
다단계 상호연결망에서 효율적인 멀티캐스팅 알고리즘과 성능 평가

An Efficient Multicasting Algorithm and Its Performance Evaluation in Multistage Interconnection Networks

김진수*, 장정환**

건국대학교 컴퓨터·응용과학부*, 부산외국어대학교 디지털미디어학부**

Jinsoo Kim(jinsoo@kku.ac.kr)*, Jung-Hwan Chang(jhchang@pufs.ac.kr)**

요약

본 논문에서는 영역 부호화 방식을 사용하는 다단계 상호연결망(MIN)에서의 효율적인 멀티캐스팅 알고리즘을 제안한다. 제안된 알고리즘은 원하는 멀티캐스트 목적지로 메시지를 보내기 위해, 네트워크에서 메시지를 최대 두 번을 순환시키는 순환 기법을 사용한다. 이 알고리즘은 멀티캐스트 메시지의 복사 단계와 라우팅 단계인 두 순환 단계로 구성된다. 첫 번째 단계에서 출발지는 다수의 목적지 영역을 포함시키는 영역으로 메시지를 전송하고, 이들 목적지는 이 단계에서 메시지를 수신하여 저장한다. 나머지 목적지들은 두 번째 단계에서 메시지를 최종적으로 수신한다. 본 알고리즘은 이 방법으로 전체 트래픽의 양을 줄임으로써 성능을 향상시킬 수 있다. 또한, 알고리즘의 성능을 멀티캐스트 목적지당 평균 순환 회수와 목적지당 평균 링크 수의 관점에서 기존의 알고리즘과 비교하여 평가한다.

■ 중심어 : | 효율적인 멀티캐스팅 | 다단계 상호연결망 | 영역 부호화 | 성능 평가 |

Abstract

In this paper, we propose an efficient multicasting algorithm in multistage interconnection networks (MIN's) employing the region encoding scheme. The proposed algorithm uses the recursive scheme to recycle a multicast message at most two times through MIN, in order to send it to its desired destinations. It is composed of two recycling phases which are the copying phase and the routing phase of the multicast message. In the first phase, a source sends the message to a region that contains the largest number of destination regions, and destinations in these regions receive and store the message in this phase. The remaining destinations can finally receive the message in the second phase. This method of the algorithm can improve its performance by reducing the delay of message and the volume of traffic. Moreover, we evaluate the performance of our algorithm in terms of the average number of recycling and the number of internal links used per destination, comparing with the previously proposed algorithm.

■ keyword : | Efficient Multicasting | Multistage Interconnection Network | Region Encoding | Performance Evaluation |

* 본 논문은 2003년도 건국대학교 학술진흥연구비 지원에 의해 수행되었습니다.

1. 서론

멀티캐스트 통신(multicast communication)은 집합체 통신의 가장 중요한 형태 가운데 하나로서[1], 하나의 출발지에서 임의의 다수 목적지로 동일한 메시지를 전송하는 일대다 통신을 의미한다. 최근 원격 교육, 다자간 영상회의, VOD(video on demand) 등과 같은 멀티미디어 통신서비스의 수요가 꾸준히 증가함에 따라, 멀티캐스트는 정보통신, 병렬 및 분산 시스템의 기반 네트워크에서 필수적으로 제공되어야 하는 통신 기능의 하나로 부각되고 있다.

다단계 상호연결망(Multistage Interconnection Network, MIN)은 다중컴퓨터 시스템에서 프로세싱 노드들 간의 연결을 위해 사용될 뿐만 아니라, 광대역 종합 정보 통신망(B-ISDN)을 위한 ATM(Asynchronous Transfer Mode) 스위칭 시스템의 내부 연결구조로도 보편적으로 사용되고 있다. 이것은 MIN의 입력단에서 출력단으로 오직 하나의 라우팅 경로만 존재하며, 이 특성으로 고속의 셀프 라우팅 기능을 제공하기 때문이다. 또한, MIN은 임의의 입력단에서 모든 출력단으로 네트워크를 한번 통과하여 접근이 가능한 특성을 지니고 있다.

MIN에서 멀티캐스트 통신을 구현하는 많은 연구가 진행되어 왔다[2]. 이들은 크게 하드웨어 접근 방법과 소프트웨어 접근 방법으로 구분할 수 있다. 전자의 대표적인 방법은 라우팅 네트워크(routing network) 앞단에 복사 네트워크(copy network)를 두어 멀티캐스트 목적지의 수만큼 복사를 한 후, 복사된 메시지를 라우팅 네트워크를 통해 원하는 최종 목적지로 라우팅하여 멀티캐스트를 수행하는 것이다[3-5]. 또한, 전송속도를 향상하기 위해, 다수의 MIN을 병렬로 연결함으로써 복사와 라우팅의 기능을 수행하는 연구도 진행되었다[6]. 그런데, 이들 하드웨어 접근 방법은 메시지의 복사를 위한 추가적인 하드웨어 비용이 들고, 복잡한 제어 기능을 필요로 한다.

소프트웨어 접근의 대표적인 방식은 스위칭 소자(Switching Element, SE)에 브로드캐스팅(broadcasting) 기능을 부가하여, 부수적인 하드웨어 없이 멀티캐스팅을 제공하는 것이다[7-11]. 이 방법에서는 네트워크 상에서 메시지를 여러 번 순환시킴으로써, 멀

티캐스트 목적지 수만큼의 메시지를 복사하면서 라우팅을 수행한다. 이러한 기법을 순환 기법(recursive scheme)이라고 하며[7], 하나의 MIN 자체가 복사와 라우팅 기능을 동시에 처리한다. 순환 기법에서 전송 지원은 메시지의 순환 회수에 의해 크게 영향을 받는다.

Cusani[7]과 Chen[8]의 방식은 멀티캐스트를 위해 많은 수의 순환을 필요로 하는 반면, Raghavendra[9]와 Park[10]은 두 번의 순환으로 멀티캐스팅을 수행하는 알고리즘을 제안하였다. Raghavendra의 알고리즘은 순환 단계별로 다른 헤더 부호화 방식(header encoding scheme)을 사용하기 때문에 스위칭 소자의 복잡도가 증가하는 단점을 지닌다. Kim[11]은 Park의 방식을 확장하여, 다수의 스위칭 소자가 고장이 발생한 경우에도 두 번의 순환으로 멀티캐스팅을 수행하는 알고리즘을 제안하였다. 이들 알고리즘의 경우, 첫 번째 순환에서 멀티캐스트 목적지 그룹 수만큼을 복사하기 위해 무작위로 선택한 중간 경유지 노드로 메시지를 라우팅한 후, 두 번째 순환에서 최종 목적지 영역들로 각각 라우팅을 수행한다. 즉, 모든 멀티캐스트 메시지는 반드시 두 번의 순환 단계를 거쳐야 한다.

본 논문은 MIN에서 효율적인 순환을 통해 성능을 향상시킨 멀티캐스팅 알고리즘을 제안한다. 이 알고리즘은 연속된 주소를 하나의 영역으로 나타내는 영역 부호화(region encoding) 방식을 사용하여 임의의 목적지를 갖는 멀티캐스트 메시지를 구성한다. 또한, 네트워크를 통해 메시지를 기껏해야 두 번 순환을 시킴으로써 멀티캐스팅을 수행한다. 첫 번째 순환 단계에서 목적지 영역을 최대로 포함하는 중간 경유지를 선택하고, 이 목적지 영역들에 속한 목적지는 한 번의 순환으로 메시지 수신을 완료하도록 개선하여, 전체 트래픽을 줄이고 성능을 향상시킨다.

본 논문의 구성은 다음과 같다. 2장에서는 MIN의 구조에 대해 기술하고, 멀티캐스트 주소를 표현하기 위한 영역 부호화 방법에 대해 설명한다. 3장에서는 중간 경유지의 선택 방법과 MIN에서의 효율적인 멀티캐스팅 알고리즘을 기술하고, 그 예를 설명한다. 4장에서는 알고리즘의 성능을 목적지당 평균 순환 회수와 멀티캐스팅에서 사용되는 내부 링크 수의 관점에서 분석하고, 기존의

알고리즘과 비교 평가한다. 그리고 5장에서 결론을 맺는다.

II. 시스템 구조와 기본 용어

1. 네트워크의 구조

크기가 N 인 다단계 상호연결망은 $n = \log_2 N$ 개의 단(stage)으로 구성되고, 각 단은 $N/2$ 개의 2×2 스위칭 소자들로 이루어진다. 각 단에서 N 개의 입출력 링크는 위부터 아래로 0에서 $N-1$ 까지 n -비트로 표현되고, $N/2$ 개의 SE는 위부터 아래로 0에서 $N/2-1$ 까지 $(n-1)$ -비트로 표현된다. 각 단의 번호는 좌측부터 우측으로 $n-1$ 에서 0로 표현된다.

[그림 1]은 논문에서 대상으로 하는 MIN 기반의 다중 컴퓨터의 예를 보인다. 그림 1에서 보는 바와 같이, 인접한 두 단간 버터플라이(butterfly) 연결 형태, 노드들과 제일 좌측의 $(n-1)$ 단 사이는 완전 셔플(perfect shuffle) 연결형태를 갖는다. 그리고 메시지의 순환을 위해, [그림 1]의 점선과 같이, 우측의 출력단에서 노드쪽으로 메시지의 순환을 위한 역방향 링크들이 존재한다. 그리고, 멀티캐스트를 지원하기 위해, 모든 SE들은 브로드캐스팅 기능을 갖는다.

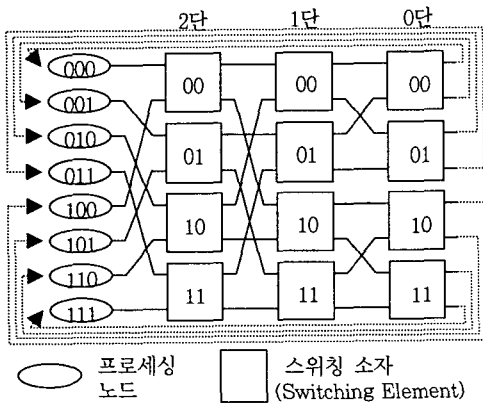


그림 1. MIN 기반의 다중컴퓨터 예 ($N=8$).

2. 영역 부호화 방식

일대다 통신 형태를 갖는 멀티캐스팅에서 목적지의 개수는 일정하지 않다. 이러한 멀티캐스트 메시지의 목적지들을 정형화된 형태로 표현하기 위해 여러 가지 방식들이 사용되는데, 다단계 상호연결망 기반의 시스템에서 널리 사용되는 방식 중 하나가 영역 부호화 방식이다 [4][10][11]. 영역 부호화 방식은 연속된 주소를 갖는 다수 목적지들을 하나의 영역(region), 즉 최소와 최대 주소의 쌍, [최소주소, 최대주소]로 표현하는 방식이다. 예를 들어 멀티캐스트 목적지가 1, 2, 3, 5일 때, 목적지들을 두 개의 영역인 [1, 3], [5, 5]로 나타낼 수 있다. 즉 하나의 영역은 1개 이상의 연속된 목적지 주소로 구성되는 집합으로 볼 수 있고, [1, 3]은 {1, 2, 3}을, [5, 5]는 {5}를 의미한다.

영역 부호화 방식을 사용하는 MIN에서 SE는 [최소주소, 최대주소]인 멀티캐스트 라우팅 헤더를 처리해야 한다. i ($n-1 \geq i \geq 0$) 단의 SE는 헤더가 최소주소 $min = m_{n-1} \dots m_i m_{i-1} \dots m_0$ 과 최대주소 $MAX = M_{n-1} \dots M_i M_{i-1} \dots M_0$ 를 갖는 메시지를 받으면, 다음과 같이 메시지를 전송한다.

- (1) $m_i = M_i = 0$ 이거나 $m_i = M_i = 1$ 이면, 헤더의 변경 없이 출력 포트 0이나 1로 메시지를 전송함
- (2) $m_i = 0$ 이고 $M_i = 1$ 이면, 출력 포트 0과 1로 메시지를 전송하되 각각 다음과 같이 헤더를 변경함.

- 출력 포트 0으로 나가는 메시지의 헤더 : $[min, M_{n-1} \dots M_{i+1} 0 1 \dots 1]$

- 출력 포트 1으로 나가는 메시지의 헤더 : $[m_{n-1} \dots m_{i+1} 1 0 \dots 0, MAX]$

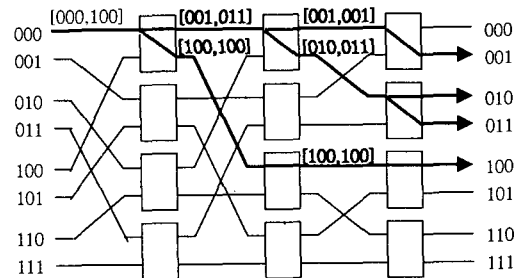


그림 2. 영역 부호화 방식의 라우팅과 헤더 예.

[그림 2]는 출발지가 0이고 목적지가 1, 2, 3, 4인 멀티캐스트 메시지의 라우팅 예를, 라우팅 경로 상에 있는 SE들의 입력과 출력 포트의 헤더 변화와 함께 보인다. 단의 수가 3이므로 출발지는 000, 목적지 영역은 [001, 100]로 표시된다. 1단의 SE 10은 헤더가 [100,100]인 메시지를 받고, $m_1=M_1=0$ 이므로 헤더의 변경 없이 출력 포트 0으로 메시지를 전송한다. 1 단의 SE 00은 헤더가 [001,011]인 메시지를 받고, $m_1=0$ 이고 $M_1=1$ 이므로, 헤더가 [001,001]인 메시지를 출력 포트 0으로 헤더가 [010,011]인 메시지를 출력 포트 1로 전송한다.

3. 용어의 정의

본 논문에서는 목적지의 개수와 분포가 가변적인 멀티캐스트 목적지들을 다수 목적지 영역들로 표현한다.

[정의-1] $D = \{R_1, R_2, \dots, R_m\}$ 이 멀티캐스트 목적지들 나타낼 경우, 이러한 집합 D 를 목적지 영역 집합이라고 정의하며, 본 논문에서 m 은 목적지 영역의 수를 표현한다. 그리고 영역 $R_i = [d_i^x, d_i^y]$ 과 $R_j = [d_j^x, d_j^y]$ ($1 \leq i < j \leq m$)의 길이, $R_i \sim R_j$ 은 $d_j^x - d_i^x + 1$ 의 값으로 정의한다.

[정의-2] 플래그는 목적지 영역의 종류를 나타내는 m -비트 Boolean 변수이며, 문자 f 로 표시한다. f_j 가 0이면 한 번의, f_j 가 1이면 두 번의 순환이 필요한 목적지 영역임을 나타낸다.

III. 효율적인 멀티캐스트 알고리즘

본 알고리즘은 최대 두 단계의 순환으로 멀티캐스팅을 수행한다. 단계 1에서는, 출발지에서 멀티캐스트 목적지 영역의 개수의 크기를 갖는 중간 경유지 영역 IR (Intermediate Region)로 메시지를 복사한다. 기존 알고리즘은 이 IR 을 무작위로 선택하는 반면, 본 알고리즘에서는 최종 목적지 영역들을 최대로 포함하는 IR 을 선택하고, 이들 목적지 영역에 대해 단계 1에서 메시지의 수신을 완료시킴으로써 단계 2에서의 라우팅을 생략한다. 단계 2에서는, IR 에 포함되지 않은 최종 목적지 영역들에 대해서만 복사된 메시지를 각각 라우팅을 한다.

1. 중간 경유지 영역의 선택

출발지는 목적지 영역을 최대한 포함할 수 있도록 IR 을 선택함으로써, 멀티캐스트 통신의 성능을 향상시킬 수 있다. [그림 3]은 이러한 IR 을 선택하고, 단계 2에서 해당 목적지 영역을 라우팅할지 여부를 나타내는 플래그를 생성하는 알고리즘을 보이고 있다. [그림 3]에서 step 1에서 3까지는 목적지 영역을 최대로 포함하는 IR 을 선택하는 과정이고, step 4에서 플래그를 생성한다. step 1.2에서, $no_region(i)=0$ 은 $|IR_i| < |R_i|$ 경우에 발생한다. 알고리즘의 시간 복잡도는 step 1에 의존한다. 목적지 영역 집합 $\{R_1, R_2, \dots, R_m\}$ 이 이미 정렬이 되어 있으므로, step 1.1에서 j 를 찾는 것은 이진 탐색 방법을 이용하여 $\log m$ 시간이면 수행이 가능하다. 따라서 이 알고리즘의 시간 복잡도는 $O(m \log m)$ 이다.

입력 : 목적지 영역 집합 $\{R_1, R_2, \dots, R_m\}$
 출력 : IR , 플래그 f

```

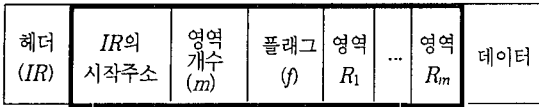
1. for  $i = 1$  to  $m$  do
1.1  $R_i \sim R_j \leq m$ 을 만족하는 최대 값  $j$ 를 구한다.
1.2  $no\_region(i) \leftarrow j - i + 1$ ; /*  $i-1 \leq j \leq m$  */
2.  $\max(no\_region(i), 1 \leq i \leq m)$ 인  $i$ 를 구한다.
3.  $R_i \subseteq IR, \dots, R_{i+no\_region(i)} \subseteq IR$  인  $IR$ 을 생성한다.
4. for  $t = 1$  to  $m$  do /* 플래그  $f = f_1 \dots f_m$  생성 */
4.1 if  $(i \leq t \leq i+no\_region(i)-1)$  then  $f_t \leftarrow 0$ 
4.2 else  $f_t \leftarrow 1$ 
        
```

그림 3. 중간 경유지 선택과 플래그 생성 알고리즘.

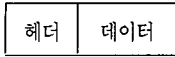
예를 들어, 목적지 영역 집합이 $\{[1, 3], [5, 5], [7, 8], [12, 14]\}$ 이면, $m=4$ 이므로 배열 no_region 은 $\{1, 2, 1, 1\}$ 이 된다. 따라서 step 2에서 선택된 i 는 2가 되어, IR 은 $[5, 8]$ 로 선택되고 플래그는 1001로 생성된다.

2. 멀티캐스트 메시지의 형식

[그림 4]는 알고리즘의 단계 1과 단계 2에서 사용되는 두 가지 종류의 메시지 형식을 보이고 있다.



(a) 메시지 형식1 (단계 1)



(b) 메시지 형식2 (단계 2)

그림 4. 단계별 멀티캐스트 메시지 형식.

[그림 4](a)는 단계 1에서의 메시지 형식이며, 출발지 노드에서 중간 경유지 영역으로 라우팅할 때 사용된다. 여기서 굵은 사각형안의 필드들은 단계 2에서 라우팅 시 필요한 추가적인 정보이다. [그림 4](b)는 단계 2에서의 메시지 형식이며, 중간 경유지 영역의 노드에서 최종 목적지 영역으로 라우팅할 때 사용된다.

3. 멀티캐스팅 알고리즘의 기술

제안하는 알고리즘은 [그림 5]와 같이 두 단계로 구성된다.

가정

1. 멀티캐스트 목적지는 목적지 영역 집합으로 구성된다.
2. 플래그는 $f = f_1 \cdots f_m$ 로 표시한다.
3. 상대주소 $rel = (\text{자신의 노드 주소} - IR \text{의 시작 주소}) \bmod m$ 이다.
4. 단계별 메시지는 그림 4의 형식으로 구성된다.

알고리즘

단계 1 :

- 송신자: 출발지 노드는 멀티캐스트 메시지의 복사 및 전송을 위해, 그림 3의 알고리즘에서 선택한 $IR=[s, s+m-1]$ 로 메시지를 라우팅한다.

- 수신자: IR 에 속한 중간 경유지 노드는 메시지를 받은 후, 다음을 만족하는 노드 dn 은 최종 목적지로서 메시지를 저장한다.

• $dn \in R_x$ 이고 $f_x = 0$ ($1 \leq x \leq m$)

단계 2 :

- 송신자: IR 에 속한 중간 경유지 노드 $s+rel$ 은 단계1에서 수신한 메시지를 다음 조건에 따라 대응되는 목적지 영역 R_{rel} 으로 라우팅한다($0 \leq rel \leq m-1$).

• $f_{rel} = 1$ 이면, 라우팅을 수행한다.

• $f_{rel} = 0$ 이면, 자신의 메시지를 무시한다.

그림 5. 효율적인 멀티캐스팅 알고리즘 (EMA).

단계 1에서, 목적지 영역의 개수 크기(= m)의 중간 목적지 영역 IR 을 [그림 3]의 알고리즘에 따라 선택한 후, IR 로 메시지를 송신한다. IR 에 포함된, 즉 $R_i \subseteq IR$ ($1 \leq i \leq m$)인, 목적지 영역 R_i 에 속한 목적지 노드들은 그 플래그의 값이 $f_i=0$ 이므로, 이 값에 따라 메시지를 최종 수신하여 자신의 지역 메모리에 저장한다.

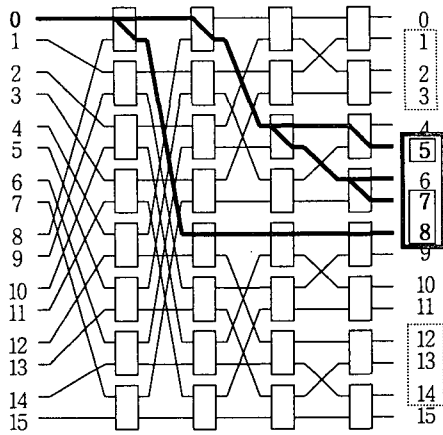
단계 2에서, IR 에 속한 노드들은 플래그 값에 따라, 단계 1에서 수신된 메시지를 최종 목적지 영역들로 각각 송신한다. 메시지에 포함된 IR 의 시작 주소를 이용하여, 상대주소 값 $rel = (\text{자신의 노드 주소} - IR \text{의 시작 주소}) \bmod m$ 을 계산하고, 이 값에 해당하는 플래그 비트 값(f_{rel})이 1이면 라우팅을 하고, f_{rel} 이 0이면 무시한다. 또한 이 상대주소 rel 을 이용하여, 자신이 라우팅해야 하는 목적지 영역을 선택할 수 있으며, 선택된 영역을 새로운 라우팅 헤더로서 생성한 메시지를 라우팅한다.

4. 알고리즘의 예

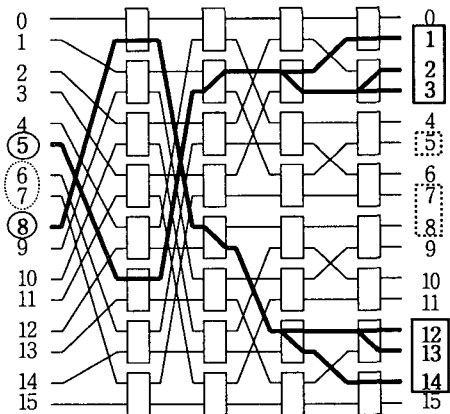
[그림 6]은 출발지 0에서 목적지 1, 2, 3, 5, 7, 8, 12, 13, 14로 멀티캐스트 메시지를 전송할 경우, 알고리즘의 단계별 라우팅 예를 보이고 있다. 목적지 영역 집합은 $\{[1, 3], [5, 5], [7, 8], [12, 14]\}$ 이며, $m=4$, $IR=[5, 8]$, $f=1001$ 이 된다.

단계 1에서는 [그림 6](a)와 같이, 0에서 [5, 8]로 메시지를 라우팅한다. [그림 6](a)에서, 굵은 사각형은 IR 을 나타내며, 실선 사각형과 같이 IR 에 포함된 [5, 5]와 [7, 8]은 이미 멀티캐스트 메시지의 수신이 완료된 목적지 영역임을 알 수 있다. 반면에, 점선 사각형은 아직 메시지를 받지 못한 나머지 목적지 영역 [1, 3]과 [12, 14]를 표시하고 있다.

[그림 6](b)는 단계 2의 라우팅 경로를 보여준다. 플래그 f 가 1001이므로, IR 에 속한 중간 경유지 노드 5와 8은, 실선 원과 사각형과 같이, 각각 목적지 영역 [1, 3]과 [12, 14]로 멀티캐스트 메시지를 전송하고 있다. 반면, 점선 원과 사각형과 같이, 중간 경유지 노드 6과 7은 메시지를 송신하지 않으며, 목적지 영역 [5, 5]와 [7, 8]은 메시지를 받지 않는다.



(a) 단계 1 (메시지의 복사/전송)



(b) 단계 2 (메시지의 전송)

그림 6. 알고리즘의 단계별 멀티캐스팅 예.

IV. 성능 평가 및 분석

멀티캐스트 알고리즘의 성능은 그 알고리즘에 의해 표현되는 멀티캐스트 트리의 깊이(depth)와 크기(volume)에 의해 평가될 수 있으며, 이들 평가 척도는 모두 값이 작을수록 우수한 성능을 보이는 것이다. 순환을 이용하는 멀티캐스트 알고리즘의 경우, 멀티캐스트 트리의 깊이는 목적지 당 평균 순환 횟수로 측정되며, 트리의 크기는 목적지 당 사용되는 링크의 수로 측정된다. 이 장에서는 본 논문에서 제안한 알고리즘(EMA)과 기존에 제안된 Park[10]의 영역 부호화 방식 알고리즘(MA라고 표

기)의 성능을 상기 두 척도에 대하여 비교 평가한다.

1. 목적지당 평균 순환 횟수

멀티캐스트 목적지 당 평균 순환 횟수는 멀티캐스트 트리의 깊이에 해당되며, 즉 메시지의 전송지연(latency)을 의미한다. MA의 경우, 모든 멀티캐스트 목적지에 대해 두 번의 순환을 통해 멀티캐스팅을 수행되므로, 목적지 당 평균 순환 횟수가 2로서 항상 일정하다. 그러나 EMA의 경우, 두 번째 순환에서 제외되는 목적지 영역 $R(R \subseteq IR)$ 에 속한 목적지에 대해서는 1이며, 그 이외의 모든 목적지 경우에 대해서는 2가 된다. 목적지 영역집합을 $D = \{R_1, R_2, \dots, R_m\}$ 로 나타낼 경우, 목적지당 평균 순환회수 ANR은 다음과 같은 수식으로 표시된다.

$$ANR(EMA) = \left(\sum_{i=1}^m |R_i| \times 2 - \sum_{R \in IR} |R| \right) / \sum_{i=1}^m |R_i|$$

[그림 7]은 네트워크의 크기 N , 즉 단의 수 $n (= \log_2 N)$ 의 변화에 따라, 두 알고리즘의 목적지 당 평균 순환 횟수에 대해 비교한 결과이다. 그림에서 p 는 출력 링크가 멀티캐스트 목적지에 포함될 확률이며, p 가 0.05, 0.10, 0.25인 경우에 대한 결과를 보이고 있다. 네트워크 크기가 큰 경우는 멀티캐스트 목적지 수가 많을수록 더 우수한 성능을 보이는 반면, 작은 네트워크에서는 멀티캐스트 목적지 수가 적을수록 더 성능이 우수함을 알 수 있다. 이는 p 가 높아지고 네트워크의 크기가 커져서 목적지 수가 많아짐에 따라, 중간 경유지 영역의 크기 IR 이 커져서 더 많은 수의 목적지 영역을 포함할 가능성이 높아지기 때문으로 판단된다.

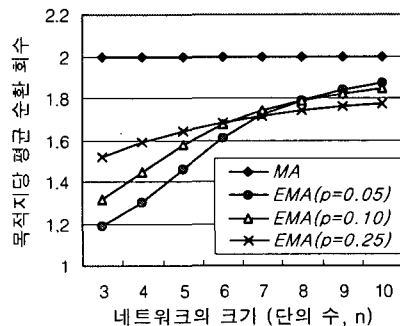


그림 7. 목적지당 평균 순환 회수.

2. 목적지당 사용되는 평균 내부 링크 수

목적지 당 사용되는 평균 내부 링크 수는 멀티캐스트 트리의 크기에 해당되며, 즉 멀티캐스트 메시지에서 발생하는 내부 트래픽의 양을 의미한다. 이것의 분석을 위해, 알고리즘 A에 의해 사용되는 전체 링크 수는 $TNL(A)$ 로 표시하고 평균 링크 수는 $ANL(A)$ 로 표시하고, 이를 구하고자 한다.

하나의 멀티캐스트 메시지의 목적지 영역 R 을 $[a,b]$ 로 표시하고, a 와 b 의 n -비트 표현에서 일치하는 최상위 비트의 개수를 $s(a,b)$ 라고 표시하자. 이 메시지는 $(n-s(a,b)-1)$ 단에 있는 SE에서 첫 번째로 브로드캐스팅 기능을 수행한다. 그 SE에서 출력포트 0번과 1번을 통과하여 도달하는 목적지의 수 $Up(a,b)$ 와 $Down(a,b)$ 는 각각 다음과 같이 표시된다.

$$Up(a, b) = 2^{n-s(a,b)-1} - a \text{ mod } 2^{n-s(a,b)-1}$$

$$Down(a, b) = b \text{ mod } 2^{n-s(a,b)-1} + 1$$

따라서, 이 메시지가 사용하는 전체 링크의 수 $NL(R)$ 는 다음과 정리할 수 있다.

$$NL(R) = s(a, b) + \sum_{i=0}^{n-s(a,b)-1} \left(\lceil \frac{Up(a, b)}{2^i} \rceil + \lceil \frac{Down(a, b)}{2^i} \rceil \right) \quad (1)$$

EMA과 MA의 단계 1에서 중간 경유지 영역을 IR_1 과 IR_2 라고 하면, 각 알고리즘에서 사용되는 링크 수는 $NL(IR_1)$ 와 $NL(IR_2)$ 로 나타낼 수 있다. 두 알고리즘에서 중간 경유지 영역의 선택에 있어 차이는 있으나, 그 크기는 동일 ($|IR_1| = |IR_2|$) 하므로, $NL(IR_1)$ 와 $NL(IR_2)$ 는 같다고 할 수 있다.

MA의 단계 2에서 사용되는 링크 수는, 식(1)을 사용하면, $\sum_{R \in D} NL(R)$ 이다. 또한, EMA의 단계 2에서는 이미 단계 1에서 수신 완료된 목적지 영역들을 제외해야 하므로, 단계 2에서 사용되는 전체 링크 수는 $\sum_{R \in D} NL(R) - \sum_{R \in IR} NL(R)$ 이다. 따라서 $TNL(MA)$ 과

$TNL(EMA)$ 은 식(1)을 사용하여 각각 다음과 같이 표현된다.

$$TNL(MA) = NL(IR_1) + \sum_{R \in D} NL(R) \quad (2)$$

$$TNL(EMA) = NL(IR_1) + \sum_{R \in D} NL(R) - \sum_{R \in IR} NL(R) \quad (3)$$

따라서 멀티캐스트 목적지당 사용되는 평균 링크 수 $ANL(MA)$ 와 $ANL(EMA)$ 은 식(2)와 (3)를 사용하면 다음과 같다.

$$ANL(MA) = TNL(MA) / \sum_{R \in D} |R|$$

$$ANL(EMA) = TNL(EMA) / \sum_{R \in D} |R|$$

[그림 8]은 네트워크의 크기와 p 의 변화에 따라, 두 알고리즘에서 사용되는 내부 링크의 수의 관점에서 비교한 결과를 보여준다. 멀티캐스트 목적지 수가 적은 경우($p = 0.05$)에는 네트워크 크기가 작을수록 성능의 차이가 커지는 결과를 보였다. 멀티캐스트 메시지 당 목적지 수가 많은 경우($p = 0.25$)는 목적지 수가 적은 경우에 비해, 네트워크 크기가 사용되는 평균 링크 수의 단순 차이에 주는 영향은 작았다. 그리고 p 에 상관없이, 네트워크 크기에 따른 링크 수의 증가 폭선이 선형에 가깝게 나타났다. 또한 모든 경우에 EMA가 MA에 비해 좋은 성능을 보임을 알 수 있다.

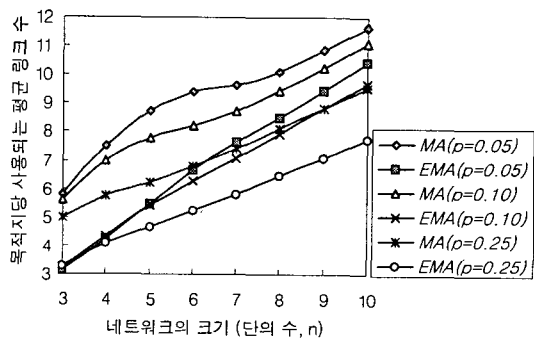


그림 8. 목적지당 사용되는 평균 내부 링크 수.

V. 결론

본 연구에서는 순환 구조를 갖는 다단계 상호연결망에서 효율적으로 멀티캐스팅을 수행하는 알고리즘을 제안하였다. 제안된 알고리즘은 영역 부호화 방식을 사용하여 임의의 멀티캐스트 목적지를 표현하였다. 이 알고리즘은 기존의 여러 멀티캐스트 알고리즘과 같이, 멀티캐스팅을 수행하기 위해 네트워크에서 메시지를 순환시키는 기법을 사용하였다. 이 순환 기법을 이용한 알고리즘들은 최소 두 번의 순환을 통해 멀티캐스팅을 수행하였다.

본 논문에서는 두 번의 순환을 하는 기존 멀티캐스트 알고리즘에서, 중간 경유지 영역을 무작위로 선택하고, 이 최종 목적지 영역이 이미 중간 경유지 영역에 포함하고 있을 경우에도 두 번째 단계의 라우팅을 한다는 것을 관찰하였다. 그리고 두 번째 순환에서 이 최종 목적지 영역을 라우팅하지 않도록 개선하고, 가능하면 다수의 목적지 영역을 포함하는 중간 경유지를 선택하도록 개량함으로써, 전체 트래픽을 줄이고 성능을 향상시켰다. 또한, 알고리즘의 성능을 멀티캐스트 목적지당 평균 순환 회수와 목적지 당 평균 링크 수의 관점에서 분석하였다. 실험을 통해 기존에 제안된 알고리즘과 비교한 결과, 두 평가 척도에 대해 모두 우수한 성능을 보임을 알 수 있었다. 이 알고리즘은 다단계 상호연결망의 고장 허용 멀티캐스팅에도 쉽게 확장 적용할 수 있을 것으로 사료된다.

참고 문헌

[1] J. Duato, S. Yalamanchili, and L. Ni, *Interconnection Networks: An Engineering Approach*, Morgan Kaufmann, 2002.
 [2] N. F. Mir, "A Survey of Data Multicast Techniques, Architectures, and Algorithms," *IEEE Communications Magazine*, Vol.39, No.9, pp.164-170, Sep, 2001.
 [3] J. S. Turner, "Design of a Broadcast Packet Switching Network," *IEEE Trans. on Commun.*, Vol.36, pp.734-743, June, 1988.

[4] T. T. Lee, "Nonblocking Copy Networks for Multicast Packet Switching," *IEEE JSAC*, Vol.6, pp.1455-1467, Dec, 1988.
 [5] W. D. Zhong, J. Kaniyil, and Y. Onozato, "A Copy Network with Shared Buffers for Large-Scale Multicast ATM Switching," *IEEE/ACM Transactions on Networking*, Vol.1, No.2, pp.157-165, Apr, 1993.
 [6] Y.-W. Deng and W.-T. Chen, "Efficient Parallel Multicast ATM Switch," *IEE Proceedings-Communications*, Vol.147, No.2, pp.123-132, Apr, 2000.
 [7] R. Cusani and F. Sestini, "A Recursive Multistage Structure for Multicast ATM Switching," *Proc. of IEEE Infocom*, pp.1289-1295, Apr, 1991.
 [8] X. Chen and V. Kumar, "Multicast Routing in Self-Routing Multistage Networks," *Proc. of IEEE Infocom*, pp.306-314, Apr, 1994.
 [9] C. S. Raghavendra, X. Chen, and V. P. Kumar, "A Two Phase Multicast Routing Algorithm in Self-Routing Multistage Networks," *Proc. of Int'l. Conference on Communications*, pp.1612-1618, June, 1995.
 [10] J. Park and H. Yoon, "Novel Recursive Multicast Algorithms in Multistage Interconnection Networks for ATM Switches," *Journal of High Speed Networks*, Vol.8, No.4, pp.265-280, 1999.
 [11] 김진수, 장정환, "다단계 상호연결망에서 영역부호화 방식을 사용하는 고장 허용 멀티캐스팅 알고리즘", *한국정보과학회 논문지(시스템 및 이론)*, 제29권, 제3/4호, pp.17-124, 2002.

저자 소개

김진수(Jinsoo Kim)

정회원



- 1983년 2월 : 서울대학교 컴퓨터 공학과(공학사)
- 1985년 2월 : 한국과학기술원 전산학과(공학석사)
- 1998년 8월 : 한국과학기술원 전산학과(공학박사)

- 1985년 4월~2000년 2월 : KT 선임연구원
- 2000년 3월~현재 : 건국대학교 컴퓨터·응용과학부 교수

<관심분야> : 병렬처리, 고속통신망, 센서네트워크

장정환(Jung-Hwan Chang)

정회원



- 1983년 2월 : 경북대학교 전자공학과(공학사)
- 1985년 2월 : 한국과학기술원 전산학과(공학석사)
- 1998년 8월 : 한국과학기술원 전산학과(공학박사)

- 1985년 4월~2000년 8월 : KT 선임연구원
- 2000년 9월~현재 : 부산외국어대학교 디지털미디어 학부 교수

<관심분야> : 상호연결망, 그래프 이론, 통신망 보안