

인터넷 멀티캐스트 라우팅에서 QoS를 지원하는 공유 트리 접근법

박세훈* (shpark@archi.snu.ac.kr), 안상현** (ahn@venus.uos.ac.kr)

* 서울대학교 컴퓨터공학과

** 서울시립대학교 전산통계학과

A QoS Supporting Shared Tree Approach for The IP Multicast Routing

Se-Hun Park*, Sanghyun Ahn**

* Dept. of Computer Engineering, Seoul National Univ.

** Dept. of Computer Science & Statistics, Univ. of Seoul

요약

인터넷 멀티캐스트 라우팅을 위해 PIM, CBT, DVMRP 등의 여러 가지 프로토콜이 연구되고 있다. 본 논문에서는 이와 같은 멀티캐스트 라우팅에서 QoS를 보장하기 위한 새로운 방법을 제시한다. 이 방법은 멀티캐스트 라우팅의 두 가지 접근법 중 공유 트리 접근법에서 지연시간에 대한 QoS를 보장하는 방법으로 PIM-SM을 개선한 형태를 취한다. 동적인 멤버십에 의해 주어진 지연시간 상한값을 만족하지 않는 상황이 발생하면 공유 트리의 RP(Rendezvous Point)를 다시 선택하고 트리를 재구성함으로써 QoS를 보장하는 방법을 사용하며, 시뮬레이션을 통해 이 방법의 적절성을 확인하였다.

1. 서론

인터넷에서 멀티캐스트 라우팅을 지원하기 위한 연구는 주로 두 가지 접근법을 통해 진행되고 있다. 첫 번째는 밀집 방식(dense mode) 접근법으로 멀티캐스트 그룹의 구성원들이 밀집되어 분포되어 있고 대역폭이 충분한 경우에 적절한 접근법이다. 그룹 내의 각 송신자에 대해 이 송신자를 루트로 하는 별개의 트리를 설정하기 때문에 송신자-기반(source-based) 트리 접근법, 또는 data-driven 접근법이라고도 불린다. 이 접근법의 대표적인 프로토콜로는 DVMRP(Distance Vector Multicast Routing Protocol)[1], PIM-DM(Protocol Independent Multicast - Dense Mode)[2] 등이 있다. 두 번째는 희소 방식(sparse mode) 접근법으로 멀티캐스트 그룹의 구성원들이 희소하게 분포되어 있고 대역폭이 충분하지 않은 경우에 적절한 접근법이다. 공유 트리를 가지고 각 수신자들이 공유 트리에 join하는 방법으로 트리를 설정하기 때문에 공유트리 접근법, 또는 수신자-기반(receiver-initiated) 접근법이라고도 한다. 이 방법의 대표적인 프로토콜로는 CBT(Core-Based Tree)[3], PIM-SM(Protocol Independent Multicast - Sparse Mode)[4] 등이 있다.

이 두 가지 접근법 중 현재는 희소 방식 접근법에 연구가 집중되고 있는데, 이는 밀집 방식 접근법이 가지는 단점들 때문이다. 밀집 방식 접근법은 각 송신자마다 distribution 트리가 존재하므로 각 구성원들이 가져야 할 정보가 많고, 또한 트리를 구성하기 위해 플러딩 혹은 브로드캐스팅에 의존한다. 따라서 구성원이 많아지고, 망이 커질수록 비효율적이다. 그러나 희소 방식 접근법은 하나의 공유 트리만 유지하므로 라우터가 유지해야 할 정보가 적고, 또한 트리를 구성하는 데 플

러딩과 같이 망 자원을 많이 쓰는 방법을 쓰지 않는다. 따라서 보다 확장이 용이(scalable)하고 현재의 인터넷 환경에 적절하다. 그러나 희소 방식 접근법도 데이터 전달이 항상 공유 트리의 루트를 지나가야 하므로 루트가 병목으로 작용하게 되는 단점이 있다.

인터넷 멀티캐스트 라우팅에서 최근의 주제는 QoS를 지원하는 것이다. 다른 인터넷 응용들과 마찬가지로 대부분의 인터넷 멀티캐스트 라우팅의 응용들이 QoS 지원을 요구하고 있다. 예를 들어 화상회의와 같은 응용에서는 모든 종단간의 최대지연이 어느 수준 이하이기를 원한다.

본 논문은 인터넷 멀티캐스트 라우팅의 두 가지 접근법 중 희소 방식 접근법, 즉 공유 트리 접근법을 기반으로 한 프로토콜에서 QoS를 지원하는 방법을 제시한다. PIM-SM 프로토콜을 기반으로 하여 종단간 최대지연에 대한 QoS 요청을 해결하는 개선된 프로토콜을 제시한다. 이 개선된 프로토콜은 QoS 요청에 대해 능동적으로 대처함으로써 QoS 요청을 반드시 해결하는 것을 목표로 한다.

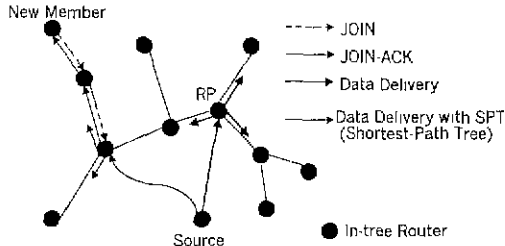
2장에서는 관련된 연구들을 개괄적으로 설명한다. 3장에서는 본 논문의 핵심적인 제안들을 상세하게 설명하고 시뮬레이션의 결과를 간단하게 제시한다. 그리고 4장에서 결론을 맺는다.

2. 관련 연구

PIM-SM은 RP(Rendezvous Point)를 루트로 하는 공유 트리를 이용하여 멀티캐스트 데이터를 배포한다. 멀티캐스트 그룹에 새로이 참여하는 구성원(실제로는 구성원이 연결된 서버망의 DR(Designated

Router))은 미리 알려진 RP에게 Join 메시지를 보내어 참여를 알린다. Join 메시지는 유니캐스트 메시지로써 전달되며, RP로 전달되는 도중에 이미 공유 트리의 일원인 라우터에 도달하면 더 이상 전달되지 않는다

Join 메시지를 받은 라우터는 Join에 대한 승인메시지를 되돌려 보내고 이를 새 구성원이 반응으로써 Join 과정은 완성된다. 멀티캐스트 데이터를 보내려는 송신자는 그 데이터를 유니캐스트 메시지로 캡슐화하여 RP에 보내고, 이를 받은 RP는 데이터를 역캡슐화하여 트리 전체에 전달한다[4]. PIM-SM의 동작에 대한 간단한 개념도는 [그림 1]과 같다.



[그림 1] PIM-SM의 동작

위와 같은 과정에서 PIM-SM은 QoS에 대한 고려를 전혀 하지 않는다. 본 논문은 이러한 PIM-SM을 QoS 제공을 위해 개선하는 방법을 제안한다.

본 논문과는 다른 방식으로 인터넷 멀티캐스트 라우팅에서 QoS를 지원하기 위해 제안된 프로토콜이 몇 가지 있다. 그 중 한가지인 QoS-MIC(Quality of Service sensitive Multicast Internet protoCol)[5]는 루트가 없는 공유 트리를 사용한다. PIM-SM의 RP나 CBT의 core 같은 루트 역할의 라우터가 없는 대신 관리자(Manager) 라우터가 있어서 그룹을 관리하고 구성원들의 Join에 관여한다. QoS 지원을 위해 각 연결에 대해 대체 경로를 제공하고, 이를 위해 두 가지의 Join 방법을 제공한다. 또한 필요할 경우 특정 송신자에 대해서 PIM-SM에서처럼 송신자-기반 트리로 전환도 가능하다. 그러나 Join 과정의 오버헤드가 크고, Join에 걸리는 시간이 길다는 단점이 있다.

다른 방법인 QoS Extension to CBT[6]는 QoS를 지원할 수 있도록 기존의 CBT를 확장한 것이다. 다양한 QoS 요청을 처리할 수 있고 CBT 외의 다른 멀티캐스트 라우팅 프로토콜에도 적용 가능하다. 그러나 Join 시에 주어진 QoS를 만족하지 못하는 경우 무조건 Join을 거부하며, 공유 트리를 유지하기 위해 라우터가 가지고 있어야 할 상태가 많아 확장성(scalability)이 떨어진다.

3. QoS를 지원하는 멀티캐스트 공유 트리 구조

본 논문에서 고려하는 QoS 요구사항은 중단간 지연시간에 집중 될 것이다. 화상회의와 같은 멀티캐스트의 주된 응용들에서 중요한 것은 송신자의 영상 및 음성을 수신하는 수신자들이 거의 비슷한 시간대에 보고 듣는 것이다. 이때 다른 수신자보다 좀 더 빨리 수신하는 것은 별 문제가 되지 않으나, 다른 수신자들에 비해 훨씬 나중에 수신하는 것은 문제가 된다. 이 문제에 있어서 QoS의 다른 요소(metric)들도 이에 영향을 끼치기는 하지만, 직접적인 영향을 끼치는 것은 지연 시간이다. 따라서 본 논문의 목표는 다음과 같다

목표: 멀티캐스트 공유 트리에서 모든 두 구성원 사이의 지연시간이 항상 응용에 의해 미리 결정된 상한값 내에 있다.

이 상한값을 초과하는 구성원 쌍이 발견되면 멀티캐스트 트리를 수정함으로써 문제를 해결한다

3.1 가정

문제 해결을 위해 다음과 같은 몇 가지 가정을 하였다

- 모든 송신자는 곧 수신자로서 공유 트리의 구성원이다.
- 지연시간의 척도는 거리(distance), 즉 경로의 길이를 사용한다. 이는 구현의 편이를 위한 것으로, 망에서 각 링크의 지연시간을 구하는 방법을 제공한다면 쉽게 실제 지연시간으로 수정할 수 있다
- RP가 될 수 있는 라우터, 즉 candidate-RP(C-RP)가 망에 다수 존재한다. RP가 되기 위해서는 일반 라우터에 비해 좋은 성능을 가져야 한다.

3.2 QoS를 지원하는 공유 트리 접근법

사전에 그룹에 참여할 구성원들과 망의 구성을 알 수 있다면 RP 선정 시에 QoS 요구사항에 직결한 라우터를 RP로 선정함으로써 QoS 문제를 해결할 수 있을 것이다. 그러나 동적 멤버십 하에서는 사전에 정보를 알기가 어려우므로 상황에 따라 RP를 다시 선택하여 트리를 재구성하는 방법을 선택하였다.

각 트리내 라우터(in-tree router)들은 자신으로부터 하위 단말노드까지의 지연시간 값들 중 최대의 값을 유지한다. 구성원의 Join/Leave에 따라 각 라우터들이 유지하고 있는 최대지연시간 값은 갱신되고, 만약 이 갱신된 값의 조합이 상한값을 벗어나게 되면 트리를 재구성한다 따라서 각 트리내 라우터들이 PIM-SM에서 필요로 하는 상태들 외에 추가로 유지해야 하는 상태는 다음과 같다.

$d_{max}(u)$: 트리내 u 에서 하향(downstream) 인터페이스로 도달 가능한 가장 먼 단말 라우터까지의 지연시간

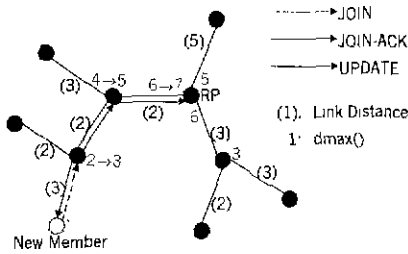
단 RP는 $d_{max}(root)$ 대신 다음의 상태들을 유지한다.

$d_{max,i}(root)$ 하향 인터페이스 i 로 도달 가능한 단말 라우터중 가장 먼 것까지의 지연시간

3.2.1 새로운 구성원의 Join

지연시간에 대한 QoS를 지원하기 위해 기존 PIM-SM의 Join 방법은 다음과 같이 변경된다

새로이 Join하고자 하는 구성원은 RP를 향해 Join 메시지를 보낸다. 이 메시지가 RP로 향하는 중간에 트리내 라우터에 도달하면 트리내 라우터는 자신의 $d_{max}(i)$ 값을 갱신하고, Join-ACK 메시지를 되돌려 보내며, $d_{max}(i)$ 값이 증가한 경우 Join-Update 메시지를 상향으로 내 보낸다. Join-ACK 메시지를 받은 라우터들은 트리내 라우터로 전환되면서 $d_{max}(i)$ 값을 가지게 되고 새로운 구성원을 포함하는 트리를 구성하게 된다. 이와 병행하여 Join-Update 메시지는 계속 RP로 향하



[그림 2] 새로운 구성원이 Join하는 경우에 대한 예

면서 중간의 라우터들의 $dmax()$ 값을 갱신하고, 만약 갱신이 필요 없을 경우 전송이 중단된다 [그림 2]에 위에 설명한 Join 과정에 대한 간단한 예가 있다

이를 위해 Join 및 Join-Update 메시지는 전송된 경로의 지연시간 값을 알아낼 방법을 제공해야 한다. 이는 IP 헤더의 TTL(Time To Live)을 이용하여 쉽게 구현할 수 있다.

3.2.2 구성원의 Leave

그룹에서 떠나고자 하는 구성원은 Leave 메시지를 부모 라우터에 보낸다 부모 라우터에 더 이상의 구성원이 없으면 부모도 Leave 메시지를 보낸다 이는 떠나려는 구성원 외에 다른 구성원이 있는 라우터 까지 진행되고, 이 라우터는 바로 하위의 트리내 라우터들에게 직접적인 질의를 통해 $dmax()$ 값을 갱신한다. 이때 $dmax()$ 값이 바뀌게 되면 Leave-Update 메시지를 상향으로 전송한다 이 Leave-Update 메시지는 계속 RP로 향하면서 직접적인 질의를 통해 $dmax()$ 값을 갱신하고, 만약 갱신이 필요 없을 경우 전송이 중단된다.

이를 위해 Leave 및 Leave-Update 메시지는 Join의 경우와 마찬가지로 전송된 경로의 지연시간 값을 알아낼 방법을 제공해야 한다.

3.2.3 RP 재선택 및 트리 재구성

새로운 구성원의 Join에 의해 최대 지연시간의 가능한 조합 중에서 상한값을 초과하는 조합이 있음을 RP가 발견하게 되면 RP를 다시 선택하여 상한값을 만족하는 공유 트리로 변경한다 이 과정은 다음과 같다. 여기서 구성원 라우터는 자신의 서브망에 구성원을 가진 라우터를 말한다.

- 1) RP는 ReSel-Req 메시지를 멀티캐스트한다.
- 2) ReSel-Req 메시지를 받은 구성원 라우터들은 더 이상의 전송을 중단하고 각 C-RP 들에게 ReSel-Join 메시지를 전송한다.
- 3) 각 C-RP는 구성원 라우터들로부터 받은 ReSel-Join을 종합하여 수렴된 지연시간 정보를 ReSel-Reply에 담아 RP로 보낸다.
- 4) RP는 C-RP들 중에서 가장 좋은 것을 새로운 RP로 선택하고 새 RP의 주소를 담아 ReSel-Complete 메시지를 멀티캐스트한다
- 5) ReSel-Complete 메시지를 받은 구성원 라우터들은 더 이상의 전송을 중단하고 새로운 RP에 기존의 Join 메시지를 보낸다.
- 6) 각 Join은 321에서처럼 처리되어 새로운 트리를 구성하게 된다.

위 과정이 진행되는 동안 데이터 손실을 막기 위해 기존의 공유 트리는 그대로 유지된다. 따라서 재선택 과정 중에 상한값을 초과하는

경우가 잠시 발생하긴 하지만 멀티캐스트 데이터 전송은 계속된다. 또한 ReSel-Join은 Join, Leave 메시지와 마찬가지로 전송된 경로의 지연시간 값을 알아낼 방법을 제공해야 한다.

3.3 시뮬레이션

노드 수와 평균 차수를 변수로 하여 가상의 망을 무작위로 구성하여 시뮬레이션을 수행하였다. 여러 가지 다양한 조건으로 C-RP들을 선정한 후, 망 내에서 무작위로 구성원들을 하나씩 선정하여 Join시키는 방법을 취하였다. 지연시간의 상한값은 망의 직경으로 지정하여 상한값을 초과하지 않는 트리 구성이 가능하도록 하였다. 최초의 RP는 무작위로 선정하였다.

평균적으로 처음 2~4 구성원에서 상한값을 초과하는 조합이 발생하여 트리 재구성이 이뤄졌고, 트리 재구성이 수행될 때마다 다음 재구성이 발생할 때까지 추가적인 구성원 Join 회수가 늘어났다. 결과적으로 상한값을 만족하지 않는 조합이 초반에 빈번하게 발생함을 알 수 있다.

추가적인 오버헤드, 즉 PIM-SM에 비해 추가된 메시지의 발생회수는 구성원의 수와 C-RP의 수에 비례하여 증가하였다. 그러나 구성원의 수에 대해서는 어느 정도에 이르면 증가가 현저히 둔화되었다.

4. 결론

본 논문에서는 멀티캐스트 리우팅의 주요한 QoS 요구사항인 중단 간 지연시간에 대한 QoS를 지원하는 프로토콜을 PIM-SM을 기반으로 하여 설계하였다.

수신자 간의 수신 속도의 격차는 멀티캐스트 응용에서 중요한 변수가 된다. 시뮬레이션에서 살펴본 것처럼 QoS에 대한 고려 없는 RP 선정은 지연시간 상한값을 만족하지 않는 조합이 쉽게 발견되어 작은 트리 재구성을 요구한다. 그러나 QoS를 고려한 RP 선정이 어려우므로 본 논문에서는 필요시에 RP를 다시 선택하여 공유 트리를 재구성하는 방법을 제시하였다. 결과적으로 본 논문에서 제시하는 프로토콜은 트리 재구성에 따르는 오버헤드를 용인하는 대신, 구성원의 분포와 그룹 규모에 따라 적절한 RP를 찾는 공유 트리를 구성하고, 주어진 지연시간 상한값을 항상 만족시킨다.

참고 문헌

- [1] D. Waitzman, S. Deering, C. Partridge, "Distance Vector Multicast Routing Protocol", RFC 1075, November 1988.
- [2] D. Estrin, D. Farnacci, A. Helmy, V. Jacobson, L. Wei, "Protocol Independent Multicast - Dense Mode (PIM-DM): Protocol Specification, Proposed Experimental RFC, September 1996.
- [3] A. Ballardie, "Core Based Trees (CBT) Multicast Routing Architecture", RFC 2201, 1997
- [4] D. Estrin, D. Farnacci, S. Deering, D. Thaler, A. Helmy, M. Handley, V. Jacobson, C. Liu, P. Sharma, L. Wei, "Protocol Independent Multicast - Sparse Mode (PIM-SM). Protocol Specification", RFC 2362
- [5] M. Faloutsos, A. Banerjee, and R. Pankaj, "QoS-MIC: Quality of Service sensitive Multicast Internet protocol", SIGCOMM '98
- [6] J. Hou, H.-Y. Tyan, B. Wang, and Y.-M. Chen, "QoS Extension to CBT", Internet Draft, February 1999