정 과정

그룹이 동적으로 변화하는 과정에서 하나의 지역 그룹 대표자에게 소수의 수신자만이 연결되는 형태가 반복되어 깊이가 깊어지는 트리가 형성될 위험이 있다. 이 경우 트리의 깊이에 따른 부하가 증가하는 특성 때문에 비효율적인 그룹 관리가 될 수 있다. 그러므로 본 논문에서는 지정된 임계값 M<sub>thresh</sub>를 이용하여 간단한 그룹 재조정을 동적으로 수행할 수 있도록 한다. 하나의 지역 그룹 대표자는 자신의 M<sub>thresh</sub>값을 자신의 부모 노드에게 주기적으로 전송하고 초기 트리 형성을 벗어나면 부모노드는 M<sub>thresh</sub>/n<sup>4)</sup> 값 이하인 지역그룹을 통합한다. 그림 3은 이 과정을 나타낸 알고리즘이다.

```
Local Group Tree Merge
if N(LGi) < Mthresh/n and N(LGj) < Mthresh/n //i≠j,
if tree level of i and j are the same
    LGHi becomes new LGH // LGHi : 왼쪽 지역그룹대표자
    merge LGi and LGj
    // LGj 의 모든 노드 LGHi 로 Join
if i is a parent of j
    LGHi becomes new LGH
    merge LGi and LGj
    // LGj 의 모든 노드 LGHi 로 Join
```

그림 3 불균형 트리를 만들 여지가 있는 지역 노드간의 병합 과정

이렇게 국소적으로 그룹을 통합하는 것을 궁극적으로는 전체 트리의 깊이가 깊어지는 것을 감소시킬 수 있게 되어 불균형 트리로 인해 발생하는 오버헤드를 줄일 수 있다.

#### 4. 결론 및 향후 연구 과제

본 논문은 트리 기반의 신뢰성을 보장하는 멀티캐스트 프로토콜을 위한 효과적인 트리 구축 및 관리 방법을 제안했다. 현재 대부분의 트리 기반 프로토콜에서 사용하는 ERS 방법은 간단하면서도 견고하게 그룹 수신자를 연결시켜주는 장점이 있지만 그룹 생성 및 참여 과정에서 광역에 산재한 그룹의 경우 메시지 오버헤드가 커지는 단점이 있다.

이에 본 논문에서는 새로운 노드가 그룹에 참여할 때 지역 대표자 뿐 아니라 지역 그룹 내에서 새로운 노드와 가장 가까운 일반 수신자들도 지역 대표자에 대한 정보를 전송해 줌으로써 그룹 내의 메시지 오버헤드를 줄일 수 있도록 하였다. 또한, 지역 그룹 대표자의 연결 시 불균형 트리를 최대한 억제하기 위하여 간단한 트리 재조정 방법을 제안한다. 성능평가가 수행 중이지만 제안한 방법이 트리에서 발생하는 오버헤드를 감소시킬 것이 예측된다. 현재 보다 효율적인 그룹 대표자 선정 방안과 제안한 모델의 성능평가 작업이 진행중이다.

#### 5. 참고문헌

- [1]. J. C. Lin, S. Paul, RMTP : A Reliable Multicast Transport Protocol, IEEE Infocom 96, March 1996
- [2]. S. Floyd, V. Jacobson, L. Zhang 외, A Reliable Multicast Framework for Light-weight Sessions and Application Level Framing, IEEE/ACM Transactions on Networking, November 1996
- [3]. R. Yavatkar, J. Griffioen, M. Sudan, A Reliable Dissemination Protocol for Interactive Collaborative Applications, Proceedings of ACM Multimedia '96
- [4]. B. Neil Levine, J.J. Garcia-Luna-Aceves, A Comparison of Known Classes of Reliable Multicast Protocols, Proceedings of International Conference on Network Protocol (ICNP-96), 1996
- [5]. M. Hofmann, Enable Group Communication in Global Networks, Proceeding of Global Networking 97, Canada, June 1997
- [6]. T. Speakman, D. Farinacci, S. Lin, A. Tweedly, Pretty Good Multicast(PGM) Transport Protocol Specification, Internet Draft, January 1998
- [8]. M. Kadansky, D. Chiu, J. Wesley, J. Provino, Tree-based Reliable Multicast(TRAM), Internet-Draft, September 1999
- [9]. B. Wetten, L. Vicisano, R. Kermode 외, Reliable Multicast Transport Building Blocks for One-to-Many Bulk-Data Transfer, Internet Draft, June 1999

4) 최적의 n 값 산출이 필요하나 현재 3을 초기값으로 부여함