

신뢰적인 멀티캐스트를 위한 라우터 기반 버퍼 관리

박선옥,⁰ 안상현
서울시립대학교 컴퓨터·통계학과
(sopark95, ahn)@venus.uos.ac.kr

Router based Buffer Management for Reliable Multicast

Soen-ok Park⁰ Sang-hyun Ahn
Dept. of Computer Science & Statistics, Univ. of Seoul

요 약

하나의 송신자가 동일한 데이터를 다수의 수신자들에게 전송하는 인터넷 방송 서비스, 파일 전송 서비스, 뉴스 중계 서비스 등에 있어서 멀티캐스트는 효율적인 수단을 제공한다. 이러한 멀티캐스트의 효율성이 부각되면서 확장성 및 신뢰성을 보장하기 위한 다양한 프로토콜들이 제안되었다. 최근에는 망의 트리 구조를 알고 있는 라우터를 이용하여 지역 그룹내의 응답자(Replier)를 선정하여 손실된 패킷에 대한 재전송을 처리해 주는 방법에 대한 연구가 진행되고 있다. 이 경우, 지역 그룹내의 응답자는 망의 상태에 따라 임의적으로 선택되므로 멀티캐스트 그룹내의 모든 수신자들은 응답자가 될 가능성을 가지고 있다. 따라서 모든 수신자들은 손실 복구를 위한 버퍼를 가지며 수신한 패킷을 저장하게 된다.

본 논문에서는 지역 그룹내의 모든 수신자가 수신했을 패킷을 응답자의 버퍼에서 지우는 시점을 제안한다. 이 기법은 LSM[1] 모델을 기반으로, 지역 그룹을 대표하는 삭제자(Eraser)로부터의 ACK를 기반으로 응답자는 불필요한 패킷을 판단하고 버퍼에서 해당 패킷들을 비우게 된다.

1. 서론

하나의 송신자가 다수의 수신자들에게 데이터 전송을 하는데 있어서 멀티캐스트는 효율적인 수단을 제공한다. 즉, 송신자 입장에서는 같은 데이터를 다수의 수신자들에게 반복해서 전송할 필요가 없으므로 송신자의 작업 부하가 감소하며, 같은 데이터가 네트워크상의 경로를 중복해서 지나는 경우를 최소화함으로써 자원을 보다 효율적으로 이용할 수 있다. 이러한 멀티캐스트의 효율성이 부각되면서 다양한 프로토콜들이 제안되었고 그에 따른 여러 가지 문제점들도 제기되었는데, 특히 확장성 문제와 더불어 신뢰적인 멀티캐스트 전송 문제가 IETF 산하 RMT (Reliable Multicast Transport) WG에서 활발히 진행되고 있다.

최근에는 망의 트리 구조를 알고 있는 라우터를 이용하여 손실 복구를 수행하는 연구가 진행되고 있다. 라우터를 이용한 손실 복구 방법을 제안한 LSM[1]은 수신자가 손실된 패킷을 요청하면 라우터가 손실 복구 요청 패킷을 자신의 하위 링크나 상위 링크로 전달함으로써 손실 복구를 지원한다. 이러한 방법은 그룹 멤버들을 사용하여 지역 그룹을 구성하는 방법보다 지역 그룹 구성이 간단하고, 라우터가 손실 복구를 요청한 링크로 손실 복구 패킷을 전달하기 때문에 노출(Exposure)문제를 감소시킨다.

LSM[1]을 비롯한 지역 그룹내의 응답자(Replier)를 사용하여 손실된 패킷에 대한 재전송을 처리해 주는 방법의 경우, 응답자는 지역 그룹내의 모든 수신자들이 이미 수신한 패킷에 대해서는 더 이상 복구 요청을 받지 않을 것이므로 자신의 버퍼에 더 이상 해당 패킷들을 저장하고 있을 필요가 없다. 그러나, 응답자들은 손실된 패킷에 대한 수신자들의 NACK (Not

Acknowledgement) 정보만을 기반으로 손실된 패킷에 대한 복구를 처리하므로 불필요한 패킷을 판단하기 힘들다. 하지만, 응답자가 모든 수신자로부터의 ACK (Acknowledgement) 패킷을 수신하고 이를 바탕으로 불필요한 패킷을 판단하는 방법은 정확성을 제공하기는 하지만 ACK 폭주(ACK implosion) 문제를 초래하게 된다.

본 논문에서는 지역 그룹내의 모든 수신자가 수신했을 패킷을 버퍼에서 지우는 시점을 제안한다. 멀티캐스트 그룹 멤버를 대표하는 삭제자(Eraser)로부터의 ACK를 기반으로 응답자는 불필요한 패킷을 판단하고 버퍼에서 해당 패킷들을 비우게 된다.

응답자는 지역 그룹 내에서 망의 상태에 따라 임의적으로 선택되므로 모든 수신자들은 응답자가 될 가능성을 가지고 있다. 따라서, 모든 수신자들은 손실 복구를 위한 버퍼를 가지며 수신한 패킷을 저장하게 된다. 응답자의 버퍼에 저장되어 있는 불필요한 패킷들은 자원의 손실을 초래하며 또한, 응답자의 버퍼가 일정 크기로 제한이 된다면 손실 복구를 요청한 수신자가 재전송 패킷을 수신할 때까지 소요되는 손실 복구 시간(loss recovery time)이 증가하게 된다.

논문은 2장에서 관련 연구를, 3장에서 제안한 방법의 대략적인 동작 방법을, 4장에서는 구체적인 삭제자 선정 방법을 설명하고, 5장에서 결론 및 향후 연구 과제를 제시한다.

2. 관련연구

현재까지 제안된 대부분의 신뢰적인 멀티캐스트 프로토콜들은 버퍼 관리에 대해서는 고려하지 않고 있다. 하지만, 불필요한 패

킷을 여전히 버퍼에 저장하고 있는 것은 자원 손실을 초래한다. RMTP[4]는 보조기억장치에 전체 멀티캐스트 데이터를 저장하는 방법을 사용하고 있으나 확장성을 제공해주지 못하며, 메모리에서 패킷을 지우기 위해 타이머를 사용하고 있는 Search Party[7]는 타이머를 어떻게 설정하는지에 대해서는 언급하지 않고 있다.

신뢰적인 멀티캐스트의 버퍼 관리를 위해 새롭게 제시된 방법으로는 Gossip-Style Stability 프로토콜[2]이 있다. 그룹내의 모든 수신자들이 수신한 패킷을 “안정하다” 라고 말하고 있으며 안정한 패킷을 감지하는 방법을 제안하고 있다. 모든 수신자로부터의 상태 정보를 먼저 수신한 수신자가 불필요한 패킷을 결정하고 이를 다른 수신자들에게 다시 알려주는 방법을 사용하고 있다. 하지만 수신자들은 수신한 상태 정보에 자신의 상태까지 반영되도록 상태 정보를 수정해야 하는 오버헤드가 있고, 송신자가 데이터 전송전에 멀티캐스트 그룹 멤버를 미리 알지 못하는 경우에는 확률적인 방법을 사용하므로 신뢰적인 패킷 감지가 어렵다.

3. 제안한 버퍼 관리 기법 개요

논리적인 계층 트리를 구성하는 프로토콜은 손실된 패킷을 복구하기 위해 그룹 멤버들을 사용하여 계층적인 트리를 구성한다. 하지만 논리적인 트리는 물리적인 트리와 다르기 때문에 자식-부모 관계가 바뀔 수 있으며 이 경우 오히려 지역 오류 복구가 올바르게 동작하지 않게 된다는 단점이 있다. 따라서 본 논문의 응답자 구성방법과 동작원리는 이러한 LSM[1]과 같은 물리적인 망을 이용한 손실 복구 방법을 기반으로 한다.

지역 그룹내의 손실 복구를 처리해주는 응답자의 버퍼를 관리하기 위해 수신자들 중에 최대 비용을 가진 삭제자를 선정한다. 지역 그룹 멤버 중에 선정된 삭제자로부터의 ACK를 기반으로 불필요한 패킷을 판단하고 버퍼로부터 해당 패킷을 제거하게 되며, TP마다 하나의 응답자와 하나의 삭제자가 선정된다.

3.1. 망 구성 요소

■ 라우터

지역 그룹 내 모든 수신자들로부터 받은 CANDIDATE[Cost, Null] 메시지를 보고 최소 비용을 가진 수신자를 응답자로 선정하고, Aggregated CANDIDATE[Min_Cost, Max_Cost] 메시지를 상위 라우터에게 전송한다. 이 메시지를 수신한 상위 라우터에서는 하위 링크로부터 받은 정보들을 이용하여 최대 비용을 가진 수신자를 삭제자로 선정하게 된다

■ 응답자

지역 그룹내의 모든 수신자들의 패킷 손실에 대한 복구를 처리하며, 응답자 선정 방법이나 패킷에 대한 손실 복구 방법은 앞 서 말한 LSM[1]을 기반으로 한다.

삭제자로부터의 ACK를 기반으로 불필요한 패킷을 판단하고 지역 그룹내의 모든 수신자들이 버퍼에서 해당 패킷을 비우도록 멀티캐스트로 알려준다.

■ 삭제자

응답자 구성 시에 지역 그룹내의 최대 비용을 가진 수신자를 삭제자로 선정하게 된다. 응답자는 지역 그룹내의 모든 수신자로부터 ACK를 수신하여 불필요한 패킷을 판단하여야 하지만, 이러한 방법은 불필요한 패킷을 판단하는데 걸리는 시간이 길며, ACK 폭주(ACK implosion) 문제를 초래하여 확장성 측면에서 좋지 못한 결과를 가져온다. 따라서 본 논문에서는, 가장

성능이 좋지 않은 수신자가 수신한 패킷은 다른 수신자들도 수신하였을 것이라 가정하고 삭제자를 기반으로 응답자의 버퍼에서 불필요한 패킷을 제거하도록 한다.

라우터는 그룹 멤버가 변경될 때마다 상위 링크로 변경정보를 전송하여 TP를 재구성하고 응답자와 삭제자를 재선정 하도록 한다

3.2. 버퍼에서 지워진 패킷에 대한 손실 복구 요청

응답자는 지역 그룹을 대표하는 삭제자로부터의 ACK를 기반으로 불필요한 패킷을 판단하므로, 패킷 손실에 대한 복구 요청을 직접 처리해 주지 못할 수도 있다. 이러한 경우에는 상위 응답자에도 해당 패킷이 없을 확률이 높으므로 송신자에게 직접 손실 복구를 요청한다.

4. 삭제자 선정 기법 소개

삭제자의 ACK만을 기반으로 불필요한 패킷을 판단하므로 삭제자 선정 방법은 굉장히 중요하다. 삭제자 선정 시 사용되는 수신자 비용은 대역폭, 홉 수, 링크 지연 시간 등 여러 가지 요소를 반영 할 수 있다. 망의 상태에 따라 트리는 재구성되며 응답자와 삭제자도 함께 재선정 되어진다.

4.1. 응답자 / 삭제자 선정 알고리즘

하위 모든 링크로부터 CANDIDATE[Cost, Null] 메시지를 수신하게 되면 라우터는 새로운 CANDIDATE[Min_Cost, Max_Cost] 메시지를 생성하고 상위 링크로 해당 메시지를 전송하며 송신자에게 도달할 때까지 반복된다. CANDIDATE 메시지를 수신한 각각의 라우터들은 하위 링크의 상태정보를 저장하기 위한 테이블을 구성하고 유지한다. 이 테이블을 기반으로 응답자와 삭제자가 선정되고, 선정된 응답자와 삭제자에 대한 정보도 함께 테이블에 저장된다.

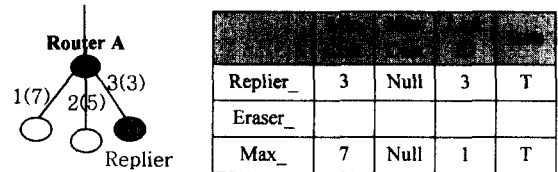


그림 1. 말단 라우터 A의 상태정보 테이블

그림 1은 하위 모든 링크로부터 CANDIDATE[Cost,Null] 메시지를 수신한 말단 라우터 A의 상태 정보 테이블을 보여준다. 각 링크에 있는 숫자는 링크 ID를 의미하며 괄호 안에 있는 숫자는 비용을 의미한다. Replier_ 레코드에는 선정된 응답자에 대한 비용, 링크 정보가 기록된다. Max_ 레코드에는 추후 삭제자 선정 시에 참고할 만한 정보들이 저장되며, 삭제자가 선정되면 Eraser_ 레코드에 정보가 유지된다.

각각의 라우터는 두개 이상의 응답자와 삭제자 링크를 가질 수 없으며 손실 복구 요청 패킷을 수신하게 되면 응답자 링크로 해당 패킷을 전송하게 된다.

CANDIDATE[Min_Cost, Max_Cost] 메시지를 받은 TP에서는 하위의 TP를 위해서 ERASER_SELECT 메시지를 보내게 된다. 그림 1의 말단 라우터 A가 상위 TP로부터 ERASER_SELECT 메시지를 수신하게 되면, 1번 링크가 Eraser 링크로 선정되며 상태정보 테이블의 Eraser_ 레코드에 해당 정

보가 저장된다.

하위 TP는 ERASER_SELECT 메시지를 수신하게 되면 아래와 같은 방법으로 삭제자를 선정한다. 그림 2는 TP에서의 가능한 하위 링크 타입을 보여주고 있으며, 각 링크 타입에 따라서 다른 동작에 의해 삭제자가 선정된다. Er1과 Rep1은 각각 TP1의 삭제자와 응답자를 각각 의미하며 상위 TP로부터 ERASER_SELECT 메시지를 수신함으로써 TP의 삭제자가 선정된다.

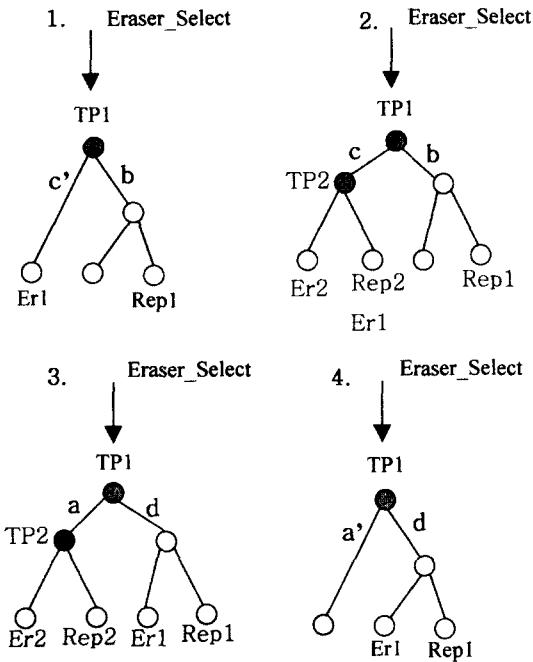


그림 2. TP에서의 가능한 하위 링크 타입

각 라우터에서 수신하게 되는 CANDIDATE 메시지와 저장되어 있는 테이블의 상태정보를 기반으로 링크 타입이 결정되며 링크 타입에 따라 서로 다른 제어 메시지들이 하위 라우터로 전송되어진다. 표 1은 서로 다른 링크 타입에 따라 라우터가 전송하게 되는 메시지들을 보여주고 있다.

표 1. 링크 타입에 따라 전송되어지는 메시지

Cost	Max : O Min : X		Max : X Min : X		Max : X	Max : O
	Max Cost = Null?	O	X	O	X	무관
링크 타입	C'	C	A'	A	B	D
결과 전송 메시지	없음	ERASER_SELECT, LINK_STATE	없음	ERASER_SELECT	없음	없음

4.2. 라우터에서 메시지 수신 시 취하게 되는 동작

삭제자 선정을 위해 사용되는 메시지와 하위 라우터들이 이러한 메시지들을 수신하였을 경우에 취하게 되는 동작들은 다음과 같다.

- CANDIDATE 수신: 라우터의 상태정보 테이블의 Replier 레코드의 Min_Cost보다 메시지의 Min_Cost가 더 크고, 메시지의 Max_Cost가 Null이 아니면 메시지를 수신한 링크로 ERASER_SELECT 메시지를 전송한다.
- ERASER_SELECT 수신: 만약 상태정보 테이블의 Max_레코드의 Sate가 True이면, 해당 링크로 ERASER_SELECT 메시지를 보내고, 그렇지 않으면 응답자 링크로 ERASER_SELECT-REPLIER 메시지를 보냄과 동시에, Max_레코드의 State는 True로 바꾸어 준다
- LINK_STATE 수신: LINK_STATE 메시지를 수신한 라우터는 해당 메시지가 말단 라우터에 도달할 때까지 응답자 링크로 전달하게 되고, 말단 라우터에 LINK_STATE 메시지가 도달하게 되면, 상태정보 테이블의 Max_레코드의 State를 False로 바꾼다. 본 논문에서 사용하는 말단 라우터라는 용어의 의미는 멀티캐스트 트리 구성 시에 중간에 존재하는 라우터가 아니라 수신자와 직접 연결되어 있는 라우터를 의미한다.
- ERASER_SELECT_REPLIER 수신 : 응답자 링크로 ERASER_SELECT 메시지를 전달한다. 이 메시지를 수신한 응답자는 상위 TP의 응답자에 대해 삭제자로서 동작하게 된다.
- ERASER_GIVEUP 수신 : 수신자가 이 메시지를 수신하게 되면 더 이상 삭제자로서 동작하지 않는다.

5. 결론

본 논문에서는 라우터를 기반으로 지역 그룹내의 모든 수신자가 수신했을 패킷을 감지하는 방법을 제안하였다. 이 기법은 LSM[1] 모델을 기반으로, 지역 그룹을 대표하는 삭제자(Eraser)로부터의 ACK를 기반으로 응답자는 불필요한 패킷을 판단하고 자신의 버퍼에서 해당 패킷들을 지울 뿐만 아니라 그룹내의 다른 수신자들에게 알림으로써 불필요한 자원 손실을 막는다.

6. 참고 논문

- [1] C. Papadopoulos and G. Parulkar, "An Error Control Scheme for Large-Scale Multicast Applications", Proc. INFOCOM'98.
- [2] K. Guo, I. Rhee, "Message Stability Detection for Reliable Multicast," INFOCOM 2000.
- [3] Levine, B. and Garcia-Luna-Aceves, J., "A comparison of reliable multicast protocols," ACM Multimedia Sys., August '98
- [4] S. Paul, et al., "RMTP: A Reliable Multicast Transport Protocol for High-Speed Network," Proceedings of the Tenth Annual IEEE Workshop on Computer Communications, September 1995.
- [5] S. Pingali, D. Towsley and J. Kurose, "A Comparison of Sender-initiated and receiver-initiated Reliable Multicast Protocols," SIGMETRICS '94.
- [6] T. Speakman, et al., "Pragmatic General Multicast Transport Protocol Specification", draft-speakman-pgm-spec-00.txt
- [7] A. Costello and S. McCanne, "Search party: Using randomcast for reliable multicast with local recovery", Proceedings of IEEE INFOCOM '99, March 1999.
- [8] S. McCanne, S. Floyd, "NS (Network simulator)," <http://www.nrg.ce.lbl.gov/ns>, 1995.