신뢰성 있는 다대다 멀티캐스트 프로토콜의 에러 복구 성능 비교 분석*

이승익 김성훈 이경민 강경란 이동만
한국정보통신대학원대학교 공학부
{silee, kimsh, kmlee0312, korykang, dlee}@icu.ac.kr

An Analysis of Recovery Mechanism Performance for Many-to-many Reliable Multicast Protocols

Seungik Lee, Sunghoon Kim, Kyungmin Lee, Kyungran Kang, and Dongman Lee
School of Engineering, Information and Communications University

요 약

인터넷 보급의 확산 및 네트웍 성능의 향상으로 멀티캐스트 기술을 응용한 서비스가 일대다의 정보 전달 서비스 위주에서 다중 사용자 네트웍 게임 (MUG) 등과 같이 인터넷을 사용한 다자간 상호작용 서비스로 그 응용 서비스 범위를 넓혀가게 되었다. 본 논문에서는 이러한 다자간 상호작용을 지원하기 위한 다대다 멀티캐스트 세션을 위한 신뢰성 있는 멀티캐스트 프로토콜에서 발생하는 비용을 기술하고, 기존의 일대다 멀티캐스트 세션을 위해 제안된 신뢰성 있는 멀티캐스트 프로토콜들(RMTP, TMTP)과 다대다 멀티캐스트 세션을 위해 제안된 프로토콜들(SRM, Lorax, GAM)에 대한 시뮬레이션을 통하여 에러 복구 기법의 처리 비용을 측정함으로써 다대다 멀티캐스트 세션에 적합한 신뢰성 있는 멀티캐스트 프로토콜을 제안한다.

1. 서론

1992년 인터넷 멀티캐스트가 처음 인터넷에 적용되기 시작한 즈음에는 멀티캐스트 응용 서비스로써 비디오 컨퍼런스 등과 같은 비신뢰적 다대다 (Many-to-many) 멀티캐스트 서비스가 주로 사용되었다. 점차 멀티캐스트에 대한 관심이 높아지면서 그 적용 범위가 확대되고 신뢰성 있는 정보 전달 서비스가 주요 관심 응용 분야로 떠오르게 되어, RMTP[1], TMTP[2], PGM[3], TRAM[4] 등과 같은 일대다 (One-to-many) 신뢰성 있는 멀티캐스트 프로토콜들이 제안되었다

인터넷의 보급이 확산되고 네트웍 성능이 향상됨에 따라 일방적인 정보 전달이 아닌 인터넷 상의 다자간 상호작용에 대한 요구가 늘어가고 있다. 이러한 요구에 따라 다대다 신뢰성 있는 멀티캐스트 프로토콜로서 Lorax[5], SRM[6], GAM[7] 등이 제안되었다.

본 논문에서는 이러한 신뢰성 있는 다대다 멀티캐스트 프로토콜의 에러 복구에 관한 처리 비용을 기술한다. 에러 복구 비용의 관점에서 다대다 멀티캐스트 세션의 신뢰성을 보장하기 위한 최적의 프로토콜을 선정하기 위해, 시뮬레이션을 통해 기존에 제안된 신뢰성 있는 멀티캐스트 프로토콜들의 성능을 분석한다. 일대다 멀티캐스트 프로토콜인 RMTP, 다대다 멀티캐스트 프로토콜인 Lorax, SRM, GAM 등에 대해 성능을 분석하며, 특히 Lorax 의 경우는 트리 구성 알고리즘만 제안하고 있으므로 재전송을 위한 프로토콜로 TMTP 를 적용하여 성능을 분석한다.

2장에서는 신뢰성 있는 다대다 멀티캐스트 프로토콜의 요구사항 분석을 통해 그 필요성을 기술하며 3장에서는 기존의 신뢰성 보장을 위한 멀티캐스트 프로토콜의 기법들에 대해 기술한다. 4장에서는 시뮬레이션에 대한 설명과 그 결과에 대해 분석하고 5장에서 본 논문의 결론을 제시한다.

2. 다대다 멀티캐스트

다대다 멀티캐스트 (Many-to-many Multicast) 는 하나의 멀티캐스트 그룹 주소 안에서 세션 참가자들이 상호 데이터를 서로 교환하는 것을 말한다. 즉, 세션 내의 개별 참가자 모두가 송신자이며 동시에 나머지 참가자들에 대한 수신자가 된다. 응용 분야로는 네트웍을 사용한 다중 사용자 게임, 채팅, 원격 교육, 멀티미디어 회의, 네트웍 상에서의 협동 작업, 중복 데이터베이스 시스템과 같은 분산시스템 등이 있다. 특히 멀티미디어 통신과는 달리 데이터의 무결성을 요구하는 분산시스템은, 일관성 있고 효과적인 상호작용을 위해 데이터의 신뢰성 있는 전달이 필수적이다. 그러나 이러한 다대다 환경의 통신에서 일대다 통신을 위한 신뢰성 보장 기법의 확장 적용은 다음과 같은 문제점을 안고 있다.

n 명의 참가자로 구성된 다대다 멀티캐스트 세션에서 하나의 참가자가 동시에 하나의 송신자와 *(n-1)*개의 수신자 역할을 한다. 한 참가자의 신뢰성 있는 전달을 위한 처리 비용 *COST(M)*을 송신자의 처리 비용 *COST(S)*와 수신자의 처리 비용 *COST(R)*로 구분하면, *COST(M) = COST(S) + (n-1) * COST(R)* 이 된다. [7] *COST(S)* 에 해당하는 비용에는 수신자로부터 전달되는 피드백 처리 비용과 피드백에 따른 재전송 비용, 그리고 재전송을 위한 기전송 패킷의 유지 비용 등이 포함된다. *COST(R)* 에는 수신 실패 시 요구되는 재전송 요청 (retransmission request) 비용, 패킷이 손실된 것을 인지한 후 해당 재전송 패킷을 수신할 때까지 소요되는 복구 지연 시간(recovery latency), 그리고 송신자로부터 재전송된 패킷 (repair packet) 의 처리 비용 등이 포함된다. 이 외에도 송신자의 수가 복수이므로 각 송신자별로 손실 시 복구를 요청할 대상에 대한 정보를 유지해야 하며, 이것 역시 *COST(R)*에 포함된다.

* 본 연구는 과학기술부가 주관하는 국가지정연구실사업 (NRL: M1-0104-00-0130) 의 지원으로 수행되었음.

일반적으로 신뢰성 있는 일대다 멀티캐스트 세션에서는 처리 비용을 분산하기 위해 *COST(S)* 를 최소화하는 프로토콜이 효과적이지만, 다대다 멀티캐스트 세션에서는 참가자의 수가 늘어날수록 *COST(R)*의 효과가 커지므로 *COST(S)* 뿐만 아니라 *COST(R)*도 줄일 수 있는 프로토콜이 요구된다.

3. 신뢰성 있는 멀티캐스트 프로토콜

3.1 Reliable Multicast Transport Protocol (RMTP)

RMTP[1] 는 정적인 ACK 트리 구조를 이용하여 신뢰성 있는 멀티캐스트를 제공한다. 전체 네트웍을 지역(region)으로 나누고, 각 지역마다 Designated Receiver (DR) 을 배치하여 에러 복구를 담당하도록 한다. 그리고, 송신자와 DR, 수신자들 사이의 계층적인 제어 트리를 구성하여 에러 복구 및 요청에 활용한다.

수신자들은 일정한 주기로 ACK 을 전송하는데, 송신자에게 전송하는 것이 아니라 자신이 속한 지역의 DR 에게 유니캐스트로 전송한다. 재전송은 해당 패킷에 대해 지역 내 손실을 겪은 수신자의 수에 따라 유니캐스트 혹은 멀티캐스트를 사용하여 이루어진다. 멀티캐스트를 사용하여 재전송하는 경우에는, 자신이 손실하지 않은 패킷에 대해서 중복 수신하는 수신자들이 발생하게 되고 수신자의 처리 비용이 증가한다. 또 일정한 주기로 ACK 을 전송하기 때문에 패킷 손실을 인지한 시점에서 해당 패킷에 대한 복구가 이루어지기까지의 복구 지연 시간이 길어진다는 단점이 있다.

3.2 Tree-base Multicast Transport Protocol (TMTP)

TMTP[2] 는 확장된 링 검색(ERS; Extended Ring Search) 기법을 사용하여 멀티캐스트 라우팅 트리에 유사한 계층적 제어 트리를 구성한다. 계층적 제어 트리는 세션 참가자들을 노드로 하며 하나의 노드는 특정 개수 *K* 개의 자식 노드의 신뢰성 있는 전달을 책임진다. 자식 노드는 부모 노드에게만 주기적으로 ACK 을 유니캐스트한다. 송신자 또는 자식 노드를 가진 노드에서 자식 노드로부터 제한된 시간 내에 ACK를 받지 못하면 자식 노드만을 포함시키는 범위의 제한적 멀티캐스트를 사용하여 재전송한다. 그리고, 빠른 복구를 위해 NACK을 사용한 복구를 병행하는데, NACK을 억제 (Suppression) 하기 위해 부모 노드를 포함하는 범위의 제한적 멀티캐스트를 사용하여 NACK 을 전송한다.

패킷 손실이 발생해도 NACK 억제 기법에 따라 일정 시간 지연이 생기게 되어 복구 지연 시간이 길어지고, 재전송할 때 멀티캐스트를 이용하므로 동일 패킷에 대해 중복 수신하는 경우가 발생하여 수신자의 처리 비용이 증가한다.

3.3 Lorax

Lorax[5] 는 재전송 기법을 제안하기보다는 제어 트리를 구성하는 기법을 제안하고 있다. 일대다 신뢰성 있는 멀티캐스트 프로토콜의 경우 송신자를 중심으로 제어 트리를 구성하게 되므로 송신자의 수가 많아지면 제어 트리의 수가 증가하고 각 수신자별로 제어 트리를 위해 저장해야 하는 정보의 양이 증가하게 된다. 이러한 비용을 줄이기 위해, Lorax 에서는 단일 제어 트리를 구성하여 모든 송신자에 대해 공유한다. 관리해야 하는 제어 트리의 수가 줄어들지만, 각 송신자와 수신자 간의 실제 패킷 전달을 위한 라우팅 트리와는 거리가 있으므로 재전송 요청 패킷과 재전송 패킷의 전달 시간이 송신자 별 제어 트리에 비해 더 길어질 수 있다.

3.4 Scalable Reliable Multicast (SRM)

SRM[6] 은 제어 트리를 사용하지 않고 수신자 주도로 손실 복구가 이루어진다. 패킷을 손실한 수신자는 NACK 을 세션 참가자들 전체에게 멀티캐스트하고 해당 패킷을 수신했던 참가자가 다시 해당 패킷을 세션 참가자들에게 멀티캐스트한다. NACK 억제를 위해 요청 타이머(request timer)를 사용하여 일정 시간 기다린 후에 NACK 을 전송하고, 중복된 복구를 막기 위해 복구 타이머(repair timer)를 이용하여 일정 시간 기다린 후에 복구를 위한 패킷을 전송한다.

요청 타이머를 송신자와의 추정 RTT(round trip time)을 사용하여 송신자에 가까운 수신자가 NACK 을 전송하도록 하고, 복구 타이머는 재전송을 요청한 참가자와의 추정 거리를 사용한다. 요청 타이머 및 복구 타이머의 길이에 따라 복구 지연 시간이 길어질 수 있으며, 멀티캐스트를 사용하여 재전송하므로 중복된 패킷을 수신하여 수신자의 처리 비용이 증가한다.

3.5 Group Aided Multicast (GAM)

GAM[7] 은 제어 트리를 사용하되 두 단계로 이루어진 제어 트리를 사용한다. 세션 참가자들을 여러 그룹으로 나누고 각 그룹마다 코어(core)를 루트로 하는 제어 트리를 하나 구성한다. 그리고, 코어들 간에는 각 코어를 루트로 하는 제어 트리를 구성한다. 패킷 손실의 패턴에 따라 멀티캐스트 라우팅 트리에서의 부모 자식 관계를 추정하여 라우팅 트리에 가장 근접하도록 제어 트리를 구성한다.

패킷 손실 시에는 제어 트리의 부모 노드에 해당하는 참가자에게 유니캐스트를 사용하여 NACK을 전송하고, 부모 노드도 재전송을 요청한 자식 노드에게 유니캐스트를 사용하여 재전송 패킷을 전송한다. 그리고, 주기적으로 부모 노드에게 ACK 을 전송하여 재전송을 위해 유지하는 패킷 공간을 해제할 수 있도록 한다.

GAM 에서는 코어의 위치가 전체 프로토콜의 성능에 중요한 역할을 하여 각 수신자가 유지해야 하는 제어 트리 정보의 양을 줄였으며 유니캐스트를 사용하여 제어 패킷을 주고 받음으로써 중복된 패킷의 수를 줄여 전체적으로 수신자 처리 비용을 줄였다.

4. 성능 분석

이장에서는 3장에서 설명한 프로토콜들에 대해 다대다 멀티캐스트 세션에서의 에러 복구 성능을 분석한다. 2장에서 기술한 것과 같이 전체 처리 비용 *COST(M)*은 송신자 처리 비용 *COST(S)*과 수신자 처리 비용 *COST(R)*로 이루어진다. 하지만, 본 성능 분석에서는 *COST(S)*와 *COST(R)*을 구분하지 않고 통합하여 전체 처리 비용을 재전송 요청 처리 비용과 송수신한 재전송 비용, 그리고 복구 지연 시간 등으로 나누어 측정하였다.

4.1 성능 분석 환경

Linux 상에서 미국의 UC Berkeley 대학에서 개발한 ns-2 (Network Simulator 2) [9] 를 사용하여 3장에서 기술한 프로토콜들을 시뮬레이션 하였다. RMTP 와 Lorax 는 발표된 논문에 근거하여 시뮬레이션 코드를 생성하였고, SRM과 GAM 은 각각 [6], [7] 에서 사용된 시뮬레이션 코드를 사용하였다. 단, Lorax 는 자체 재전송 기법을 포함하고 있지 않으므로 재전송 기법으로 TMTP의 시뮬레이션 코드를 구현하여 적용하였다.

표 1. 성능 분석 시험 환경 설정

변수명	값
노드 개수	10개, 20개, 30개
평균 회선 지연시간	20 ms
회선 대역폭	10 Mbps
회선 상의 오류 모델	모든 회선에서 1%의 비율로 균등분포에 따른 패킷 손실이 발생하도록 하는 모델
원본 패킷 수	1000 개

실험에 사용하는 네트웍 위상 (topology) 은 GT-ITM[8] 을 사용하여 transit-stub 모델의 위상을 생성하였다. 네트웍 위상 상의 모든 노드에 참가자를 위치시켰으며, 보다 다양한 경우의 성능을 고려하기 위하여 노드의 수를 10개, 20개, 30개로 변화시켰다. 회선 지연 시간을 평균 20ms 가 되게 하되 10ms 이내의 범위에서 감소 혹은 증가시킨 값으로 설정하였다. 추가 트래픽을 발생시키지 않고 각 회선 별로 1%의 패킷 손실이 발생시켜 재전송 과정이 관찰되도록 시험 환경을 구성했다. 시험 시간은 각 송신자가 1000개의 원본 패킷을 전송하고 이에 대한 복구가 완료되는

시점까지로 제한했다. 그리고 각 프로토콜의 제어 트리는 각 논문의 알고리즘에 따라 정적으로 구성했다.

전체 처리 비용 *COST(M)*은 세 가지 영역으로 나누어 볼 수 있는데, 재전송 요청을 위한 "재전송 요청 비용"과 재전송 요청에 따라 송수신된 수리(repair) 메시지를 처리하는 "재전송 처리 비용", 그리고 수신자의 재전송 요청 후 해당 패킷을 재전송 받기까지의 "복구 지연 시간"을 측정하였다.

요청 메시지는 NACK과 ACK을 모두 사용하는 프로토콜(SRM, TMTP를 적용한 Lorax, GAM)에 대해서는 송수신된 NACK의 개수를 측정하였으며, ACK에 재전송 요청 정보를 포함시키는 프로토콜 (RMTP) 에 대해서는 송수신된 ACK 중에서 에러 비트맵을 포함하는 ACK의 개수만을 측정하였다. 수리 메시지로는 송수신된 재전송 패킷의 개수를 측정하였다. 복구 지연 시간으로는, NACK 혹은 ACK 전송 후 재전송 요청에 대응하는 재전송 패킷이 도착하는 데까지 걸리는 시간을 측정하였다.

4.2 성능 분석 결과

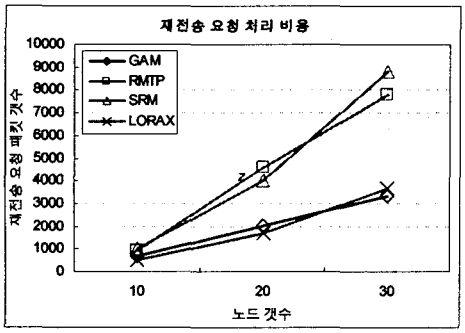

그림 1. 재전송 요청 처리 비용

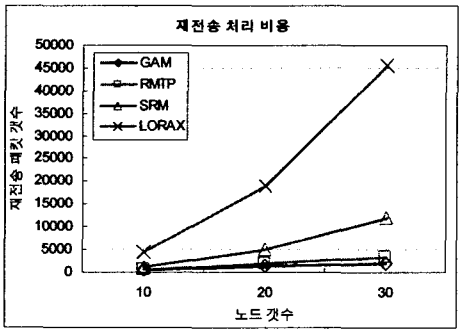

그림 2. 재전송 처리 비용

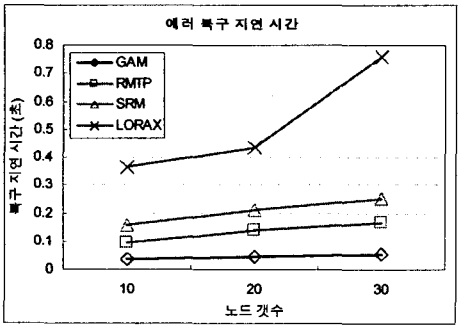

그림 3. 에러 복구 지연 시간

위 그림에서 보다시피 다른 프로토콜에 비해 GAM 이 참가자의 수가 늘어남에도 불구하고 모든 비용면에서 대체로 낮은 비용으로 신뢰성을 보장하고 있다. 재전송 요청 처리 및 재전송 처리 비용면에서는 GAM 이 자신의 부모 노드에 유니캐스트로 재전송 요청 및 처리를 하므로 일정한 시간마다 반복되는 ACK을 이용한 RMTP 나 억제 기법을 위한 NACK과 재전송의 멀티캐스트를 사용하는 SRM 보다 좋은 성능을 나타낸다. 특히 Lorax 는 재한 범위의 멀티캐스트와 NACK 억제 기법을 사용하여 재전송 요청 처리 비용이 적지만 반면에 단일의 제어 트리에 따른 지역 복구 실패에 의해 재전송 처리 비용이 급격하게 늘어나게 된다.

에러 복구 지연 시간 면에서는 마찬 가지로 Lorax 가 단일 제어 트리에 의한 지역 복구 실패로 인해 그 비용이 증가하지만 SRM 및 RMTP 에서는 유니캐스트 및 멀티캐스트를 이용한 지역 복구가 가능하여 그 비용이 적다. 특히 GAM 에서는 NACK 및 재전송의 유니캐스트와 가장 근접한 부모 노드의 지역적 에러 복구를 사용하므로 그 비용면에서 가장 좋은 성능을 나타내었다.

5. 결론

본 논문에서는 신뢰성 있는 다대다 멀티캐스트 프로토콜의 비용을 재전송 요청 비용, 에러 복구 비용, 복구 지연 시간으로 분류하고, 기존의 일대다 멀티캐스트 프로토콜과 다대다 멀티캐스트 프로토콜들에 대해 시뮬레이션을 통해 각 비용을 측정하였다.

RMTP 와 SRM, TMTP와 결합된 Lorax, GAM 에 대해서 비용을 측정한 결과 GAM 이 비교적 적은 수의 요청 메시지, 가장 적은 수의 복구 메시지, 가장 낮은 복구 지연 시간이라는 결과를 보였다. 이 결과에 따라 신뢰성 있는 전달을 위한 최소 비용을 보이는 GAM 을 다대다 멀티캐스트 프로토콜로 사용할 것을 제안한다.

신뢰성 있는 다대다 멀티캐스트 세션을 위해서는 신뢰성 있는 전달 외에 혼잡 제어 및 전송률 조정과 같은 기법이 동반되어야 한다. 그러므로, 이러한 관점에서의 분석 작업이 추가적으로 진행되어야 할 것이다.

참고 문헌

[1] J. Lin and S. Paul, "RMTP: A Reliable Multicast Transport Protocol," IEEE INFOCOM '96, San Francisco, CA, March 1996.

[2] R. Yavatkar, J. Griffioen, and M. Sudan, "Reliable Dissemination Protocol for Interactive Collaborative Applications," ACM Multimedia'95, pp333-344, November 1995.

[3] T. Speakerman, D. Farinacci, S.Lin, A.Tweedly, N.Bhaskar, R.Edmonstone, R.Sumanasekera, and L.Vicisano, "PGM: Relable Transport Protocol Specification," Internet RFC 3208, December 2001.

[4] M. Kadansky, and J. Wesley. "TRAM :A tree-based reliable multicast protocol," Sun Microsystems Technical Report TR-98-66, July 1998.

[5] B. N. Levine, S. Paul, J. J. Garcia-Luna-Aceves, "Organizing multicast receivers deterministically by packet-loss correlation," the sixth ACM international conference on Multimedia, pp201-210, September 1998.

[6] S. Floyd, V. Jacobson, C. Liu, S. McCanne, and L. Zhang, "A Reliable Multicast Framework for Light-weight Sessions and Application Level Framing," IEEE/ACM Transactions on Networking, Volume 5, Number 6, pp784-803, December 1997.

[7] W. Yoon, D. Lee, H.Youn, S. Lee, S. Koh, "A Combined Group/Tree Approach for Many-to-Many Reliable Multicast," IEEE INFOCOM'02, June 2002.

[8] K.I. Calvert, M.B. Doar, and E.W.Qegura, "Modelling Internet Topology," IEEE Communications Magazine, Volume 35, Nuber 6, pp160-163, June 1997.

[9] "The Network Simulator: ns-2," http://www.isi.edu/nsnam/ns/