

오버레이 멀티캐스트 네트워크에서 종단 호스트 멀티캐스트 트리 프로토콜 기법에 관한 연구

노경택*, 이기영**

End-Host Multicast Tree Protocol in Overlay Multicast Networks

Kyung Taeg Rho*, Ki Young Lee**

요 약

다중 유니캐스트 전송기법보다 멀티캐스트 전송기법이 장점이 많은 것이 사실이지만 IP 멀티캐스트 프로토콜은 간편한 관리제어 목적으로 몇몇 네트워크 영역으로 국한된다. 영역간 멀티캐스트 배치는 기술적인 면과 관리적인 면으로 인하여 진척이 느려지고 있다. 이에 대안으로 최근에 오버레이 멀티캐스팅 기술이 제시되었고, 본 논문에서 우리는 종단 호스트 멀티캐스트 트리 프로토콜 (EMTP)을 제안한다. EMTP는 리프 노드에 도달할 때까지 한번에 트리의 두 레벨을 검색하면서 또 한편으로는 가입하고자 하는 노드와 트리상의 가장 가까운 노드를 찾아낸다. 비록 해당 위치의 노드가 허용 가능한 차수가 없을지라도 그 노드를 잠재적인 부모노드로 선택 가능하게 함으로써 차후에 스위칭 횟수를 줄여 트리의 안정성을 높이고 데이터 전송 시간의 감소를 가져온다.

Abstract

While the advantages of multicast delivery over multiple unicast deliveries is undeniable, the deployment of the IP multicast protocol has been limited to network domains under single administrative control. Deployment of inter-domain multicast delivery has been slow due to both technical and administrative reasons. For this reasons overlay multicasting technologies are recently proposed. We propose a End-Host Multicast Tree Protocol (EMTP) that searches two levels of the tree at a time until reaching to a leaf node and can select the nearest node that a new node try to join in the tree. Even if there is no available degree of the desirable node that a new node want to join, selecting the node as a potential parent by EMTP can reduce the number of tree switching with robustness of the tree and bring to reduce the data delivery time.

▶ Keyword : end-host multicast, IP multicast, Potential Parent

I. 서론

멀티캐스팅은 하나의 참가자가 그룹의 다른 참가자들에게 데이터를 효과적으로 전송하기 위한 네트워크 기술이다. IP 멀티캐스팅은 IPv4에서 실행되는 멀티캐스팅이다[1]. IP 멀티캐스팅에서 패킷이 송신자에 의해서 한번 전송되면, 물리적인 네트워크 상에 불필요한 중복 없이 모든 호스트에게 전달된다. IP 멀티캐스팅은 송신자가 데이터를 전송하고 수신자가 전송된 데이터를 받기 위해서 그룹에 가입하기 위한 특별한 주소를 사용한다[2]. 라우터들은 그룹의 상태를 유지하며, 필요한 라우팅을 실행한다. 그러나, 기술적인 문제와 관리적인 문제로 인하여, IP 멀티캐스팅은 인터넷에 널리 보급되지 못하였다[3].

멀티캐스팅을 위한 다른 접근 방식으로 오버레이 멀티캐스팅이 있다. 이 방식은 데이터 전달을 위해서 오버레이 멀티캐스트 트리를 생성하고 유지한다. 이들의 주요 아이디어는 멀티캐스트 그룹의 참가자들이 가상네트워크를 형성하기 위해서 서로 연결하는 것이다. 그 멤버들은 가상의 네트워크 상에서 스스로 멀티캐스트 데이터를 라우팅하고, 이 데이터는 모든 멤버에게 전달된다. 멀티캐스팅의 기능이 라우터에서 중단 호스트로 이동한 것이다. 즉, 기능이 네트워크 계층에서 전송 또는 응용 계층으로 이동한 것이다[5]. 오버레이 멀티캐스팅은 일반적인 유니캐스트 전달 기능을 제외하고 라우터로부터 어떠한 지원도 요구하지 않는다.

오버레이 멀티캐스팅에는 Overcast[4], Narada[6][13], ALMI[7], Yoid[8][14], HMTP[5], NICE[9], CAN[10] 등이 있다.

오버레이 멀티캐스팅은 단점을 지니고 있는데, 이는 IP 멀티캐스팅보다 네트워크 자원을 더 많이 소모한다는 것이고 수 백 호스트 이상으로 확장되기 어렵다. 그러나 이 방식은 중단 호스트에서 실행됨에 따라서 쉽고 빠르게 보급될 수 있다. 오버레이 멀티캐스팅의 확장성과 효율성은 이들의 분산된 트리의 질에 의해서 영향을 받는다. 그러므로, 그룹의 멤버들간에 효과적인 트리를 구성하는 것이 중요하다. 이때 하나의 호스트가 동시에 10개 이상의 연결을 지원하기 위한 대역폭을 지니지 못하기 때문에 트리를 구성할 때는 노드의 제한된 차수를 고려해야 한다[11].

노드의 제한된 차수를 고려한 기존 연구로써 HMTP[5]가 있다. HMTP는 트리 기반의 중단 호스트 멀티캐스팅 프로토콜이다. HMTP는 트리의 총 비용을 줄이면서 모든 호스트마다 제한된 차수를 최대한 유지하기 위해서 가까운 멤버들끼리 밀집한 구조를 형성하기 위해 노력한다. HMTP의 단점은 그룹의 규모가 클 때, 새로운 멤버가 그룹에 가입하는데 시간이 오래 걸리며, 또한 최적의 부모 노드를 선택하려고 노력하지만 실제로 자신과 가장 가까운 잠재적인 부모 노드를 선택하는 것이 어렵다는 것이다. 이러한 이유는 HMTP가 잠재적인 부모 노드를 선택하기 위해서 한번에 트리를 한 레벨씩 검색하기 때문이다.

본 논문의 구성은 다음과 같다. 2장에서는 기존 접근방안과 그에 따른 문제점을 알아보고, 3장에서는 우리가 최근 제안했던 FJM[12]를 보다 개선한 EMTP 알고리즘을 제시한다. 끝으로 4장에서 결론 및 향후계획으로 맺으려고 한다.

II. 관련 연구

중단 호스트 멀티캐스팅 방식의 확장성과 효율은 멀티캐스트 트리의 질에 의존한다. 중단 호스트 멀티캐스트 트리를 만드는 가장 쉬운 방법은 기존 라우팅 알고리즘을 이용한 송신자를 루트로한 shortest-path 트리를 만드는 것이다. 그러나 이 방식은 송신자와 수신자 사이에 유니캐스트 연결로 이루어진 송신자를 루트로한 스타형 토폴로지를 생성한다. 이러한 트리는 송신측 근처에 심대한 확장성과 견고함의 문제를 내포한다. 트리 생성기법외에 중단 호스트 멀티캐스팅 방식은 동적으로 변하는 그룹 멤버십을 다뤄야 하고, 멤버를 발견하는 기법도 요구된다.

HMTP에서 각 그룹멤버는 사용자영역에 디몬 프로세스(호스트 멀티캐스트 에이전트)가 존재한다. 이 디몬 프로그램은 중단 호스트에서 호스트 멀티캐스팅 기능을 지원한다. 모든 멤버노드들은 HMTP를 이용하여 양방향 공유트리를 구성한다. 공유트리상에서 두 멤버가 터널로 연결되어 있다면 서로 이웃상태인 것이다. 특정노드는 트리의 루트로 할당되고 그 밖에 모든 노드는 하나의 이웃을 부모로 한다. 부모노드가 아닌 이웃들은 자식노드이다. 루트로부터 각 멤버까지 트리를 따라 유일한 하나의 루프가 존재하지 않는 패스가 존재한다. 이 패스의 리스트를 루트패스라 불린다.

일단 하나의 멤버가 트리에 접속하면 그것은 멀티캐스트 전송에 참여한다. 이웃으로부터 전달된 패킷은 복제되어 다른 모든 이웃으로 전달된다. 결국은 트리상에 모든 노드로 전달되어진다. 모든 멀티캐스트 그룹은 Host Multicast Rendezvous Point (HMRP)를 가진다. 새로운 멤버 N는 HMRP에 질의를 보내서 공유된 트리의 루트노드를 알게된다. 새로운 멤버는 루트로부터 시작해서 트리의 각 레벨에서 노드 자신과 가까운 노드 Z를 찾는다. 만약 노드 Z의 자식노드 수가 허락된 트리의 차수제한보다 작다면, 노드 N은 Z의 자식 노드로써 그룹에 가입한다. 만약 그렇지 않다면 다음레벨에서 위와 같은 절차를 반복하며, 노드 Z의 자식노드들 중에서 잠재적인 부모 노드를 찾기 위해 노력한다.

HMTMP는 다음과 같은 문제점을 지니고 있다. 첫째, HMTMP는 트리의 매 레벨에서 한 번씩 검색을 수행하기 때문에 그룹의 멤버가 많을 때 가입 지연 시간이 길다는 것이다. 이러한 긴 가입 지연 시간은 그룹의 멤버가 빠르게 변할 때 데이터 전송에 영향을 끼치게 된다. 둘째, HMTMP에서 새로운 노드는 최적의 부모 노드를 선택하려고 노력하지만 실제로 자신과 가장 가까운 부모노드를 선택하는 것은 어렵다. 이유는 트리의 각 레벨을 검색하여 자신과 가까운 노드 중에서 차수가 허용되는 노드에게 가입을 수행하기 때문이다. 이는 새로운 노드가 그룹에 가입한 이후에 트리의 최적화를 위해서 트리스위칭[11]을 발생시킨다. 트리 스위칭은 네트워크 트래픽 상태를 변화시키며, 스위칭 동안에 패킷손실 및 중복을 야기 할 수 있기 때문에 최소화해야 하며, 이를 위해서는 새로운 노드는 좀 더 최적의 부모 노드에 선택하여 그룹에 가입을 해야 한다.

HMTMP와 같이 <그림 1>의 한 레벨 방식은 기존 멤버의 부분목록으로부터 가장 가까운 멤버를 선택함으로써 그룹에 가입한다. 그 결과 생성된 트리는 전체적으로 랜덤하지 않고 또한 모든 그룹 멤버의 목록을 유지할 필요가 없지만 한 레벨씩 검사방식으로 인하여 가입시간이 길어지는 단점이 있다. 이러한 HMTMP의 문제점을 개선 방안으로 FJM (Fast Join Mechanism)을 제시하였다[12]. FJM은 잠재적인 부모 노드를 찾기 위해서 <그림 1>의 두 레벨 검색 방식같이 트리를 한번에 두 레벨씩 검색한다. 주요한 아이디어는 그룹을 유지하기 위해서 평소의 메시지 교환 회수는 그대로 유지하면서, 새로운 멤버가 그룹에 가입하려고 할 때 메시지의 양을 증가시켜서 빠르게 가입시키고, 메시지의 양이 증가한 만큼 더 많은 노드를 검색함으로써, 좀 더 최적화된 부모 노드를 선택할 기회를 부여하여 새로운 노드가 그룹에 가입한 후에 트리의 스위칭 횟수를 줄이는 방식

이다. 이를 위해서, 기존의 멤버가 Request 메시지를 그들의 자식 노드에게 전달하는 방식을 사용한다.

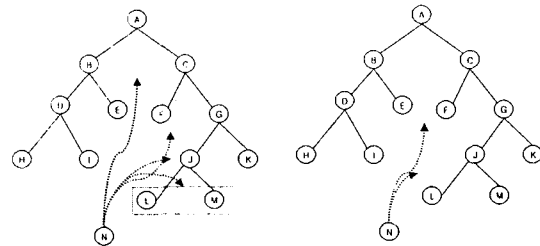


그림 1. 한 레벨 검색방식과 두 레벨 검색방식의 비교
Fig 1. Comparisons of one-level search method and two-level search method

III. 제안 방안

FJM은 HMTMP 보다 빠른 가입과 더불어 보다 최적화된 부모 노드를 선택하는 장점이 있지만, 새로운 노드가 가입하려 할 때 루트에서 시작해서 두 레벨 검색방식을 이용하여 최종 리프 노드에 도착하던지 아니면 중간에 차수가 허용되는 노드를 만날 때 가입하게 됨으로써 자신과 가장 가까운 노드에 가입하지 못 함으로써 트리의 질이 떨어지는 문제점을 가지고 있다. 이에 대한 개선 방안으로 End-Host Multicast Tree Protocol (EMTP)에서 필요한 메시지 타입에 대한 설명과 함께 EMTP 기법을 제시한다.

표 1. 요청 메시지 및 응답 메시지
Table 1. request message and response message

이름	의미	메시지 수신 후 동작
oMSG	최초의 요청	<ul style="list-style-type: none"> • 유효 차수가 있는 경우 : acceptMSG를 새로운 멤버에 전송 • 유효 차수가 없는 경우 : fMSG를 자식 노드들에게 전송
fMSG	전달된 요청	<ul style="list-style-type: none"> • 유효 차수가 있는 경우 : acceptMSG를 새로운 멤버에 전송 • 유효 차수가 없는 경우 : addrMSG를 새로운 멤버에게 전송
addrMSG	가입불능	목록의 노드의 rtt를 측정하여 가장 가까운 노드에게 oMSG를 전송
acceptMSG	가입가능	새로운 멤버는 acceptMSG를 보낸 노드와 lolNFO에 존재하는 노드 중에서 가까운 노드에 가입
lolNFO	거리정보	루트 패스를 통해서 새로운 멤버와의 최단거리노드

EMTP에서 사용하는 메시지에 대한 설명 및 동작방식을 (표 1)에서 볼수 있다. 새로운 멤버는 oMSG 메시지를 전송하여 가입이 가능한지 여부를 알아본다.

oMSG 메시지를 받은 멤버에게 허용된 차수가 없으면, fMSG메시지로 변환하여 자신의 자식 노드에게 메시지를 전달한다. 응답으로써 addrMSG 메시지를 받은 새로운 멤버는 받은 노드의 목록들 중에서 자신과 가장 가까운 노드를 검색한다.

만약 새로운 멤버가 acceptMSG 메시지를 받았다면 acceptMSG를 보낸 노드와 loINFO에 존재하는 노드 중에서 가까운 노드에 가입한다.

1. EMTP의 가입 기법

- 1.1 가입하고자 하는 새로운 멤버 N은 EMRP (End-Host Rendezvous Point)에게 질의를 보내서 공유된 트리의 루트를 알아낸다. minRTT에 무한대 값을 저장한다.
- 1.2 노드 N은 루트 노드를 잠재적 부모 노드로 지정하고, 가입 메시지로 oMSG 메시지를 루트 노드에게 보낸다.
- 1.3 oMSG 메시지를 받은 노드에게 허용된 차수가 없다면, 전달 fMSG 메시지로 변환해서 자신의 자식 노드들에게 전달한다.
- 1.4 fMSG 메시지를 받은 노드는 자신과 자식들의 주소를 addrMSG 메시지에 담아서 노드 N에게 보낸다.
- 1.5 노드 N이 addrMSG 메시지를 받으면, 받은 목록의 노드들과의 거리를 ping 메시지를 통한 rtt 값을 알아내고, 그 중에서 가장 가까운 거리를 지닌 노드의 최소 rtt 값이 지금까지의 minRTT와 비교하여 작다면 minRTT에 저장하고 또한 그 노드의 위치정보를 loINFO에 기록해둔다. 그리고 지금까지의 잠재적 부모 노드를 추적하기 위해 스택 S에 저장한다. 또한 잠재적 부모 노드에게 oMSG를 보내고 스택 3로 이동한다.
- 1.6 만약 oMSG 메시지를 받은 노드 Z에게 허용된 차수가 있다면, acceptMSG를 노드 N에게 전송하고, 노드 N은 노드 Z와 지금까지의 minRTT 값을 지닌 loINFO 값에 해당하는 노드 X와의 거리를 비교한다.
- 1.7 만약 노드 Z와의 거리가 짧다면 노드 Z의 자식 노드로 가입하고, 그렇지 않다면 노드 X를 자신의 부

모 혹은 자식 노드로 구성하여 트리를 적절히 재배치한다.

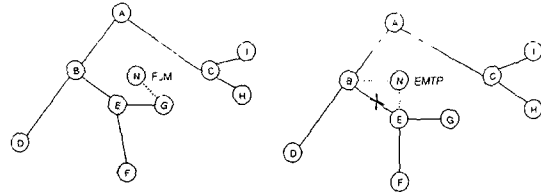


그림 2. FJM과 EMTP기법을 이용한 새로운 노드 N 가입 경우
Fig 2. In cases a new node N try to join a tree using FJM and EMTP

(그림 2)의 이차원 트리의 경우에 있어서 새로운 노드 N가 노드 B와 노드 E 사이에 위치하여 있을 때 노드 N은 FJM 기법을 적용할 때 루트노드 A에서 시작하여 노드 B, E를 거치면서 허용된 차수가 없기에 결국은 노드 G, F 중에서 가까이 존재하는 노드 G를 부모노드로 선택한다. 모든 노드는 부모노드보다 루트에서 멀어져야 안정적 트리구조를 갖는 것에 위배된다. 한편 제안된 EMTP 기법을 적용할 때 노드 B의 허용차수가 없다 하더라도 새로운 노드 N은 노드 B를 부모노드로 선택하고, 노드 E를 자신의 자식노드로 선택함으로써 트리의 레벨이 높아질수록 루트에서 멀어지도록 함으로써 전체적으로 데이터 전송지연을 줄임과 동시에 보다 견고한 트리를 만들기 위한 트리 스위칭 횟수를 줄일 수가 있다.

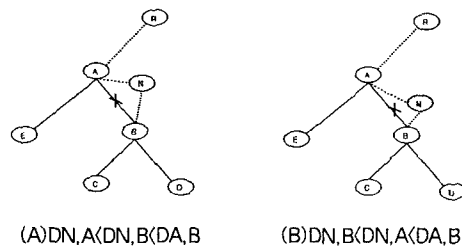


그림 3. 새로운 노드 N가 노드 A를 부모노드로 선택하는 경우
Fig 3. In cases a new node N select a node A as a parent node

또한 (그림 3) 첫 번째 트리와 같이 가입하고자 하는 노드 N이 노드 A와의 거리 DN,A가 최소 rtt값을 갖는 경우와 두 번째 트리와 같이 노드B와의 거리 DN,B가 최소 rtt 값을 갖는 경우에 있어서, 다른 트리를 이용한 오버레이 멀티캐스트 기법 즉, HMTTP와 FJM 경우 노드 N은 노드 D에 가입하게 되어 다시 루트노드 방향으로 돌아감으로써 전송지연을 초래하게 되고, 결국은 나중에 트리의 질을 높이기 위해 트리 스위칭을 필요로 하게 되는 단점이 있다.

이와 달리 노드 A를 부모 노드로 선택하고 노드 B를 자식노드로 배치하여 최소경로를 지닌 트리에 가깝도록 트리를 구성한다.

EMTP는 HMTP와 FJM처럼 각 노드는 루트노드까지의 패스정보를 보유하고 있어서 그룹에 속한 모든 노드에 관한 정보를 유지할 필요가 없다. FJM의 장점을 유지하면서 단점을 보완하여 공유된 트리의 루트노드에서 시작하여 새로운 노드가 추가될 때 가장 가까이 존재하는 노드를 부모 또는 자식 노드로 선택하면서 자식 노드가 부모 노드보다 루트에서 멀어지도록 하는 안정된 방사형 트리 구조를 유지한다. 따라서 두 레벨 검색방식으로 인한 probe 메시지가 증가되는 점도 있지만 보다 빨라진 검색속도와 더불어 트리 스위칭 횟수를 줄여주는 질 높은 트리를 만들어 낸다.

IV. 결론 및 향후 계획

우리는 두 레벨에 걸친 검색을 이용하여 기존 공유트리의 적절한 위치에 새로운 노드를 가입시키는 종단 호스트 멀티캐스팅 트리를 만드는 기법을 제시하였다. 각 멤버노드는 이웃한 자신의 부모노드와 자식노드들의 정보를 가지고 있으며, 새로운 노드는 트리에 가입하기 위해 전체 그룹 멤버가 아닌 해당 부분 그룹 리스트에 속한 노드들에게 ping을 이용한 가입하기 적합한 가까운 노드를 찾아내고 그 노드의 자식 노드로서 가입한다. 우리가 FJM에서 보여준 장점을 유지하면서 트리 스위칭 횟수를 줄이려 한 점이 본 논문의 목적이다. 향후 비록 허용 가능한 차수가 없더라도 그 위치가 새로운 노드가 가입할 적당한 위치일 경우 그곳에 존재하는 노드를 부모노드로 선택하고 그 노드의 자식노드를 새로 가입할 노드의 자식노드로 두는 기법인 EMTP를 FJM과 HMTP에 대하여 시뮬레이션을 통한 비교 분석을 하려고 한다.

참고문헌

- [1] S. Deering and D. Cheriton, "Multicast routing in datagram internetworks and extended LANs", ACM Transactions on Computer Systems, 8(2):85110, May 1990.
- [2] D. Thaler, etc., "The Internet Multicast Address Allocation Architecture", IETF RFC 2908, Sep. 2000.
- [3] C. Diot et al., "Deployment Issues for the IP Multicast Service and Architecture", IEEE Networks magazine's Special Issue on Multicast, Jan. 2000.
- [4] John jannotti, David K. Gifford, Kirk L. Johnson, M. Frans Kaashoek, and James W. O'Toole, Jr., "Overcast: Reliable multicasting with an overlay network", Proceedings of the Fourth Symposium on Operating Systems Design and Implementation, pages 197212, October 2000.
- [5] B. Zhang, S. Jamin, and L. Zhang, "Host multicast: A framework for delivering multicast to end users", In Proceedings of IEEE Infocom, June. 2002.
- [6] Yang-hua Chue, Sanjay G. Rao, and Hui Zhang, "A case for EndSystem multicast", Proc. of ACM IGMETRICS'00, pages 112, June 2000.
- [7] Dimitrios Pendarkis, Sherlia Shi, Dinesh Verma, and Marcel Waldvogel, "ALMI: an application level multicast infrastructure", 3rd USENIX Symposium on Internet Technologies and Systems, March2001.
- [8] Paul Francis, "Yoid: extending the Multicast Internet Architecture", Technical report, NTT, April 2000.

[9] S. Banerjee, B. Bhattacharjee, and C. Kommareddy, "Scalable application layer multicast", In Proceedings of ACM Sigcomm, August 2001.

[10] S. Ratnasamy, P. Franscis, M. Handley, R. Karp, and S. Shenker, "A scalable content-addressable network, In Proceedings of ACM Sigcomm, August 2001.

[11] D. Helder and S. Jamin, "End-host multicast communication using switchtrees protocols", Global and Peer-to-Peer Computing on Large Scale Distributed Systems, 2002

[12] 조성연, 노경택, 박명순 정보처리학회논문지 C 제 10-C 권 제 5호(2003년 10월) p625~p634 오버레이 멀티캐스팅에서 트리의 스위칭을 고려한 빠른 멤버 가입 방안에 관한 연구

[13] Y-H. Chu, S. G. Rao, S. Seshan, and H. Zhang, "Enabling Conferencing Applications on the Internet using an Overlay Multicast Architecture", In Proceedings of ACM SIGCOMM, August 2001.

[14] P. Francis, "Yoid: extending the Multicast Internet Architecture", 1999. White paper <http://www.aciri.org/yoid/>.

저자 소개



노 경 택

1986년 중앙대학교 컴퓨터학과 이
학사

1989년 뉴저지공대 컴퓨터학과 이
학석사

1989년 사우스캐롤라이나대학교
컴퓨터학과 박사수료

1999년 고려대학교 컴퓨터학과 박
사수료

현 재 서울보건대학 인터넷정보
과 교수

<관심분야> 멀티캐스팅,
모바일컴퓨팅, 웹서버시스템



이 기 영

현 재 서울보건대학 인터넷정보
과 교수

한국OA학회 논문지 제3권 제4호
참조