

멀티캐스트 데이터를 위한 분산적인 오류제어 기법

正會員 宋明烈*, 李俊豪*, 李載用*, 李相培*

A Distributed Error Control Scheme for Multicasting Data

Myong-Lyol Song*, Jun-Ho Lee*, Jai-Yong Lee*, Sang-Bae Lee* Regular Members

이 논문은 부분적으로 1995년도 국방과학연구소 및 학술진흥재단의 연구비 지원에 의한 것임.

要 約

멀티캐스트 데이터 전송에 사용되는 대부분의 오류제어 방법은 멤버들이 탐지한 오류를 복구하는데 송신자가 중심적인 역할을 담당한다. 본 논문에서는, 분산적인 고속 광역망과 같은 환경에서 그룹통신을 위한 멀티캐스트 에러제어 기법 DECMC(Distributed Error Control scheme for MultiCasting data)를 제시하였다.

이 기법은, 망 전체에 걸쳐 분산되어 있는 LGM(Local Group Manager)이 이웃멤버들에게 발생된 오류를 복구할 수 있기 때문에, LGM이 송신자의 부하를 부분적으로 담당함으로써 송신자와 멤버들 사이에 교환되는 제어메시지 수를 줄인다. 패킷의 전파범위를 제어하는 time-to-live를 조절함으로써, 그룹통신을 위한 멀티캐스트 트리가 몇개의 작은 subnetwork으로 나뉘어 지고, 각 subnetwork 내의 LGM이 이웃멤버들의 상태를 감시하고 그들이 요청한 데이터를 재전송한다.

제안된 오류제어 방법으로써, 멀티캐스트 트리에 발생하는 제어메시지 수의 감소, throughput 증가, 국부적인 문제로 인해 전체 멤버에게 미치는 영향 차단, 오류의 신속한 복구 등과 같은 효과를 얻는다. 컴퓨터 시뮬레이션을 통해 제안된 방법의 효용성을 확인하였다.

ABSTRACT

Most error control schemes for multicasting data are sender-oriented so that sender takes a major role in recovering the errors occurred to some members. In this paper, a DECMC(Distributed Error Control scheme for MultiCasting data) for group communication, especially, in a distributed high speed wide area network, is proposed.

In the scheme, the LGM's(Local Group Manager) distributed throughout a network can recover the error occurred to their neighbor members so that the load of the sender partly moves to the LGM's in order to reduce the number of control messages

* 연세대학교 전자공학과
論文番號 : 95298-0831
接受日字 : 1995年 8月 31日

between the sender and the other members. By adjusting the time-to-live field controlling the scope of the propagation of a packet, a multicast tree for group communication can be divided into several small subnetworks, and then the LGM in each subnetwork monitors the states of its members, retransmits some required data to them.

The performance of group communication is improved by the following effects : reduction of control messages in multicast tree, increase of throughput, prevention of local problems from affecting entire members in the group, and fast error recovery. Computer simulation results show the efficiency of the proposed scheme.

I. 서 론

최근 고속전송과 대용량 저장기술의 발전으로 인해 CSCW(Computer Supported Cooperative Work), Desktop Conference, Multimedia Mail 등과 같은 분산 멀티미디어 응용분야의 시스템들이 개발되고 있다. 이러한 분야의 응용 시스템들은 그룹통신 기능을 필요로 하며 효율적인 데이터의 전송을 위해 멀티캐스트 기능의 지원을 받아야 한다.⁽¹⁾

그룹통신에 필요한 멀티캐스트 기능을 지원하기 위해서는 두가지 측면을 고려해야 한다. 망계층의 멀티캐스트 기능과 수송계층의 멀티캐스트 기능을 고려해야 하는데, 망계층에서는 멀티캐스트를 위한 최적의 경로를 설정하는 하는 방법에 관한 연구⁽²⁾가 주종을 이루고 있고 수송계층에서는 멀티캐스트 데이터 전송시 발생하는 메시지 수, 순서보장(ordering), 신뢰성(reliability) 등과 같은 문제들을 해결하기 위한 연구가 진행되어 왔다^(3, 4, 5, 6, 7, 8, 10, 11, 12).

특히 오류제어를 위해서는, 기존의 수송계층 멀티캐스트 프로토콜들은, 송신자가 그룹내에서 중심적인 역할을 담당하면서, 수신자의 부담을 줄이고 망에서 발생하는 전체 메시지 수를 줄이도록 설계되었다. 예를 들어, MTP, XTP, VMTP 등과 같은 프로토콜에서 사용되는 멀티캐스트 오류제어기법이 이러한 부류에 속한다. 이러한 프로토콜의 대부분은 재전송되는 메시지 수와 제어를 위한 메시지 수를 줄이는데 초점을 맞추고 있다. MTP는 신뢰성이 높은 망에서 그룹이 master, consumer, producer로 구성된다고 가정하고 NAK방식으로 오류제어를 수행한다. Consumer가 오류발견시 unicast 주소로 NAK를 전송하면 producer가 해당 데이터를 멀티캐스트로 재전송한다⁽⁴⁾. 그러나 이 프로토콜은 전송한 데이터를 일정기간 후에는 버리므로 신뢰성

은 보장하지 못한다. XTP에서는, 오류제어를 위해 bucket 알고리즘을 사용하는데, damping과 slotting 기능을 함께 사용하여 발생하는 제어용 패킷 수를 감소시켜 송신자의 부담을 경감시킨다. 또한, list-based 멀티캐스트 기법은, 송신자가 수신자들의 목록을 관리하며 수신자들의 멀티캐스트 데이터 수신 상태를 확인하는 방법인데, k개의 수신자까지 신뢰성을 보장할 수도 있다⁽⁵⁾. VMTP에서는 송신자에 가장 먼저 도착되는 acknowledgment만으로도 멀티캐스트된 데이터가 수신자들에게 성공적으로 전송되었다고 판단함으로써 재전송되는 메시지 수를 상당히 줄인다⁽⁶⁾. 그러나 이 프로토콜은 데이터 전송에 대한 신뢰성은 보장할 수 없다.

앞에서 언급한 멀티캐스트 프로토콜들이 효율적으로 동작할지라도, 분산 고속 광역망에서 많은 수의 멤버를 포함하는 환경에서 동작할 경우, 버퍼의 크기, 망에 대한 부하 및 신뢰도(reliability) 등과 같은 문제들이 야기될 수 있다. 이러한 환경에서는, round-trip time 이 크기 때문에 송신자가 재전송을 대비하여 규모가 큰 버퍼를 관리해야 하고, 많은 멤버들이 송신자에서 처리되어야 할 제어용 패킷을 발생시키므로, 멀티캐스트된 데이터의 신뢰도가 떨어지게 된다. 이 논문에서는, 고속 광역망에 많은 수의 멤버들이 분산적으로 분포되어 있는, 그룹통신에서 발생하는 멀티캐스트된 데이터에 대해 분산적으로 오류제어를 수행하기 위한 방법 DECMC(Distributed Error Control for MultiCast)를 제안한다. 이 방법을 적용함으로써, 광역망에서 재전송되는 메시지 수가 상당히 줄어들고 송신자의 버퍼 크기를 적게 준비해도 되는 효과를 얻는다. LGM(Local Group Manager)이 멀티캐스트된 데이터를 관리 제어하기 때문에, 전체적으로 볼 때, 데이터 전송에 대한 신뢰도가 향상된다. II 장에서는 그룹통신의 한 모델을 정하여 멀티캐스트 데이터에 대해 분산적으로

오류제어하는 방법을 기술하고, III 장에서는 제안된 오류 제어 방법에 대한 해석을 다루고, IV 장에서는 시뮬레이션을 통해 제안된 방법에 대한 특성을 알아보고, V 장에서 결론 및 향후 연구방향을 제시한다.

II. 제안된 DECMC 기법

그림 1에 제안된 DECMC 방법이 적용될 수 있는 그룹통신 모델을 나타냈다. 이 모델에서, 각 subnetwork은 라우터(router)를 통해서 서로 연결되어 있고 멤버들은 자신이 속한 subnetwork에 연결되어 있다. 각 subnetwork안에서는 첫번째로 그룹에 가입(join)하는 멤버(member)가 LGM의 역할을 동시에 수행한다. 송신자(sender)로부터 멀티캐스트된 메시지는 라우터의 망계층 멀티캐스트 기능의 도움을 받아 수신자인 멤버들에게 전달된다.

제안된 DECMC scheme에는 두 종류의 독립적인 상호작용이 이루어진다. 그 중 하나는 송신자와 LGM 사이의 상호작용이고, 다른 하나는 임의의 subnetwork i 에서 LGM i 와 그 LGM의 이웃 멤버(멤버 j)들 사이의 상호작용이다.

정상적으로 동작할 때, 송신자는 망의 멀티캐스트 멤버들에게 데이터를 멀티캐스트로 전송하고 주기적으로 제어메세지를 LGM에게 전송한다. LGM은 수신된 데이터를 저장하면서 제어메세지를 수신할 경우 응답메세지를 송신자에게 전송한다. 또한 LGM은, 자기의 이웃 멤버들의 상태를 판단하기 위해, 주기적으로 제어메세지를 이웃 멤버들에게 전송한다. 한 subnetwork 내의 멤버들은 자기와 이웃하는 LGM으로부터 제어메세지를 수신하면 응답메세지를 그 LGM에게 전송한다. LGM은 이웃 멤버들 모두로부터 응답메세지를 수신하면, 이웃 멤버들에 의해 수신화인된 데이터를 LGM 내에 유지하고 있는 queue로부터 제거시킨다. LGM이 이웃 멤버들 가운데 오류가 발생했거나 분실된 데이터가 존재한다고 판단할 경우에는 Go-back-N 방법에 따라서 해당된 데이터를 재전송한다. 제어메세지와 응답메세지가 망 전체에 걸쳐서 범람하는 것을 방지하기 위해, 메시지가 전파되는 범위를 제어하는 field가 메시지 내부에 정의된다.

1. 메세지 형식

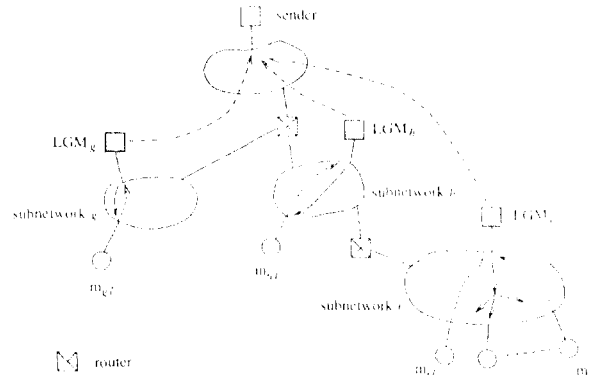


그림 1. DECMC 방법이 적용된 그룹통신 모델
Fig. 1. A group communication model for the DECMC scheme.

이 절에서는, 어떤 망에 접속되어 있는 end host에서 수행되는 수송계층 프로토콜에 DECMC방법이 적용될 경우 그 프로토콜에 의해 발생하는 메세지들을 정의하고, 그 메세지들이 어떤 그룹에 속해 있는 멤버들에게 도착할 경우에 관찰되는 메세지의 특성에 대해 분석한다.

데이터 전송 및 멤버들의 상태정보 교환을 위해, 수송계층 프로토콜에 *Data*, *Control*, *ACK* 메세지를 정의한다.

- *Data*(*DstAddr*, *SrcAddr*, *ttl*, *age*, *seq*, ...)
- *Control*(*DstAddr*, *SrcAddr*, *ttl*, *age*, ...)
- *ACK*(*DstAddr*, *SrcAddr*, *ttl*, *age*, *rseq*, ...)

여기서 각 변수들은 다음과 같은 의미를 갖는다. *DstAddr* : 패킷의 목적지 주소, *SrcAddr* : 패킷의 발신지 주소, *GrpAddr* : 그룹주소, *ttl*(time-to-live) : 패킷의 전파범위 제어에 사용되는데, 라우터를 통과할 때마다 1씩 감소된다. *age* : 어떤 패킷이 통과해야 할 라우터 수를 나타내는데, 라우터는 이 필드를 처리하지 않는다. *seq* : *Data* 메세지에 전송되는 데이터 중 첫번째 byte를 나타내는 순서번호(sequence number). *rseq* : 멤버가 수신되기를 기다리고 있는 데이터의 순서번호.

DstAddr, *SrcAddr*, *ttl*은 망계층 패킷에 해당되는 field에 대응되고 망계층 프로토콜에 의해 처리되는 반면, *age*, *seq*, *rseq* field는 end host의 수송계층 프

로토클에 의해 처리된다. 이럴 경우, 그룹주소와 time-to-live field를 처리하는 기존의 망계층 프로토콜을 변경하지 않고 수송계층 프로토콜에 DECMC 방법을 포함하는 것이 가능하다.

표 1. DECMC 방법에 의해 발생하는 메시지들의 특성
Table 1. Characteristics of the messages generated by the DECMC scheme

	Data	Control	Acknowledgment
Sender	$t_{tl}=ttl_{max}$ $age=agemax$	$t_{tl}=ttl_{max}$ $age=agemax$	
LGM	$t_{tl}=1$ $age=0$	$t_{tl}=1$ $age=0$	$t_{tl}=ttl_{max}$ $age=agemax$
Member			$t_{tl}=1, ttl_{max}$ $age=0, agemax$

표 1에 DECMC를 사용하는 그룹에서 송신자, LGM, 멤버들에 의해서 발생하는 메시지들의 특성을 나타냈다. 패킷의 t_{tl} field는, 패킷이 라우터를 통과할 때마다 1씩 감소되므로, 멀티캐스트되어 임의의 멤버에 의해 관찰될 때는 1과 ttl_{max} 사이의 정수값을 갖는다. 그러나 age field는 라우터에 의해 처리되지 않고 end host의 수송계층 프로토콜에 의해서 처리되기 때문에 발생당시의 값을 그대로 유지하고 있다.

2. 송신자

송신자는 멤버들의 상태를 감시하기 위해서 매 global switching time 마다 주기적으로 제어메세지 Control를 전송한다. 송신자는, 이 제어메세지에 대한 응답으로서 각 LGM이 발생시킨, 수신된 응답메세지 ACK에 포함되어 있는 $rseq$ 값을 이용하여 차기에 재전송해야 할 데이터를 결정한다. 이러한 것들이 오류제어를 위한 DECMC 방법의 주요 기능이다.

Global switching time은 round-trip time (RTT)에 의해 결정되는데, RTT는 송신자가 어떤 데이터를 수신자에게 전송한 후 수신자로 부터 응답이 수신될 때까지 걸리는 시간을 말한다⁽⁹⁾. 본 논문에서는 XTP에서 사용된 방법으로 global switching time을 계산하는데, XTP에서는 송신자와 수신자 사이에 측정된 RTT 값들 중 가장 큰 값을 토대로 global

switching time을 정한다⁽³⁾.

송신자가 재전송해야 할 데이터의 순서번호는 식 (1)에 의해 결정된다.

$$\widehat{rseq}_T = \min\{rseq_{L_1}, rseq_{L_2}, \dots, rseq_{L_g}\}. \quad (1)$$

여기서, $rseq_{L_i}$ 는 LGM i 가 송신자에게 재전송을 요구한 데이터의 순서번호이다. 만일 어떤 LGM이 비정상적인 상태에 있을 경우에, 정상적인 상태로 동작할 때보다, 송신자가 더 많은 수의 ACK를 받을 수도 있다. 그 이유는 그러한 LGM과 동일한 subnetwork에 속해 있는 멤버들이 ACK를 LGM에게 전송하지 않고 송신자에게 전송하기 때문이다. 이 경우에 \widehat{rseq}_T 는 식 (2)와 같이 결정된다.

$$\widehat{rseq}_T = \min\{rseq_{L_1}, rseq_{L_2}, \dots, rseq_{L_g}, rseq_{m_1}, rseq_{m_2}, \dots, rseq_{m_n}\}. \quad (2)$$

여기서, $rseq_{m_j}$ 는 비정상적인 LGM m 과 동일한 subnetwork에 속해 있는 멤버 j 가 송신자에게 재전송을 요구한 데이터의 sequence number이다.

3. 라우터

라우터는 자기 자신을 경유하는 멀티캐스트 트리를 형성하는데 필요한 정보를 포함하는 표를 유지하고, 이 표를 참조하여 수신된 패킷을 목적지에 따라 적절한 경로를 선택하여 전송하는 기능을 수행한다. 이 때, $t_{tl} \geq 2$ 인 패킷은 t_{tl} 값을 1씩 감소시킨 후 전송하고, $t_{tl} = 1$ 인 패킷은 버린다.

DECMC 방법에서는 라우터에서 t_{tl} 을 제어하는 기능을 이용하여 패킷이 전파되는 범위를 조절한다. 멤버가 $t_{tl} = 1$ 인 패킷을 전송할 경우, 이 패킷은 동일한 subnetwork에 속해 있는 LGM이나 이웃하는 멤버들에게는 전달되지만, 인접한 라우터를 통과하여 다른 subnetwork의 멤버들에게는 전달되지 않는다. 따라서, $t_{tl} = 1$ 인 패킷에 대해서는 멀티캐스트 트리를 멤버들 간의 거리가 1 hop인 몇개의 작은 subnetwork으로 나눌 수 있게 된다.

4. 멤버

멤버들은 송신자나 자신이 속한 subnetwork의 LGM이 전송한 데이터를 수신하고, Control 메세지를 수신할 경우 자신의 수신상태를 ACK 메세지를 통해 알

린다. 멤버에 도착된 패킷은 내용, 중복, 순서에 대해서 오류가 발생했는지 조사된다. 이때, 오류가 발생된 패킷은 버리고 오류가 없는 패킷은 상위계층으로 건네주면서 이 패킷으로 부터 seq 를 추출하여 멤버의 상태변수를 갱신하는데 사용한다.

에러제어를 위해 DECMC 방법이 적용될 경우, 멤버들의 동작은 자신이 속한 subnetwork의 LGM의 상태에 따라 달라질 수 있다. LGM이 정상적으로 동작하고 있다고 판단되면, 멤버들은 다음과 같은 ACK 메시지를 전송한다.

$ACK(GrpAddr, MemAddr, ttl=1, age=0, rseq, \dots)$.

$ttl = 1$ 이기 때문에 이 메시지는 subnetwork 내의 멤버들에게만 멀티캐스트되고, $rseq$ 는 LGM이 재전송해야 할 데이터를 결정할 때 사용된다. LGM이 비정상적인 상태에 있다고 판단될 경우, 멤버들은 송신자에게 다음과 같이 유니캐스트로 응답한다. 여기서 $SndAddr, LGMAddr, MemAddr$ 은 각각 송신자, LGM, 멤버의 유니캐스트 주소를 의미한다.

$ACK(SndAddr, MemAddr, ttl=ttlmax, age=agemax, rseq, \dots)$.

멤버들은 수신된 Control 메시지에 포함되어 있는 age 와 멤버 자신이 관리하는 timer를 이용하여 인접한 LGM의 동작 상태를 판단하여 LGM과 송신자 중 누구에게 응답할 것인가를 결정한다. LGM이 전송한 Control 메시지가 수신될 때마다 timer를 주어진 초기값으로 reset시킨다. 만일 timer-out이 발생하면 LGM의 Control 메시지를 수신하지 못한 결과 이므로 LGM이 비정상적인 상태에 있다고 판단한다. LGM이 전송한 제어메시지와 송신자가 전송한 제어메시지는 $age = 0$ 와 $age = agemax$ 로 각각 구분된다.

5. 지역망 그룹 관리자

DECMC방법에서 가장 중요한 역할을 수행하는 지역망 그룹 관리자 LGM은 송신자로 부터 수신된 데이터를 저장하고, 송신자의 Control 메시지에 대해 ACK 메시지로써 자기의 수신상태를 알린다. 또한, 같은 subnetwork에 속해 있는 멤버들의 상태를 감시하기 위해 Control 메시지를 멤버들에게 전송하고 멤버들이 요구한 데이터를 재전송하는 기능을 수행한다.

어떤 subnetwork 내의 멤버들을 위한 오류복구는

다음과 같은 세 단계를 반복함으로써 수행된다.

1 단계 : 제어메시지 전송

LGM은 매 local switching time 마다 자기가 속한 subnetwork의 멤버들의 상태를 감시하기 위해 제어메시지를 멀티캐스트로 전송한다.

$Control(GrpAddr, LGMAddr, ttl=1, age=0, \dots)$

2 단계 : 멤버들의 상태정보 수집

1단계에서 전송된 제어메시지에 대해 멤버들이 응답한 메시지를 수집한다.

$ACK(GrpAddr, MemAddr, ttl=1, age=0, rseq_i, \dots), i=1, \dots, m$.

3 단계 : 요구된 데이터 재전송

$\widehat{rseq} = \min(rseq_1, rseq_2, \dots, rseq_m)$;
if ($\widehat{rseq} < rseq_l$) then
 send Data($GrpAddr, LGMAddr, ttl=1, age=0, \widehat{rseq}, \dots$);
else

재전송해야 할 데이터의 순서번호 \widehat{rseq} 는 수집된 $rseq$ 값들 즉, $rseq_1, rseq_2, \dots, rseq_m$ 중 최소값으로 결정된다. 만일 LGM이 재전송할 데이터를 갖고 있는 경우에는 멀티캐스트로 그 데이터를 재전송해 주고 그렇지 않은 경우에는 전송하지 않는다.

III. 해 석

1. Sequence Number에 대한 해석

이 절에서는 DECMC 방법이 적용될 때 subnetwork 내의 멤버들의 수신상태를 나타내는 $rseq$ 의 특성에 대해서 분석한다.

Subnetwork i 에 LGM을 포함하여 n 개의 멤버가 있다고 하자. 순서번호는 유한개의 비트로 표현되기 때문에 유한한 값을 가지면서 순환하지만, 해석의 편의를 위해 무한한 값을 갖으면서 단조증가한다고 가정하자. 제어메시지가 LGM에 도착될 때 subnetwork i 의 모든 멤버들의 상태를 관찰한다. LGM이 재전송해야 할 데이터의 순서번호는 식 (3)과 같이 결정된다.

$$\widehat{rseq}_{L_i} = \min\{rseq_{i2}, rseq_{i3}, \dots, rseq_{in}\}. \quad (3)$$

여기서, \widehat{rseq}_{L_i} 와 $rseq_{ij}$ 는 각각 subnetwork i 에서 LGM이 재전송해야 할 데이터와 멤버 j 가 재전송 요구한 데이터의 순서번호를 나타낸다.

송신자가 전송한 패킷이 멤버에 도착되고 멤버가 이 패킷에 대해 오류를 검사할 때, 오류를 발견할 확률이 모든 멤버에게 동일하다고 하자. Subnetwork i 에서 LGM의 $rseq_{L_i}$ 와 \widehat{rseq}_{L_i} 를 비교하자.

$$\begin{aligned} P(rseq_{L_i} \geq \widehat{rseq}_{L_i}) &= 1 - \prod_{j=2}^n P(rseq_{L_i} < rseq_{ij}) \\ &= 1 - \left(\frac{1}{2}\right)^{n-1} \\ P(rseq_{L_i} < \widehat{rseq}_{L_i}) &= \left(\frac{1}{2}\right)^{n-1}. \end{aligned}$$

그러므로, $n \geq 2$ 에 대해

$$P(rseq_{L_i} \geq \widehat{rseq}_{L_i}) \geq P(rseq_{L_i} < \widehat{rseq}_{L_i}). \quad (4)$$

이것은 subnetwork 내에 2개 이상의 멤버가 존재할 경우, 수신상태가 가장 좋지 않은 멤버 보다 LGM의 수신상태가 확률적으로 볼 때 더 좋다는 것을 의미한다.

이제 송신자에서 순서번호의 특성을 알아보자. DECMC 방법이 적용될 때 송신자가 재전송해야 할 데이터의 순서번호를 \widehat{rseq}_T 라 하자. XTP와 같이 송신자를 중심으로 오류제어를 수행하는 프로토콜에서 송신자가 재전송해야 할 데이터의 순서번호를 \widehat{rseq}_T 라 하자. 식 (1)로 부터,

$$\widehat{rseq}_T = \min\{rseq_{L_1}, rseq_{L_2}, \dots, rseq_{L_g}\}.$$

송신자 중심의 오류제어 방법에서,

$$\widetilde{rseq}_T = \min\{rseq_1, rseq_2, \dots, rseq_G\}.$$

여기서 $rseq_G$ 는 그룹 내의 임의의 멤버가 재전송을 요구하는 데이터의 순서번호이다. Subnetwork내에서 $rseq$ 특성을 알아 볼 때 적용했던 것과 비슷한 방법으로 다음과 같은 관계를 구할 수 있다.

$$P(\widehat{rseq}_T \geq \widetilde{rseq}_T) \geq P(\widehat{rseq}_T < \widetilde{rseq}_T). \quad (5)$$

여기서 등호는 각 subnetwork에 단 하나의 멤버만 존재하는 경우에 성립한다. 각 subnetwork에 멤버수가 적어도 2개 이상인 경우, DECMC 방법에서 송신자는 송신자 중심의 다른 오류제어 방법 보다 더 큰 순서

번호의 데이터를 멤버들로 부터 요구받는다. 그러므로 송신자가 재전송해야 할 데이터의 양이 줄어들게 된다.

2. 오류복구시간에 대한 해석

이 절에서는 DECMC방법에 대한 지연시간의 특성을 분석한다. 특히, 임의의 subnetwork 내의 멤버들이나 LGM에 의해 탐지된 오류가 복구되는데 걸리는 시간에 초점을 맞추어 분석하고 XTP의 지연특성과 비교한다.

DECMC방법에서는 Go-back-N 오류제어 기법을 사용한다. 송신자는 global switching time T_{gsW} 마다 제어 메시지를 전송하는데, 멤버들이 요구한 데이터는 이 제어 메시지를 전송하기 직전에 재전송한다. 또한 LGM도 local switching time T_{lsW} 마다 송신자와 같은 방식으로 제어 메시지와 데이터를 재전송한다. 그림 2는 어떤 그룹에 참여하고 있는 멤버들, LGM 및 송신자 사이에서 오류복구를 위한 상호작용을 나타냈다.

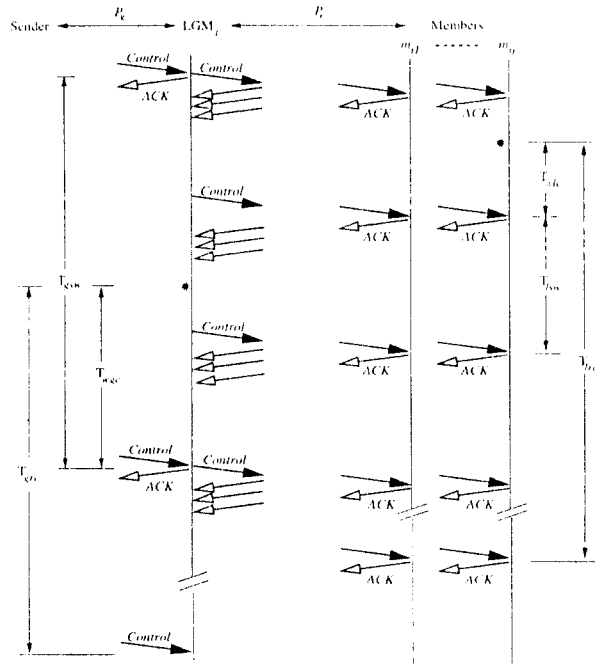


그림 2. DECMC 방법에서 멤버들 간의 상호작용.
Fig. 2. Interactions between the members under the DECMC scheme.

송신자와 임의의 subnetwork에 연결된 라우터 사이에서 P_g , 라우터와 subnetwork내의 멤버들 사이에서 P_e 의 확률로 각각 오류가 발생한다고 가정하자. 임의의 시각에 LGM에 의해 오류가 탐지될 경우, LGM이 그 오류를 송신자로 부터 복구받는데 걸리는 시간은

$$E(T_{err}|T_{err}) = \sum_{i=1}^{\infty} [T_{wgc} + (i-1)T_{cst}] P_i \quad (6)$$

여기서 P_i 는 송신자가 i 번 재전송한 후에 에러가 복구된 확률이다. 즉,

$$P_i = [1 - (1 - P_g)(1 - P_e)]^{i-1} (1 - P_g)(1 - P_e)$$

이고, T_{wgc} 는 LGM에서 오류를 탐지한 후 송신자로 부터 첫번째 제어 메시지가 LGM에 도착할 때까지 대기하는 시간이다. 그러므로

$$E(T_{err}|T_{err}) = T_{wgc} + \frac{1 - (1 - P_g)(1 - P_e)}{(1 - P_g)(1 - P_e)} T_{cst}$$

LGM이 오류를 탐지하는 시각이 $(0, T_{gsw})$ 에서 균일하게 분포된다고 가정하면,

$$E(T_{err}) = \left[\frac{1}{2} + \frac{1 - (1 - P_g)(1 - P_e)}{(1 - P_g)(1 - P_e)} \right] T_{cst} \quad (7)$$

이제 임의의 멤버가 오류를 탐지했을 경우, 멤버에서 오류복구에 소요되는 평균시간 $E(T_{irc})$ 을 구하자. 송신자가 데이터를 i 번 전송한 후 LGM과 멤버가 올바르게 수신할 경우를 각각 A_i 와 B_i 로 표현하자. 송신자가 데이터를 i 번째 전송시 LGM에는 올바르게 수신되지만 멤버에게는 오류가 감지되었다고 할 때, 멤버가 그 데이터를 올바르게 수신할 때까지 LGM이 멤버에게 재전송한 횟수를 N_m 으로 표현하자. 그러면,

$$E(T_{irc}|T_{irc}, T_{err}) = \sum_{j=1}^{\infty} [T_{wgc} + (j-1)T_{cst}] P(N_m = j)P(A_j) + \sum_{i=2}^{\infty} [T_{wgc} + (i-2)T_{cst}] P(B_i|A_i)P(A_i) + \sum_{i=2}^{\infty} \sum_{j=1}^{\infty} [T_{gsw} + (i-2)T_{cst} + jT_{cst}] P(N_m = j)P(A_i, B_i|A_i)P(A_i) \quad (8)$$

여기서, T_{wgc} 는 송신자가 전송한 데이터에 대해, 멤버가 오류를 탐지한 후 LGM이 전송한 첫번째 제어메시지가 멤버에 수신될 때까지 멤버가 대기하는 시간을 의

미한다. 그러므로 멤버에 의해 탐지된 오류가 복구되는 데 소요되는 평균시간 $E(T_{irc})$ 를 구할 수 있다.

$$E(T_{irc}) = (1 - P_g)(1 - P_e) \left(\frac{1}{2} + \frac{P_e}{1 - P_e} \right) \frac{1}{f} T_{cst} + [P_g + (1 - P_g)P_e] \left(\frac{1}{2} + \frac{1}{1 - \alpha} \right) T_{cst} + [1 + (1 - P_g)P_e] (1 - P_g)(1 - P_e) P_e \frac{1}{1 - \alpha} \left(\frac{1}{2} + \frac{1}{1 - \alpha} + \frac{1}{f} \frac{1}{1 - P_e} \right) T_{cst} \quad (9)$$

여기서

$$\alpha = (1 - P_g)P_e^2 + P_g$$

이고

$$f = \frac{T_{cst}}{T_{gsw}} \quad (10)$$

그림3에서 곡면들은 에러확률의 변화에 따라 오류복구에 소요되는 시간을 의미하는데, global switching time으로 정규화된 것이다. 0.5 수준의 평면을 갖는 곡면은 LGM의 오류복구시간 $E(T_{irc})/T_{gsw}$ 을 나타내는데, 이 값은 송신자와 LGM 사이의 RTT가 감소하지 않으면 더이상 줄어들지 않는다. 밑에서 부터 두번째, 첫번째 곡면은 각각 $f=2$ 와 $f=4$ 에 대한 멤버의 오류복구시간 $E(T_{irc})/T_{gsw}$ 을 나타낸다. $E(T_{irc})/T_{gsw}$ 는 f 와 반비례 관계에 있고, 또한 $f > 1$ 에 대해

$$E(T_{irc}) < E(T_{gsw}) \quad (11)$$

인 관계가 있음을 알 수 있다.

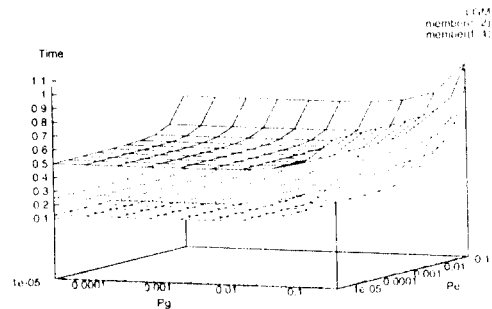


그림 3. 정규화된 오류복구 시간.
Fig. 3. Normalized error recovery time.

XTP에서 사용되는 오류제어 알고리즘의 성능을 살펴 보면, XTP의 모든 멤버들은, 제안된 DECMC 방법의 LGM처럼, 송신자와 직접 상호작용하므로 오류복구에 소요되는 시간은 식(7)과 같이 주어질 수 있다. 반면에 DECMC에서 멤버들의 오류복구시간은 식(9)와 같이 표현된다. 그러므로 송신자와 멤버들 사이에서는 큰 값의 RTT를 갖는 반면 멤버들 사이에서는 작은 값의 RTT를 갖는 그룹에서는, DECMC 방법을 사용하는 멤버들이 XTP의 멤버들 보다 더 작은 값의 오류복구시간을 갖는다.

IV. 시뮬레이션 결과

제안된 DECMC방법의 동작을 관찰하고 성능을 측정하기 위한 시뮬레이션 모델을 그림 4에 나타냈다. 멤버들에게 탐지된 오류를 복구하는데 LGM이 얼마나 효율적으로 동작하는가를 알아보기 위해, LGM과 멤버들에 의해 유지되는 상태변수 *rseq*의 변화를 관찰한다. 제안된 방법의 성능을 평가하기 위해, LGM이나 멤버들에 의해 탐지된 오류가 복구에 소요되는 시간을 측정한다. 시뮬레이션 모델에서 송신자와 subnetwork 내에 5개의 멤버가 DECMC 방법을 사용하고 있고, global switching time과 local switching time은 각각 $T_{gsw} = 20.0$ 과 $T_{lsw} = 4.0$ 이 사용된다. 송신자와 subnetwork에 인접한 라우터 사이에 오류가 발생할 확률은 $P_g = 0$, 또한 이들 사이의 전송지연은 멤버들 사이의 전송지연 보다 크다고 가정한다. 다음과 같은 경우에 대해서 시뮬레이션을 통해 *rseq*의 특성을 구한다.

- 경우 1 $P_{eL} = \max\{P_{eL}, P_{e2}, \dots, P_{e5}\}$.

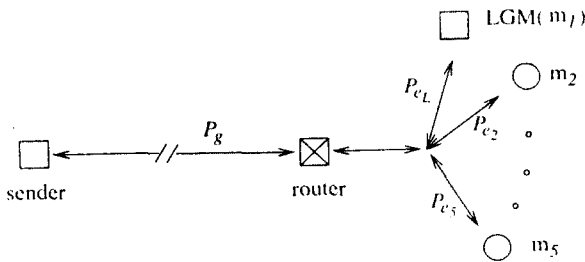


그림 4. DECMC를 위한 시뮬레이션 모델.
Fig. 4. A simulation model for the DECMC scheme.

- 경우 2 $P_{eL} = P_{e2} = \dots = P_{e5}$.
- 경우 3 $P_{e_{i \neq L}} = \min\{P_{eL}, P_{e2}, \dots, P_{e5}\}$.
- 경우 4 $P_{eL} = \min\{P_{eL}, P_{e2}, \dots, P_{e5}\}$.

먼저, LGM의 효과에 대해서 알아보자. LGM은 경우 1에서 멤버들 가운데 오류를 탐지할 확률이 가장 큰 반면 경우 4에서 가장 작고, 경우 2에서 모든 멤버들이 동일한 확률로 오류를 탐지한다. 또한 경우 4는 LGM이 아닌 다른 멤버가 가장 적은 확률로 오류를 탐지할 경우에 그 멤버의 효과를 알아보기 위함이다. 송신자가 제어메세지를 전송하는 시각마다 모든 멤버들의 *rseq*를 관찰한다.

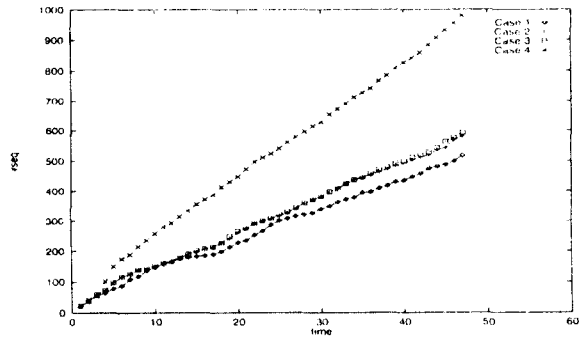


그림 5. 멤버들의 평균 rseq.
Fig. 5. The average rseq of the members.

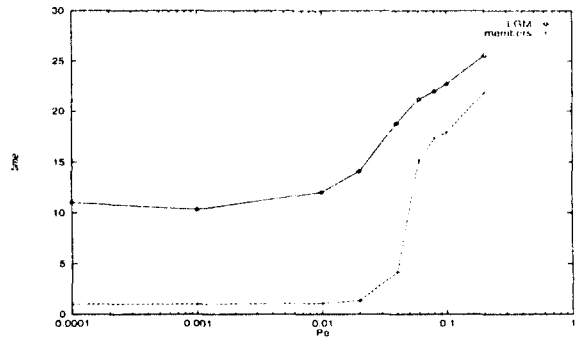


그림 6. 멤버들의 평균 오류복구 시간.
Fig. 6. The average error recovery time of the members.

그림 5에 관찰된 $rseq$ 의 특성을 나타내는 결과를 나타냈다. 여기서, 경우 1에는 $P_{e_L} = 0.10$, $P_{e_2} = \dots = P_{e_5} = 0.08$, 경우 2에는 $P_{e_L} = P_{e_2} = \dots = P_{e_5} = 0.08$, 경우 3에는 $P_{e_5} = 0.06$, $P_{e_L} = P_{e_2} = \dots = P_{e_4} = 0.08$, 경우 4에는 $P_{e_L} = 0.06$, $P_{e_2} = \dots = P_{e_5} = 0.08$ 을 사용했다. 그림에서 경우 4의 결과가 다른 경우의 결과에 비해서 상대적으로 높은 값의 $rseq$ 를 유지하고 있음을 볼 수 있다. 그 이유는 LGM이 멤버들의 오류를 복구하기에 충분한 데이터를 보유하고 있지만 다른 경우에는 그렇지 못하기 때문이다.

그림 6은 오류확률의 변화에 따라 오류복구에 소요되는 시간 T_{grc} , T_{irc} 를 측정된 결과를 나타냈다. 이 결과는 subnetwork 내의 모든 멤버들이 동일한 확률로 오류를 겪는다는 가정하에서 구한 것이다. 각 곡선은 오류 확률 P_e 가 감소함에 따라 단조감소하며 각각의 극한값에 수렴하고 있다. $P_e \leq 0.01$ 인 영역에서 T_{irc} 는 T_{grc} 의 10%를 넘지 않으면서 곡선상의 특별한 변화를 볼 수 없다. 이것은 3장의 지연시간 분석에서 구한 그림 3에서 $P_k = 0$ 인 단면에 나타나는 곡선과 유사함을 보여주고 있다.

위에서 구한 시뮬레이션 결과로 부터, 오류확률이 충분히 낮은 환경에서 DECMC방법이 사용될 경우, 오류로 인해 LGM보다 작은 값의 $rseq$ 를 갖는 멤버는 LGM과 같은 수준까지 신속하게 복구될 수 있다는 것을 알 수 있다. LGM이 다른 멤버들 보다 상태가 좋지 않은 경우에는 DECMC 방법으로 성능에 큰 변화를 주지 못한다. 그러므로 임의의 subnetwork에서 가장 좋은 상태를 유지하는 멤버에게 LGM의 역할을 담당시키는 방법을 고려해야 한다.

V. 결 론

본 논문에서는 멀티캐스트 데이터 전송에 있어서 발생하는 오류를 분산적으로 제어하는 방법을 제시하였다. 제안된 멀티캐스트 기능을 이용하는 그룹통신 모델을 나타냈고 그 그룹의 구성요소의 기능에 대해서 기술하였다. 순서번호의 특성과 오류복구시간에 대한 분석을 통해, 제안된 에러제어 방법은 멤버들이 광범위한 영역에 걸쳐서 흩어져 있고 각 subnetwork안에 몇개의 멤버들이 포함되어 있는 그룹에 적절함을 알 수 있었다.

향후에는 멀티캐스트 통신에서 발생할 수 있는 다양한

문제점들을 해결하기 위한 방법들에 대해 연구할 것이다. 특히, 멀티캐스트에서 congestion에 대처하는 문제, quality-of-service (QOS), 성능을 향상을 위한 망계층 멀티캐스트 프로토콜과 수송계층 멀티캐스트 프로토콜의 상호작용 등과 같은 분야를 다룰 것이다.

참고문헌

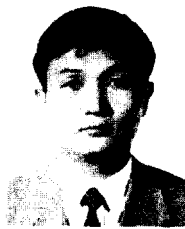
1. R. Braudes, S. Zabele, "Requirements for Multicast Protocols", RFC 1458, TSAC, May, 1993.
2. D. Waitzman, C. Partridge, and S. Deering, "Distance Vector Multicast Routing Protocol", RFC 1075, BBN STC, Stanford University, November, 1988.
3. Strayer, W., B. Dempsey, and A. Weaver, *XTP: The Xpress Transfer Protocol*, Addison-Wesley, 1992.
4. S. Armstrong, A. Freier, and K. Marzullo, "Multicast Transport Protocol", RFC 1301, Xerox, Apple, Cornell University, February, 1992.
5. S. Deering, "Host Extensions for IP Multicasting", RFC 1112, Stanford University, August, 1989.
6. D. Cheriton, "VMTP: Versatile Message Transaction Protocol", RFC 1045, Stanford University, February, 1988.
7. C. Topolcic, Editor, "Experimental Internet Stream Protocol, Version 2 (ST-II)", RFC 1190, CIP Working Group, October, 1990.
8. L. Zhang, S. Deering, D. Estrin, S. Shenker, and D. Zappala, "RSVP : A New Resource Reservation Protocol", *IEEE Network Magazine*, September, 1993.
9. D. Borman, R. Braden, and V. Jacobson, "TCP Extensions for High Performance", RFC 1323, Network Working Group, May, 1992.
10. C. Szyperski and G. Ventre, "Efficient Multicasting for Interactive Multimedia Applications", Technical Report TR-93-017.

International Computer Science Institute, Berkeley, March, 1993.

11. Shun Yan Cheung and Mostafa H. Ammar, "Using Destination Set Grouping to Improve the Performance of Window-Controlled Multipoint Connections", Technical Report GIT-CC-94-32, Georgia Institute of Technology,

August, 1994.

12. Rajesh Talpade and Mostafa H. Ammar, "Single Connection Emulation (SCE): An Architecture for Providing a Reliable Multicast Transport Service", GIT, College of Computing, Tech Report GIT-CC-94-47, October, 1994.



宋 明 烈(Myong Lyoi Song) 정회원

1963년 4월 27일생
 1985년 2월 : 연세대학교 전자공학과(공학사)
 1988년 2월 : 연세대학교 대학원 전자공학과(공학석사)
 1996년 2월 : 연세대학교 대학원 전자공학과(공학박사)

1984년 12월~1986년 1월 : 삼성전자 연구원
 1988년 12월~1991년 7월 : 육군사관학교 일반직 교수
 ※주관심 분야 : 멀티캐스트 프로토콜, 고속통신 프로토콜, 이동통신



李 俊 豪(Jun Ho Lee) 정회원

1964년 3월 16일생
 1988년 2월 : 연세대학교 전자공학과(공학사)
 1990년 8월 : 연세대학교 대학원 전자공학과(공학석사)

1996년 2월 : 연세대학교 대학원 전자공학과(공학박사)
 ※주관심 분야 : 초고속 통신망 설계, WDM, ATM 트래픽 제어

李 載 用(Jai Yong Lee)

정회원

1955년 3월 25일생

1977년 : 연세대학교 전자공학과(공학사)
 1984년 : 미국 Iowa State University, Computer Eng. (석사)
 1987년 : 미국 Iowa State University, Computer Eng. (박사)
 1977년~1982년 : 국방과학연구소 연구원
 1987년~1992년 : 포항공과대학 전자계산학과 조교수
 1992년~1994년 : 포항공과대학 전자계산학과 부교수
 1994년~현재 : 연세대학교 전자공학과 부교수

李 相 培(Sang Bac Lee)

정회원

1935년 9월 15일생

1958년 : 공군사관학교(이학사)
 1961년 : 서울대학교 공과대학(공학사)
 1964년 : 미국 Stanford 대학교(공학석사)
 1975년 : 영국 Newcastle 대학교(공학박사)
 1969년~1979년 : 서울대학교 공과대학 조교수
 1986년~1987년 : IEEE KOREA SECTION CHAIR-MAN
 1989년 1월~1989년 12월 : 대한전자공학회 부회장
 1990년 1월~1990년 12월 : 대한전자공학회 회장
 1992년 : IEEE KOREA COUNCIL CHAIRMAN
 1979년~현재 : 연세대학교 전자공학과 교수