

# 스위치 기반의 비정규적 네트워크에서의 차별적인 다중 전송

## (Differential Multicast in Switch-Based Irregular Topology Network)

노 병 권 <sup>†</sup>      김 성 천 <sup>\*\*</sup>  
(Byoung-Kwon Roh) (Sung-Chun Kim)

**요 약** 워크스테이션들의 네트워크(NOWs)와 같은 스위치 기반의 비정규적 네트워크는 고속 스위치를 이용하여 프로세서 노드나 네트워크 호스트들을 특정 위상에 구애받지 않고 연결한 것으로 우수한 유연성(flexibility)과 확장성(scalability)의 장점이 있다. 그러나 메쉬와 같은 정규적인 위상에 비해 교착 상태(deadlock)에 빠지지 않는 효율적인 다중 전송을 수행하는 것은 용이하지 않다. 본 논문에서는 전송 트리의 루트 노드에 무리를 주지 않고 할당받은 서비스에 따라서 전송 단계와 적응성에 차별성을 부여하며 스위치의 채널 사용의 불균형을 개선하고 네트워크 부하 분포의 조정이 가능한 단일 단계 차별적 다중 전송(Single Phase Differential Multicast; SPDM) 기법을 제안한다.

키워드 : NOW, 위상, 비정규적 네트워크, 메쉬, 교착상태

**Abstract** Networks of Workstations(NOWs), that has features of flexibility and scalability, recently has emerged as an inexpensive alternative to massively parallel multicomputers. However it is not easier to perform deadlock-free multicast than regular topologies like mesh or hypercube.

Single phase differential multicast(SPDM) is a modified multicast algorithm with less burden of the root node. By applying quality of service(QoS), a specific node can have differentiated service and artificial change of message flow pattern is also available. As the results of performance evaluation experiments, SPDM has lower latency and lower packet concentration rate of the root node than the case of SPAM, and has ability to control network load distribution among switch nodes by controlling the assignment rate among nodes.

**Key words** : NOW, topology, irregular network, mesh, deadlock

### 1. 서 론

고성능 병렬 컴퓨팅 환경에서는 보다 나은 성과를 얻기 위해 하나 이상의 프로세서나 컴퓨터들을 상호 연결하여 병렬적으로 작업을 수행한다. 일반적으로 정규적인(regular) 위상(topology)의 상호 연결 형태는 구조의 명확함과 라우팅의 용이함 등의 장점으로 인하여 널리 이용되어 왔다. 하지만 유연성(flexibility)과 확장성(scalability) 등에서 우수함을 보이는 비정규적(irregular) 위상이 워크

스테이션들의 네트워크(Networks Of Workstations: NOWs)[1,2,3,4] 등에 적용됨으로서 차츰 주목을 받게 되었다. 정규적인 위상에 비해 비정규적 위상에서 효율적이고 교착 상태(deadlock)에 빠지지 않는 다중 전송의 수행이 용이하지 않으므로 이에 대한 연구가 진행되어 왔다.

단일 단계 적응적 다중 전송(Single Phase Adaptive Multicast ; SPAM) 기법은 이러한 비정규적인 네트워크 환경에서 트리를 기반으로 하여 교착 상태에 빠지지 않는 다중 전송을 수행한다[5]. 이 기법은 트리에 포함되지 않는 노드 사이의 연결까지 활용하여 교착 상태를 방지할 수 있고 플릿 버퍼의 크기에 무관하게 전송이 가능하다는 장점이 있지만 쉽게 루트 노드에 부담을 줄 수 있다는 잠재적인 단점을 가지고 있다.

본 논문에서는 이러한 SPAM 기법의 문제점을 개선하여 목적지가 증가하더라도 루트에 부담을 주지 않으면서

· 본 연구는 한국과학재단 목적기초연구(R01-2001-000-00356-0) 지원으로 수행되었음.

<sup>†</sup> 비 회 원 : 서강대학교 공과대학 컴퓨터학과  
porori@archilab.sogang.ac.kr

<sup>\*\*</sup> 중 심 회 원 : 서강대학교 공과대학 컴퓨터학과 교수  
ksc@arqlab1.sogang.ac.kr

논문접수 : 2000년 4월 11일

심사완료 : 2002년 5월 2일

다중 전송이 가능한 단일 단계 차별적 다중 전송(Single Phase Differential Multicast; SPDM) 기법을 제안한다. 이와 동시에 서비스의 질(QoS)을 고려함으로써 특정 호스트 노드에 다른 호스트들과 차별되는 형태의 전송이 가능하고 차별된 채널 선택에 의해 메시지 흐름의 제어가 가능해지므로 네트워크 위상에서 야기되는 채널 사용과 부하 분포의 불균형을 다소 개선한다.

## 2. 단일단계 적응적 다중전송 기법(SPAM)

### 2.1 용어정의

<그림 1>은 스위치를 기반으로 하여 임의로 구성된 비정규적 네트워크를 나타내는 신장 트리이다[6]. 원형으로 표현된 스위치 노드는 연결 가능한 노드의 수에 제한을 하지 않음으로써 트리의 내부에 위치하게 되며 사각형으로 나타나는 호스트 노드는 단 하나의 스위치와 연결되도록 제한하기 때문에 트리의 리프에 위치한다. 노드에 부여된 숫자는 네트워크에서 유일한 ID를 나타내며 트리 설정 후에 임의로 부여된 것을 가정한다.

그림에서 실선으로 나타나는 직접 트리를 구성하는 데 사용된 채널들을 트리 채널(tree channel)이라 하고, 트리를 구성하는 데 사용되지는 않았지만 트리 구성 이전의 노드간의 기존 연결을 유지하는 채널들을 크로스 채널(cross channel)이라 하며 점선으로 표시되어 있다. 또한 리프에서 루트 방향으로의 단방향 채널들의 집합을 업 서브네트워크(up subnetwork)라 하고 그 반대 방향의 단방향 채널들의 집합을 다운 서브네트워크(down subnetwork)라고 한다. 트리의 레벨이 같은 경우에는 ID가 더 작은 노드로의 방향이 업 방향이 된다. 그러므로 다음과 같은 4가지 채널의 부류(class)가 존재한다.

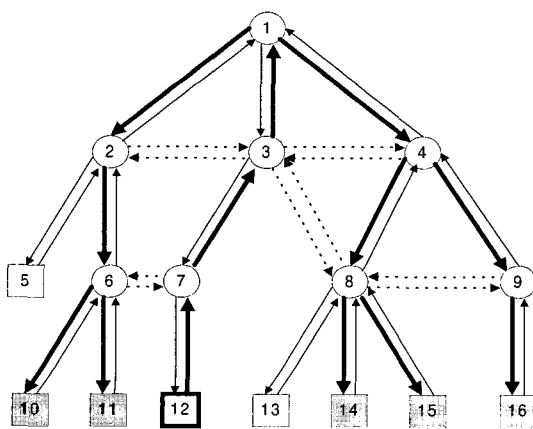


그림 1 구성된 전송 트리에서 SPAM 다중전송

업 서브네트워크에 속하는 트리 채널을 업 트리 채널(up tree channel)이라 하고, 크로스 채널을 업 크로스 채널(up cross channel)이라 한다. 마찬가지로 다운 서브네트워크에 속하는 트리 채널을 다운 트리 채널(down tree channel)이라 하고, 크로스 채널을 다운 크로스 채널(down cross channel)이라 한다.

### 2.2 SPAM 다중 전송과 문제점

단일 단계 적응적 다중 전송(Single Phase Adaptive Multicast; SPAM) 기법의 기본 라우팅 전략은 다음과 같다. 우선 교착 상태를 방지하기 위해서, 즉 전송 경로에 있어서 사이클의 생성을 피하기 위해 반드시 미리 지정된 순서로 채널의 부류를 사용해서 라우팅해야 한다. 각 스위치 노드에서 라우팅에 관한 결정은 트리 구성 시에 생성되어 목적지들로의 도달 가능성에 관한 정보를 제공하는 라우팅 테이블을 참조함으로써 이루어진다. 다운 방향으로의 도달가능성만을 고려하며 현재 노드에서 목적지가 도달 불가능할 때에는 메시지를 적응적인 특성을 반영하여 임의의 업 채널로 전달함으로써 루트 방향으로 계속 진행시킨다. 본 SPAM 기법은 전송 트리의 구성에 이용되어 노드들 사이의 결정적인 경로를 제공하는 트리 채널들 이외에도 트리에는 포함되지 않지만 스위치들 사이에 기존에 존재하던 연결을 유지하는 크로스 채널들을 적극적으로 활용함으로써 더 나은 적응성과 전송 단계의 단축을 이룩할 수 있다.

SPAM 다중 전송에서는 우선 모든 목적지 노드에 다운 트리 채널들만으로 도달할 수 있는 최저 레벨의 노드인 최소 공통 조상 노드(Least Common Ancestor; LCA)를 계산한다. 전송 시작 노드에서 보내진 메시지는 LCA에 도달하기 전까지 도달가능성을 고려하여 1개 이상의 업 서브네트워크의 채널, 0개 이상의 다운 크로스 채널, 1개 이상의 다운 트리 채널의 순서로 채널 부류를 이용한다. 원형 전송 경로의 생성을 배제함으로써 교착 상태의 발생을 피하기 위하여 반드시 이 채널 부류 사용 순서를 준수해야 한다. LCA에 도달한 메시지는 목적지에 도달 가능한 다운 트리 채널들로의 분열을 발생시켜 전송이 완료된다.

<그림 1>의 예에서는 루트 노드가 LCA에 해당하므로 반드시 루트를 경유해서 전송이 이루어지는데 네트워크의 크기가 커지고 목적지들이 전송 시작 노드의 하위 서브네트리로 희소하게 분포할수록 루트가 LCA가 될 가능성이 증가한다. 그러므로 이와 같은 전송들이 동시에 발생하는 상황을 가정하면 루트에 한순간에 과도한 부하가 집중되는 열점 현상(hot-spot effect)이 발생 가능하므로 문제점으로 지적된다. 뿐만 아니라 LCA를 계산하는 오버헤드도 문제점으로 지적된다.

### 3. 단일단계 차별적 다중전송 기법(SPDM)

#### 3.1 기본 전략

본 연구에서 제안하는 기법에서는 다중 전송시 메시지가 진행되는 도중에 부분적인 도달가능성을 고려하여 이를 발견하면 분열을 발생시켜 이들만을 목적으로 하는 새로운 메시지를 일단 보내주고, 원래의 메시지는 이들을 제외한 주소를 가지고 계속 진행하는 방식을 취한다. 다운 방향의 도달가능성만을 고려하는 SPAM 기법과의 가장 큰 차이점은 업 크로스 방향의 도달가능성도 특별히 고려함으로써 더욱 계획적인 메시지의 분열이 가능해진다는 점이다. 새로운 도달가능성을 고려했으므로 교착 상태를 피하기 위한 채널 부류 사용 순서의 제한을 확장하여 업 트리, 업 크로스, 다운 크로스, 다운 트리 채널의 순서로 사용하도록 정한다. SPAM의 다중 전송에서는 루트가 LCA로 계산될 때는 반드시 루트를 거쳐야 했지만 개선된 기법에서는 원래의 주소에서 미리 도달가능한 목적지들을 제외시키면서 진행하므로 루트 노드에 도달하기 이전에 모든 목적지들로 도달이 가능한 스위치 노드를 발견할 가능성이 커진다. 그래서 루트에 도달할 가능성은 감소하여 일반적인 트래픽 상황에서 LCA를 계산하지 않고도 루트에 부담을 주지 않는 다중 전송을 수행할 수 있다.

그러나 이와 같이 개선된 방식으로 다중 전송을 수행할 경우에 모든 흐름에 크로스 채널의 이용을 허용하면 루트에는 부리를 주지 않지만 트리 채널과 트리 채널로 연결된 상위 노드들의 이용률은 상대적으로 낮아지고 네트워크의 부하가 루트 이하의 내부 노드들, 특히 많은 호스트가 연결되어 있거나 많은 크로스 채널을 유지하고 있는 내부 스위치 노드들에 집중되는 경향이 생긴다. 그러므로 본 연구에서는 서비스의 질(Quality of Service; QoS) 개념을 적용하여 특정 호스트들의 메시지에만 업 또는 다운 크로스 채널의 이용 권한을 줌으로써 더 나은 적응성(adaptability)과 가능한 최소의 전송 단계(hop count)로 전송이 발생할 수 있도록 하였다. 즉, 네트워크 호스트 노드를 세 가지 중 하나의 서비스에 할당하고 서비스의 종류에 따라서 차별된 크로스 채널의 이용 권한을 갖도록 한다. 크로스 채널의 사용에 대한 제한은 크로스 채널을 통한 도달가능성이 숨겨짐을 의미하므로 고의로 트리 채널만을 이용하여 루트 노드를 경유하는 경로를 거치는 전송을 발생시킨다. 그러므로 할당 비율을 적절히 조절함으로써 채널 이용 부하 조절과 루트를 비롯한 스위치 노드를 통과하는 메시지 흐름의 제어가 가능해진다. 결론적으로 SPAM 기법에 비하여 업 서브네트워크에서의 모든 메시지에 대한 적응성의 측면은 약화

되지만 메시지 흐름에 대한 차별성을 부여받으므로 제안하는 기법을 '단일 단계 차별적 다중 전송'(Single Phase Differential Multicast; SPDM)이라 하였다.

#### 3.2 서비스의 종류

**Full-Adaptability(FA) 서비스 (10-20%) :** 트리 채널과 모든 크로스 채널을 이용 가능하므로 가장 나은 적응성과 최소의 전송 단계를 거친다.

**Half-Adaptability(HA) 서비스 (30-40%) :** 트리 채널과 다운 크로스 채널만을 사용 가능하다.

**Best-Effort(BE) 서비스 (40-60%) :** 트리 채널들만을 이용할 수 있으므로 결정적인 경로로 전송이 진행된다.

서비스의 할당 및 관리는 루트 노드에서 담당하고 서비스의 초기 할당은 트리 구성 과정중에 이루어진다. 루트 노드는 네트워크 상황과 호스트들의 요청에 의해서 유연성있게 할당률을 조정하는 것이 가능하며 SPAM과 같이 적응성을 중요시하고자 할 경우에는 모든 노드를 FA 서비스에 할당하면 되므로 SPDM의 특성은 차별성만에 국한되지 않는다.

#### 3.3 헤더 인코딩 및 라우팅 테이블의 구조

SPDM 기법은 NOWs 환경에 적합한[4] VCT 스위칭 방식[7]을 전제로 하고 있으므로 전송의 단위인 패킷의 헤더 인코딩은 <그림 2>와 같다. 지난 채널(last channel) 필드는 채널 부류의 지정된 사용 순서를 역행하지 않기 위해 패킷의 헤더가 마지막으로 거친 채널의 부류를 나타내며 4가지이므로 2-bit를 사용하여 <그림 3>과 같이 대응시킨다. 다음으로 해당 패킷이 어떤 서비스를 통하여 전송될 수 있는지 나타내는 'F'와 'H'의 퀄리티 비트가 존재한다. 주소는 여러 목적지들을 간편하게 나타낼 수 있는 비트 스트링 인코딩(bit string encoding) 방식[8]을 사용하는데 네트워크에 N개의 노드가 존재할 때 N-bit를 사용하여 각 자리 비트의 켜짐(1)과 꺼짐(0)으로 다중 노드들을 나타낼 수 있다. 비정규적 위상의 네트워크에서 노드의 수는 가변적일 수 있기 때문에 노드의 수가 증감할 경우를 대비하여 주소의

Last Channel	F	H	Address Length	Destination Address	Source Address	Data
--------------	---	---	----------------	---------------------	----------------	------

그림 2 패킷 헤더 인코딩

11	업 트리 채널
10	업 크로스 채널
01	다운 크로스 채널
00	다운 트리 채널

그림 3 채널 부류에 대응되는 값

channel number	channel class (2bit)
1	
2	
3	
4	reachability bit string
5	(N bit)
6	
7	
8	

그림 4 라우팅 테이블

길이(address length)를 지정하는 필드를 따로 두어서 이를 처리하도록 한다. 그러나 전송 시작 노드의 주소(source address)는 일반적인 이진수로 나타낸다.

<그림 4>는 각 스위치 노드에서 유지하고 있는 라우팅 테이블의 구조로서 스위치에 존재하는 채널 개수만큼의 엔트리를 가진다. 각 엔트리 앞의 숫자는 채널의 ID를 나타내며 채널 부류(channel class) 필드는 각 채널의 사용 여부와 현재의 트리를 구성하는 데 어떤 부류로 사용되었는지를 나타내며 양방향 채널에서 출력 채널을 기준으로 한다. 각 노드의 임의의 채널은 다른 스위치 노드 또는 호스트 노드와 연결되거나 아니면 사용되지 않은 상태로 비어있게 되며 사용중일 경우에는 마찬가지로 <그림 3>의 대응되는 값이 저장된다. 사용중인 채널의 각 엔트리에는 해당 채널을 통한 도달가능성을 마찬가지로의 인코딩 방식을 사용하여 저장한다. 단 업 트리 채널의 경우에는 도달가능성은 고려하지 않으므로 모두 '0'으로 채운다.

3.4 다중 전송 예

이제 제안하는 SPDM 기법을 <그림 5>의 호스트 노드 12에서 노드 10, 11, 14, 15, 16으로의 다중 전송의 예를 들어서 다시 설명하고자 한다. 우선 차별성을 고려하지 않고 SPAM과 마찬가지로 적응성만을 중요시하여 모든 노드들을 FA에 할당한 FA:HA:BE = 100:0:0 인 경우를 고려한다. 비트 스트링 인코딩에 의해 노드 12에서 구성되는 패킷 헤더의 목적지 주소는 '1110 0110 0000 0000'과 같다. 전송 시작 노드인 호스트 노드 12를 출발하여 스위치 노드 7에 헤더가 도착하면 이용 가능한 업 크로스 채널의 도달 가능성을 참조함으로써 목적지 중의 일부인 노드 10, 11로 일단 도달할 수 있음을 알게 된다. 그러므로 패킷의 분열이 발생하여 목적지 주소 '0000 0110 0000 0000'을 갖는 새로운 패킷이 업 크로스 채널을 통해서 다음 스위치 노드인 노드 6으로 보내진다. 한편 기존 패킷의 목적지 주소는 노드 10, 11이 리셋된

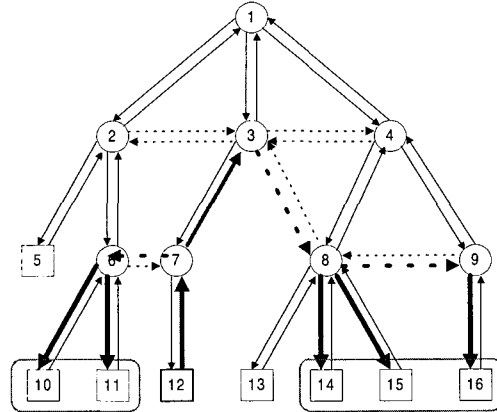


그림 5 SPDM 다중 전송 (FA:HA:BE=100:0:0)

'1110 0000 0000 0000'으로 조정되어 업 트리 채널을 통하여 스위치 노드 3으로 전달된다. 노드 6에 도착한 패킷은 다운 트리 채널들을 통하여 전송이 일단 완료되고 스위치 노드 3에 도착한 패킷의 헤더는 <그림 5>와 같이 최종적으로 전송이 완료된다. SPDM 기법에서는 부분적으로 미리 도달할 수 있음에 대한 정보를 최대한 이용하여 LCA를 미리 계산하거나 루트를 거치지 않더라도 최적의 경로를 이용해서 다중 전송을 수행한다.

다음으로 동일한 다중 전송을 일반적인 상황에서의 서비스 할당 비율 FA:HA:BE = 15:35:50인 경우에서 고려해 본다. 전송 시작 노드인 노드 12는 HA 서비스에 할당되었다고 가정하면 업 크로스 채널의 사용이 제한되는 것이 <그림 5>의 전송과의 가장 큰 차이점이다. 그러므로 노드 7에서 6으로의 도달가능성을 발견하지 못하게 되므로 업 트리 채널을 통해서 루트 방향으로 계속

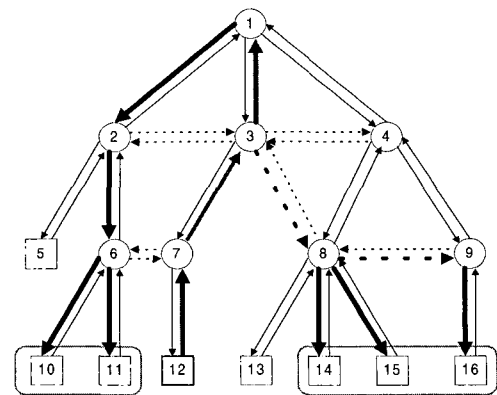


그림 6 SPDM 다중전송 (FA:HA:BE=15:35:50)

메시지가 진행하여 노드 14, 15, 16으로의 도달가능성을 먼저 발견하여 분열시키고 나머지는 루트를 점유하여 <그림 6>과 같이 전송이 완료된다.

이전의 전송에서는 무조건적으로 크로스 채널의 이용이 가능했기 때문에 일반적인 상황에서 크로스 채널의 이용률이 업 트리 채널의 이용률에 비해 상당히 높아지게 된다. 그렇지만 적절히 서비스의 할당이 이루어진 SPDM 기법에는 크로스 채널 뿐만 아니라 업 트리 채널의 이용률이 향상되어 크로스 채널이 연결된 특정 스위치에 집중되는 부하를 트리 채널로 연결된 상위 레벨의 스위치로 분산시킬 수 있는 이점이 있다. 그러므로 결과적으로 서비스의 비율이 적절히 할당된다면 네트워크의 구성 상황에서 야기되는 크로스 채널이 많이 연결된 스위치 노드에 많은 패킷이 집중되는 불가피한 현상에 대한 어느 정도의 조정 및 제어가 가능해져 부하의 상위 레벨로의 분산을 이룩할 수 있다.

**4. 성능 평가**

기존의 단일 단계 적응적 다중 전송(SPAM) 기법과 제안하는 단일 단계 차별적 다중 전송(SPDM) 기법을 시뮬레이션을 통하여 성능을 비교 분석하였다. 우선 16, 32, 64 노드의 세 가지 크기의 네트워크 각각의 경우에 서비스 할당 비율 100:0:0의 SPDM 기법, 15:35:50의 SPDM 기법, SPAM 기법에 대하여 패킷 전송 지연 시간과 루트 노드에서의 패킷의 집적도를 측정하였다. 또한 64노드 네트워크에서 SPDM 할당 비율을 15:35:50, 20:40:40, 30:40:30의 세 경우로 조정해 가며 SPAM과 성능 평가 요소들을 비교함으로써 상위 레벨의 노드들의 부하의 분포의 변화를 관찰하였다. 시뮬레이션은 Pentium-III 450MHz PC 환경에서 도구로서 MFC 6.0을 사용하였다.

**4.1 성능 평가 요소 및 기본 가정**

첫 번째 성능 평가 요소는 **목적지 노드 수에 대한 전송 지연시간**이다. 목적지의 수가 2개 이상인 다중 전송의 경우에는 목적지에 가장 늦게 도달한 패킷의 최대 지연 시간을 계산한다. 두 번째는 **스위치 노드당 통과한 패킷의 비율**로 전송 시작 노드에서 발생된 전체 패킷 수에 대해 각 스위치를 통과한 패킷의 수를 백분율로 나타낸 것이다. 이것은 모든 작업을 수행한 후 네트워크 부하가 어떻게 분포되었는가를 알아보기 위한 것으로 패킷의 분열이 발생하기 때문에 임의의 스위치에 총 작업 수보다 많은 패킷이 통과할 수도 있다.

호스트 노드의 수는 전체 노드 수의 50%에 해당하며 크로스 채널은 전체 트리 채널의 수의 약 35% 선을 유지하도록 네트워크를 임의로 설정하였다. 목적지 노드의

수는 1부터 (최대 호스트 수-1)개까지 균일하게 발생시켰으며 주된 전송 비용은 통신단계의 수(hop count)이다. 단일 통신 단계 지연 시간은 임의의 노드 사이를 통과하는 지연 시간으로 1  $\mu$ s를 가정한다. 그리고 전송 시작 노드에서 패킷 생성과 관련된 임의의 전송을 준비하기 위한 통신 준비 시간은 100 ns를 가정한다. 또한 스위치 노드에서 패킷 헤더를 디코딩한 후 라우팅 테이블을 검색하여 라우팅 결정을 내리고 출력 채널을 선택하는 시간인 패킷 헤더 처리 시간은 40 ns를 가정한다.

**4.2 성능 평가 결과 분석**

[표 4.1]은 각 네트워크의 목적지 수에 대한 전송 지연 시간의 측정에서 SPAM 기법에 대한 SPDM 할당 비율 100:0:0과 15:35:50의 평균적인 성능 향상 정도를 백분율로 나타낸 것이며 괄호 안은 최대값을 나타낸다.

표 1 평균지연시간에 대한 성능향상정도

비율	16노드	32노드	64노드	평균
100:0:0	10%(19)	13%(20)	5%(19)	9%(19)
15:35:50	2%(9)	6%(11)	4%(9)	4%(10)

<그림 7>은 특징이 가장 두드러지게 나타나는 64노드 네트워크에서 스위치 노드당 통과한 패킷의 비율을 나타낸다. 그림에서 SPAM의 경우 전체 스위치들 중에서 루트 노드에 가장 많은 패킷이 통과하였는데 네트워크가 방대해졌으므로 루트 이하 각각의 서브트리에 목적지가 존재하면 루트는 LCA가 되는 것을 피할 수 없기 때문이다. 또한 어떠한 목적지에 도달하기 위해서는 그 목적지가 연결된 스위치 노드를 반드시 거쳐야 하므로 하위 레벨의 스위치로 갈수록 세 가지 경우에 거의 차이가 나지 않음을 확인할 수 있다. 그림에서 몇몇의 노드는 다른 노드들에 비해 상당히 많은 수의 패킷이 통과하는 경향이 있는데 많은 호스트가 연결된 스위치 노드나

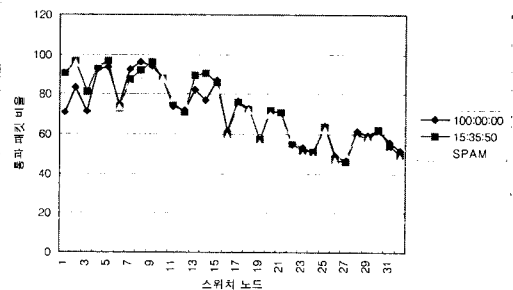


그림 7 64노드 네트워크에서 스위치 노드/통과 패킷 비율

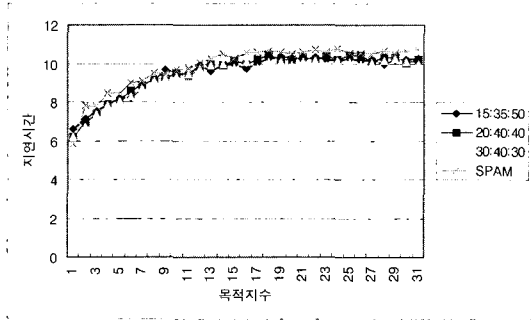


그림 8 64노드 네트워크에서 서비스 비율 변화/지연시간

트리 내부의 많은 크로스 채널이 연결된 노드일수록 현저해진다. 즉, 특정 호스트에 도달하기 위해서는 반드시 그 호스트가 연결된 스위치를 거쳐야 하고, 트리 채널 뿐만 아니라 크로스 채널을 통해서도 패킷이 전해지기 때문이다. 이것은 크로스 채널을 통한 현재 네트워크의 비정규적인 연결 상황에 의해서 야기되는 문제로서 정규적인 위상이 아닌 이상 어느 정도 불가피한 불균형이다. SPAM보다 SPDM 100:0:0의 경우에 더 심한 불균형이 일어나지만 SPDM 15:35:50에서는 조금 완화되었다.

<그림 8>은 64노드 네트워크에서 SPAM 기법과 할당 비율이 각각 15:35:50, 20:40:40, 30:40:30인 SPDM 기법의 지연 시간을 비교한 것이다. SPDM의 세 경우 모두 거의 차이가 없으며 SPAM보다 낮은 선을 유지하고 있음을 확인 가능하다. 또한 목적지 수가 어느 정도 이상이 되면 지연 시간이 거의 증가하지 않는데 이것은 하드웨어의 지원을 받는 다중 전송의 특징으로 지연 시간은 네트워크의 크기에 관계되며 목적지의 수의 영향은 적게 받는 성질을 확인할 수 있다[5].

<그림 9>는 64노드 네트워크에서 SPAM 기법과 할당 비율이 각각 15:35:50, 20:40:40, 30:40:30인 SPDM 기법의 통과 패킷 비율을 비교한 것이다. 역시 하위 레벨의 스위치에서는 거의 차이가 나지 않으며 할당 비율을 조정함에 따라서 네트워크 부하의 분포가 변화하고 있음을 나타내고 있다.

성능 평가를 통해서 얻은 결론은 다음과 같다. 첫째, 비정규적 위상에서 크로스 채널이 트리 채널의 35% 수준인 경우 개선된 SPDM 기법은 SPAM 보다 최대 20%, 평균 7% 정도의 더 나은 전송 지연 시간을 가지며 이것은 목적지의 수보다는 네트워크의 크기에 관련된다. 둘째, 개선된 기법은 SPAM에 비해 루트 노드에 주는 부담은 적지만 크로스 채널 또는 호스트 노드가 많이 연결된 노드와 그렇지 않은 노드간에 부하의 차이가 큰

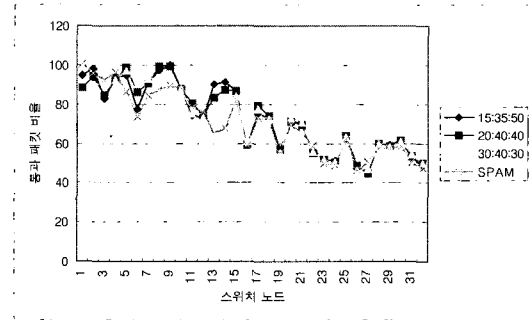


그림 9 64노드 네트워크에서 서비스 비율 변화/통과 패킷 비율

영향을 보인다. 셋째, 이러한 불균등은 네트워크의 비정규적인 위상에서 야기되는 불가피한 문제이지만 서비스 할당 비율을 조정함에 따라 어느 정도 조정과 개선이 가능하다.

### 5. 결론

본 논문에서는 워크스테이션들의 네트워크(NOWs)와 같은 스위치 기반 비정규적 위상의 네트워크에 적용할 수 있는 단일 단계 차별적 다중 전송(SPDM) 기법을 제안하였다. 이 기법은 트리를 기반으로 하며 다중 전송 패킷이 루트 방향으로 진행하는 도중에 부분적인 도달가능성을 발견하면 패킷의 분열을 발생시켜 이들로 일단 보내주고 기존의 패킷은 이들을 제외한 목적지를 가지고 계속 진행하는 방식을 취한다. 그래서 루트에 도달하기 이전에 모든 목적지로에 도달 가능성을 발견할 가능성이 커지므로 기존의 SPAM 기법의 가장 큰 문제점으로 지적된 루트 노드에서 과도한 부하의 집중을 해결하여 목적지의 수가 증가하더라도 루트에 부담 없는 전송을 수행한다. 비정규적 위상에서 크로스 채널이 트리 채널의 35% 수준인 경우 SPDM 기법은 SPAM 기법보다 최대 20%, 평균 7% 정도의 낮은 전송 지연 시간을 보이며 서비스의 질(QoS) 개념을 적용함으로써 호스트에 할당된 서비스에 따라 적응성과 전송 단계 수 등에 차별이 주어지는 전송의 형태가 가능하다. 또한 차별성 뿐만 아니라 서비스 할당의 조정에 의하여 SPAM과 같이 적응성을 강조할 수도 있다. 그리고 하위 레벨의 크로스 채널이 발달한 스위치 노드에의 상대적인 패킷 집중도를 트리 채널을 통한 루트 부근 상위 레벨의 노드들로 분산 시킴으로써 메시지 흐름과 네트워크 부하 분포의 제어가 가능하였다.

제안하는 SPDM 기법은 크로스 채널의 충분한 활용

을 전체로 하고 있으며 크로스 채널이 발달된 네트워크 일수록 더 잘 적용될 수 있고, 트리를 구성함으로써 일반적인 정규적 네트워크에도 응용이 가능하다. 비정규적 네트워크에 존재하는 크로스 채널의 비율을 다양하게 변화시킨 경우, 그리고 SPDM을 여러 정규적인 위상의 네트워크에 적용시킨 경우의 성능 평가도 교착 상태 방지 증명과 더불어 향후 연구로서 더 진행되어야 할 것으로 본다.

### 참 고 문 헌

- [ 1 ] T. Anderson, D.Culler, and D.Patterson. "A case for NOW(networks of workstations)", *IEEE Micro*, pp. 54-64, February 1995.
- [ 2 ] M. Schroeder et al. "Autonet: A high speed, self-configuring local area network using point-to-point links," Technical Report 59, DEC SRC, April 1990.
- [ 3 ] C. Clemenston, E. Pilatowicz, and K. Vong, "NOWs," <http://www.cse.ucsd.edu/users/hohlfeld/cse160/now/NOW.html>
- [ 4 ] J. Duato, A. Robles, F. Silla, and R. Bevide, "A comparison of router architectures for Virtual Cut-Through and wormhole switching in a NOW environment," In *Proceedings of the Merged 13th Int'l Parallel Processing Symposium and the 10th Symposium on Parallel and Distributed Processing*, pp. 240-247, April 1999.
- [ 5 ] R. Libeskind-Hadas, D. Mazzone, and R. Rajagopalan, "Tree-based multicasting in wormhole-routed irregular topologies," In *Proceedings of the Merged 12th Int'l Parallel Processing Symposium and the 9th Symposium on Parallel and Distributed Processing*, pp. 244-249, April 1998.
- [ 6 ] S. Varadarajan, and T. Chiueh, "Automatic fault detection and recovery in real time switched Ethernet networks," In *Proceedings of IEEE INFOCOMM '99*, pp. 161-169, March 1999.
- [ 7 ] J. Duato, et al., *Interconnection Network: An Engineering Approach*, IEEE Computer Society Press, 1997.
- [ 8 ] R. Sivaram, D. Panda, and C. Stunkel, "Multicasting in irregular networks with cut-through switches using tree based multidestination worms," In *Proceedings of the 2nd Parallel Computing, Routing, and Communication Workshop*, pp. 35-48, June 1997.



노 병 권

1998년 2월 서강대학교 전자계산학과 공학사. 2000년 2월 서강대학교 컴퓨터학과 공학석사



김 성 천

1975년 서울대학교 공과대학 공업교육학(전기전공)학사. 1979년 Wayne State Univ. 컴퓨터공학 공학석사. 1982년 Wayne State Univ. 컴퓨터공학 공학박사. 1982년 ~ 1984년 캘리포니아주립대 조교수. 1984년 ~ 1985년 금성반도체(주) 책임연구원. 1986년 ~ 1989년 서강대학교 공과대학 전자계산소 부소장. 1989년 ~ 1991년 서강대학교 공과대학 전자계산학과 학과장. 1985년 ~ 현재 서강대학교 공과대학 전자계산학과 교수(1992.9~현재). 1989년 ~ 현재 한국정보과학회 병렬처리시스템 연구회 부위원장(1989년 ~ 1993년), 위원장(1994년 ~ 1997년), 대한전자공학회 및 한국통신학회 논문지 편집위원(1991년 ~ 현재, 1993년 ~ 현재). 관심 분야는 병렬처리시스템(Parallel Computer Architecture, Interconnection Network)