

論文99-36C-11-5

PIMODS서버에서 분산 비디오스트림의 전송을 위한 상호연결망 (Interconnection Network for Routing Distributed Video Stream on Popularity-Independent Multimedia-on-Demand Server)

任綱彬*, 柳文幹*, 申俊浩*, 金庠中**, 崔景熙***,
鄭己鉉*

(Kang-Bin Yim, Moon-Gan Rew, Jun-Ho Shin, Sang-Jung Kim,
Kyung-Hee Choi, and Ki-Hyun Jung)

요 약

본 논문은 멀티미디어 서버의 부하 편중현상을 해결하기 위한 스위치를 제시하고 그 스위치에서의 트래픽 특성 분석을 위한 간단한 확률적 모델을 제안한다. 스위치는 경로설정 방안으로 우회방안을 이용하므로 스위치 내의 트래픽 부하는 우회확률에 커다란 영향을 미친다. 본 논문에서는 제안한 모델에 의거하여 스위치의 트래픽 부하에 따르는 우회확률을 추적하였다. 그리고 그 결과를 실험적 결과와 비교함으로써 확률적 모델의 타당성을 검증하였다. 확률적 모델에 의하여 스위치 안으로 유입되는 패킷의 양에 따라 발생하는 우회와 그에 따르는 스위치의 트래픽 포화지점을 예측할 수 있다.

Abstract

This paper presents an interconnection network for load balancing on a multimedia server and proposes a simple probabilistic model of the interconnection network for analysing the traffic characteristics. Because the switch uses deflection algorithm for routing, the traffic load on the switch seriously affects deflection probability. In this paper, we trace the deflection probability as a function of the traffic load according to the model. By comparing the result with the empirical result, we prove that the model is useful for estimating the deflection probability and traffic saturation point against the amount of packets getting into the switch.

* 正會員, 亞洲大學校 電子工學部

(School of Electrical and Electronics Engineering,
Ajou University)

** 正會員, 啓明文化大學 事務自動化科

(Dept. of Office Automation, Keimyung College)

*** 正會員, 亞洲大學校 情報 및 컴퓨터工學部

(School of Information and Computer Engineering,
Ajou University)

※ 본 논문은 한국학술진흥재단의 97자유공모지원사업
의 지원에 의한 결과임.

接受日字:1998年10月19日, 수정완료일:1999年10月6日

I. 서 론

멀티미디어 서버는 복수의 미디어 파일을 다중의 가입자를 상대하여 동시에 실시간으로 서비스하는 시스템이다. 미디어 파일은 주로 비디오 파일로서 그 특성상 방대한 용량과 고속의 전송률을 요구하며 실시간의 연속 서비스를 필요로 한다. 그러므로 비디오 파일을 위한 서버는 대용량의 저장 장치와 광대역의 고속 전송 구조를 가질 것이 요구된다. 더구나, 서비스 공급자의 목적이나 규모에 따라 가입자 및 파일의 수적 규모가 달라지므로 그에 대응하여 서버의 규모를 선택할 필요가 있다. 따라서, 멀티미디어 서버는 단위 증설이

가능하여야 한다.

기존의 멀티미디어 서버는 대개 범용 시스템을 이용하고 있으므로 단위 증설이 불가능하여 가입자의 수나 비디오 파일의 수가 증가하였을 경우에 대처하지 못하는 문제점을 가진다. 여타 시스템들도 마찬가지로 범용의 파일 서버를 부분적으로 개조하여 몇 개의 시스템을 통신망으로 연결한 정도이다. 그러나, 몇몇 시스템은 비디오 펌프를 장착하고 이들을 증설할 수 있도록 한 것도 있다^[7].

멀티미디어 서버의 성능을 좌우하는 주요한 요소 중의 하나는 저장장치의 입출력 대역폭이다. 기존 시스템은 거의 모든 구성요소가 단일의 버스에 연결되어 있는 구조이므로 전체적인 대역폭의 한계를 가진다. 따라서, 수많은 가입자가 선호도가 높은 비디오 파일을 동시에 요구하였을 경우의 병목현상 등을 해결할 만한 구조적인 대처 방안을 가지고 있지 못하다.

이러한 구조적인 모순에서 기인한 난점을 개선하기 위한 몇 가지 방안들이 제안되었다. 그 중에는 RAID를 이용하여 비디오 파일을 분할할당(striping)함으로써 파일 검색의 동시성에 의한 대역폭의 개선을 노린 경우나 선호도가 큰 비디오 파일의 복사본을 여러 개 두고 이들을 몇 개의 서비스 유닛에 나누어 배치하는 방안이 포함된다. 그러나, 이러한 방안도 단일버스에서 기인하는 병목현상이나 선호성에 의한 부하편중 현상을 만족스럽게 해결하지는 못하고 있다. 따라서 보다 새로운 구조의 멀티미디어 서버가 요구된다.

이러한 이유로 본 연구에서는 이미 멀티미디어 서비스에 적합한 구조의 단위 증설형 멀티미디어 서버 PIMODS(Popularity-independent multimedia-on-demand server)를 설계한 바 있다^[9]. PIMODS는 단위증설이 가능한 복수의 서비스 유닛으로 구성되며 비디오 파일의 선호도에 의한 부하편중 현상을 해결하기 위하여 유닛 간 분할할당 방안(Inter-unit striping)을 이용한다. 즉, 모든 비디오 파일들을 스케줄링을 위한 적당한 크기로 분해하여 이들을 각 서비스 유닛에 고루 분배한다. 그렇게 함으로써 선호성에 의한 저장매체의 병목현상이나 시스템 서비스 부하를 전체 서비스 유닛에 고루 분배하는 효과를 얻는다.

그러나, 본 서버는 ATM 망을 기반으로 서비스하기 때문에 이러한 파일 배치 방법은 한가지 중요한 문제점을 가진다. ATM 통신망에서는 일단 서비스가 개설

되면 해당 비디오 파일은 개설 당시의 채널을 통하여 전송되어야 한다. 그러므로, 각 서비스 유닛에 나누어 배치된 부분적인 비디오 파일이 서비스를 위하여 다시 조합되어야 한다. 따라서, 유닛 간 분할할당 방안을 이용하는 서버는 파일 조합을 위한 새로운 스위치를 요구한다.

비디오 파일의 조합을 위한 스위치는 다중의 스트림을 고속으로 처리하여야 하므로 Interconnection Network 형태의 병렬 경로를 지원하여야 하며 복수의 스트림을 효율적으로 처리하기 위한 경로설정 알고리즘을 요구한다. 본 논문에서는 이를 위하여 Meshed-pipe라는 스위치를 설계하고 그 경로설정 알고리즘으로 우회방안(deflection routing)을 이용하였다. 기존의 mesh 형태의 스위치에서는 store-and-forward나 worm-hole routing을 이용하였다. 이러한 경로 설정 알고리즘은 짧고 간헐적인 메시지 기반의 통신에서는 매우 유리한 방법이다. 그러나, 대량의 비디오 데이터를 주어진 시간 내에 고속이며 끊김 없이 전송하기 위한 스위치에서는 버퍼량의 과다 및 이로 인한 버퍼링 지연으로 실용성이 줄어든다.

본 스위치가 도입한 우회방안은 hot-potato routing이라고도 하며 임의의 노드에 도착한 패킷의 목적지를 위한 경로가 여유롭지 않을 경우 버퍼링을 하지 않고 주변의 다른 노드에 전송하는 방식이다. 본 논문에서는 meshed-pipe에서의 우회방안을 위한 전송 알고리즘과 이를 위한 단순한 확률적 모델을 제안하였다. 우회를 일으키는 패킷이 존재함에 따라서, 목적지를 향하는 최단경로를 이용하지 못하고 전송 지연을 유발하므로 이러한 우회확률을 모델링하면 스위치의 총체적인 전송 지연을 예측할 수 있다. 제안한 모델은 연구에서 정의한 실험적 모델에 의한 결과와 비교하여 서로 잘 근접함을 보임으로써 모델의 타당성을 검증하였다.

본 논문의 구성은 다음과 같다. 먼저 제 2장에서는 본 논문의 바탕이 된 멀티미디어 서버 PIMODS 및 스위치의 구조적 측면을 살펴 본다. 제 3장에서는 PIMODS에서 이용한 스위치의 경로설정 알고리즘을 설명하며 제 4장에서는 제 3장에서 정의된 알고리즘에 따라 발생하는 스위치 내의 트래픽을 위한 모델을 도출한다. 제 5장에서 제안한 모델과 실험적 모델을 비교함으로써 모델을 검증하고 마지막으로 제 6장에서는 스위치 및 모델의 유용성을 제시함으로써 결론을 맺는다.

II. PIMODS 서버

1. 단위 증설형 구조

멀티미디어 서버 PIMODS^[1]는 그림 1과 같이 복수개의 단위 증설형 서비스 유닛으로 구성된다. 각 서비스 유닛은 단위 서버로서의 PE와 가입자와의 연결을 위한 ATM I/F를 갖추고 있다. ATM I/F는 MPEG1 수준으로 압축된 영상 파일을 100여 가입자에게 전송할 수 있는 STM-1급의 통신 접속 장치이다. 이렇게 구성된 서비스 유닛은 뛰어난 처리 능력과 독자적인 저장 장치를 가지고 있어서 그 하나로서도 훌륭한 소 단위 서버로 동작할 수 있다.

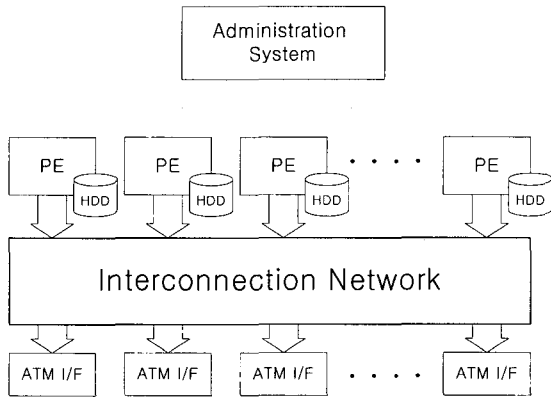


그림 1. PIMODS 서버 구조
Fig. 1. PIMODS structure.

PE의 주된 작업은 서비스 중인 가입자들의 스케줄링과 해당 가입자를 위한 비디오 파일을 저장장치로부터 읽어 ATM I/F로 전송하는 일이다. 멀티미디어 서비스 특성상 이들 두 작업은 모두 실시간으로 이루어져야 한다. 그러나, 저장장치로부터 비디오 파일을 읽으려면 프로세서의 점유가 발생하므로 가입자 스케줄링에 방해를 준다. 이러한 이유로 PE는 주프로세서와 부프로세서로 이루어져 있다. 주프로세서는 서비스 중인 비디오 파일의 유지, 관리를 담당하며 부프로세서는 저장장치로부터 비디오 파일을 읽어 ATM 셀을 만들고 이를 ATM I/F에 전송하는 일을 담당한다.

2. 유닛 간 분할할당(Inter-unit striping)

수백에서 수천의 영상과 음성이 혼합된 비디오 파일을 제공하는 멀티미디어 서비스는 기존의 문자 파일

서비스에 비하여 가입자의 선호도에 민감한 특성을 가지고 있다. 각 비디오 파일에 대한 선호도는 매우 불균일하여서 새로이 공개된 영화가 오래된 영화에 비하여 가입자의 흥미를 훨씬 더 자극한다. 한 연구에 의하면 비디오의 선호도는 다음과 같은 Zipf 분포함수로 나타난다.

$$P_{zipf}(n, M) = 1/n(1+1/2+1/3+\dots+1/M) \quad (2a)$$

Zipf 분포함수는 M개의 비디오 중에서 가입자가 n번째 선호도를 가지는 비디오를 선택할 확률을 나타낸다. M이 5000일 때 선호도 순위 50위까지의 Zipf 분포를 그림으로 나타내면 다음과 같다.

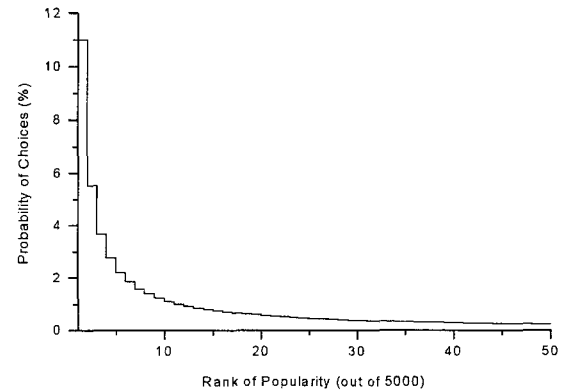


그림 2. Zipf 분포
Fig. 2. Zipf's distribution.

그림에서 보듯이 선호도 순위 3위까지의 비디오에 대한 선택 확률이 20%이상을 차지한다. 비디오 서버에서 볼 때 상위의 그림은 각 비디오 파일에 대한 상대적인 서비스 부하량을 나타내고 있다.

기존의 단위 증설형 서버에서는 비디오 파일을 배치함에 있어서, 하나의 영상 파일 전체를 하나의 단위 서비스 유닛에 할당하였다. 이러한 방법은 구현이 용이한 반면 영상 파일의 선호도 차이에 따른 서비스 부하의 편중현상이 나타난다. 즉, 5000개의 비디오 파일을 서비스하는 서버에서 하나의 서비스 유닛이 단지 가장 선호도가 높은 비디오 파일 세 개만을 서비스하여도 서버 전체 서비스 부하의 20%이상을 부담하게 된다. 이는 복수의 서비스 유닛에 대하여 소수의 유닛이 문제를 일으킬 확률이 높아지므로 서비스 유닛의 수명 단축은 물론 시스템 전체적인 안정성의 결여를 초래한다.

PIMODS는 상기의 문제를 해결하기 위한 새로운 방

안으로 유닛 간 분할 할당 기법을 이용하였다. 즉, 하나의 영상 파일을 서비스 시의 스케줄링 효율을 고려한 적절한 크기로 분할하고 이들 각각을 서비스 유닛에 균등하게 할당한다. 그렇게 함으로써 신호도가 높은 영상 파일이라 할 지라도 그에 대한 서비스 부하가 모든 서비스 유닛에 고루 분배되므로 각 서비스 유닛의 평균 서비스 부하는 동일하게 된다.

임의의 가입자가 유닛 간 분할 할당 기법에 의하여 저장된 비디오 파일을 요구할 경우 서비스의 개설은 파일의 시작부분이 담긴 서비스 유닛의 ATM I/F를 통하여 이루어진다. 이러한 상황에서 서비스가 유지되기 위해서는 모든 서비스 유닛이 자신에게 할당된 부분적인 비디오 파일을 서비스를 개설한 ATM I/F를 가진 서비스 유닛에게 순차적으로 제공하여야 한다. 즉, 모든 서비스 유닛은 일정한 시간 간격을 두고 특정의 ATM I/F에 순차적으로 연결 되어야 한다.

이러한 문제는 버스 방식을 이용하여 각 서비스 유닛으로부터의 부분 영상 정보를 다중화 할 수도 있으나 정상상태에서의 시스템에는 수많은 서비스 스트림이 움직이고 있으므로 이들을 동시에 처리하기 위해서는 복수의 서비스 유닛과 하나의 ATM I/F 사이를 동적으로 연결하기 위한 고속의 스위치가 필요하다.

3. 스위치 구조

PE와 ATM I/F 사이에 이들과 함께 단위 증설이 가능한 Interconnection Network으로서의 스위치가 위치한다. 스위치는 그림 3에서와 같이 단위 SE를 기본으로 한 meshed-pipe 구조이다. 그물망 내의 각 SE는 서로 UTOPIA 사양으로 연결되며 상, 좌, 우의 세 입력과 하, 좌, 우의 세 출력을 가진다. 또, 출력에는 패킷 전송에서의 동기화를 위한 하나의 패킷 버퍼를 가지고

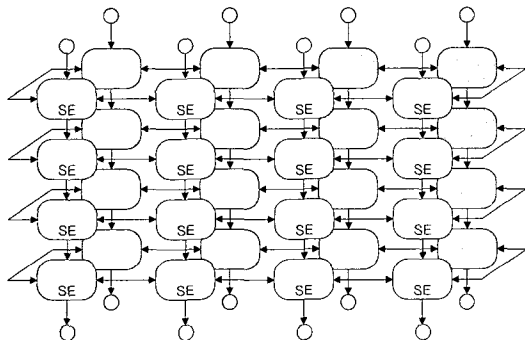


그림 3. 스위칭 엘리먼트의 연결
Fig. 3. Linkage of switching elements.

있다. SE는 크게 각 출력 방향에 대한 세개의 방향지시기와 하나의 중재기로 구성되며 제안하는 경로설정 방안을 위하여 고유의 구조를 가지는데 여기에는 신규 패킷 검출회로, 패킷 선호성 판단회로, 노드 어드레싱 회로 등이 포함된다.

III. 경로설정 방안

복수의 노드를 가지는 통신망에서 노드주소의 지정은 경로선택을 위하여 꼭 필요하다. 패킷이 통신망을 통하여 이동할 때, 패킷을 수신한 노드는 다른 노드의 도움 없이 패킷의 종착지까지 전송될 수 있도록 지원하는 방안을 가지고 있어야 한다.

그림 4에서의 스위치는 8x3의 노드로 구성된다. 각 노드는 자신의 고유 주소를 가진다. 고유 주소는 시스템의 초기화 시에 노드에 주어지며 운용 중에는 변하지 않고 고유의 값이 항상 유지된다. 이러한 고유 주소는 스위치 내에서 수평으로는 반시계 방향으로, 수직으로는 아래방향으로 하나씩 증가된 값이 주어진다. 이때 주어진 수평방향으로의 고유주소는 가입자의 서비스 요구를 받아들이는 ATM I/F의 주소로도 이용된다.

하나의 패킷이 스위치 내에서 전송될 때에는 노드 선택을 위하여 내부적으로 상대 주소가 이용된다. 패킷이 임의의 노드에 입력될 때마다 패킷이 가고자 하는 방향을 결정하기 위하여 노드의 고유주소와 패킷의 목적지 주소로부터 상대 주소가 계산되며 일단 패킷이 보내질 다음 노드의 주소가 결정되면 노드의 상대주소는 폐기된다.

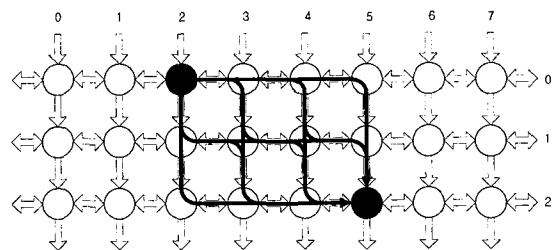


그림 4. 메시지 전송을 위한 다양한 경로
Fig. 4. Available paths for routing a message in the interconnection network.

그림 4에서 출발노드가 S, 종착노드가 D인 패킷을 고려하여 보자. 출발노드와 종착노드 간의 수평거리는 $D_x \in \{d_x; d_x = r_x(S, D)\}$ 로 표현할 수 있으며 마찬가지로

지로 수직거리는 $D_y \in \{d_y: d_y = r_y(S, D)\}$ 로 표현된다. 여기서 $r_x(S, D)$ 및 $r_y(S, D)$ 는 출발노드로부터 종착노드에 도착하기까지의 각 방향으로의 홑수를 구하기 위한 함수이다. 제안하는 스위치는 원통형을 이루고 있으므로 수평방향으로의 상대거리는 좌우 두개의 값이 존재한다. 함수 $r_x(S, D)$ 에 의하여 두 값 중에 하나가 선택되는데 조건 $d_x < M/2$ 에 의하여 보다 작은 값이 선택되며 이때 방향이 결정된다. 단, 조건에서의 M 은 원통형 그물망을 이루는 하나의 링에 속한 노드의 수이다. 구체적으로 함수 $r_x(S, D)$ 는 다음으로 정의된다.

```

dx1 = Xdest - Xnode ;
dx2 = Xnode - Xdest ;
IF dx1 < M/2 THEN
    Dx = dx1 ; dir = clockwise ;
ELSE IF dx1 > M/2 THEN
    Dx = dx2 ; dir = counterclockwise ;
ELSE
    Dx = dx1 ; dir = random ;
END IF
    
```

상기의 방식에 의하면 패킷이 망 내에서 가져야 할 필요 정보는 목적지 노드의 절대주소 뿐이다. 패킷이 수평방향으로 목적지 노드에 가까워지면 D_x 값은 하나 감소하고 멀어지면 D_x 값이 하나 증가한다. 마찬가지로 패킷이 아래로 내려가면 D_y 값은 하나 감소한다. 이렇게 하여 (D_x, D_y) 가 $(0, 0)$ 이 될 때까지 망 내를 진행한다. 즉, 패킷은 (D_x, D_y) 가 $(0, 0)$ 에 가까워지는 방향으로 진행하는 것이다.

패킷의 출발노드와 종착노드의 위치에 따라 패킷이 출력될 출구의 선택이 중요한 의미를 가진다. 만약 패킷이 종착노드가 위치한 방향과 반대방향으로는 진행하지 않는다고 가정하더라도 최단거리로 여행하는 경로는 여러 선택이 있을 수 있다. S에서 D로 향하는 패킷에 대하여 이러한 경로를 그림 4에서 회색의 화살표로 표시하였다. 이들은 수평으로 D_x 만큼, 수직으로 D_y 만큼 이동함으로써 종착노드에 닿을 수 있는, 지리적으로 모두 같은 거리를 가지는 경로이다.

정상상태의 망 내에는 무수히 많은 스트림들이 이동한다. 따라서, 수많은 패킷이 망을 흐르고 있으므로 새로운 패킷을 전송하는 중에 그 노드에서의 각 링크들

은 이미 사용되고 있는 경우가 자주 발생한다. 이는 복수의 패킷이 동일한 링크를 두고 경합을 벌이는 결과를 초래한다. 이를 해결하기 위한 방안으로서 패킷 대기를 위한 버퍼를 두고 버퍼에 대기 중인 패킷들을 적절히 스케줄링 하는 방법과 사용 중인 링크 이외의 링크로 강제 우회시키는 방법이 있을 수 있다. store-and-forward나 worm-hole 등과 같은 경로 이동 방식은 각 노드에 전송 데이터의 저장을 위한 적절한 버퍼를 두고 경쟁 패킷을 스케줄링하여 문제를 해결한다. 버퍼에 의존하는 방식은 짧고 간헐적인 메시지 전달이 요구되는 시스템에서는 효율적이나 전송률이 높고 용량이 큰 주기적 정보 전송을 다루기 위한 망에서는 버퍼링을 위한 엄청난 메모리를 요구하며 그에 따르는 전송지연을 유발하므로 버퍼의 유용성이 매우 저하한다.

버퍼링을 위한 과대한 메모리 사용을 방지하기 위하여 우회방안을 생각해 볼 수 있다. 우회방안에서는 수신된 패킷이 원하는 링크가 이미 다른 패킷에 의하여 사용되고 있을 경우 여타의 여유 있는 링크를 할당하는 방법이다. 우회방안은 경우에 따라서 트래픽을 가중시키며 전송지연을 유발하지만 고속의 망에서 좋은 성능을 발휘한다^[6].

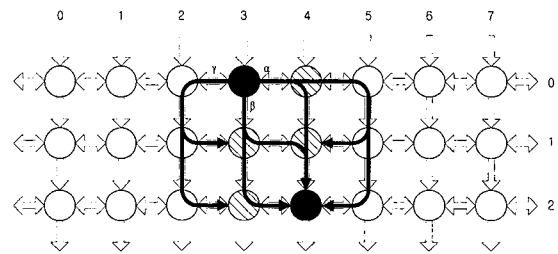


그림 5. 몇가지 우회의 예
Fig. 5. Deflected routing.

그림 5는 여러가지 우회상황을 보여준다. 노드 (3,0)에서, 오른쪽 수평링크를 사용하기를 원하는 패킷이 입력되었으나 이미 해당 링크가 다른 패킷에 의하여 α 의 확률로 사용되고 있다. 따라서, 패킷은 수평의 링크를 선택하게 되며 수평링크도 마찬가지로 β 의 확률로 다른 패킷에 의하여 이미 사용되고 있다. 그러므로 패킷은 확률 γ 로 노드 (2,0)으로 우회할 수 밖에 없다. 한번의 우회는 패킷으로 하여금 두개의 노드를 더 거치게 하므로 확률 γ 은 스위치를 통과하는 패킷들의 전송지연을 예측하는 데에 중요한 정보가 된다.

IV. 트래픽 모델링

본 절에서는 정상상태에서의 스위치 내 패킷 지연을 예측하기 위한 트래픽 모델을 정의한다. 멀티미디어 서비스의 신속성과 연속성 보장의 측면에서 패킷 지연에 대한 고려는 매우 중요하다. 우회방안에서의 패킷 지연은 우회한 수에 영향을 받는다.

우회가 발생하지 않은 경우의 평균 패킷 지연은 스위치의 규모에만 의존한다. 그러나, 임의의 노드에서 한번의 우회가 발생하면 지연된 경로는 그 노드에서의 최단경로보다 두개의 링크를 더 지나게 된다. 그러므로 우회발생 확률을 예측하면 지연된 패킷의 양과 그 지연 정도를 알 수 있다. 반대로, 패킷 지연의 한계값을 정하면 허용 가능한 최대 우회확률이 정해진다.

모델을 위하여 몇 가지 가정이 필요하다. 첫째, 스위치 내에서의 어느 노드에서도 패킷 생성은 발생하지 않는 것으로 본다. 둘째, 이상적으로는 각 노드에서 소멸되는 패킷이 존재하지 않는 것으로 본다. 그러므로, 모든 노드는 이웃하는 노드로부터 패킷을 입력 받아 다른 주변 노드로 무조건 전송하는 것으로 본다. 셋째, 정상상태에서의 트래픽은 스위치 전체에 대하여 고르게 분포하며 따라서 세 입력링크의 평균 사용률은 모든 노드에 대하여 동일한 것으로 본다. 이러한 상황에서 우회 확률은 다음의 세 단계를 수행함으로써 구한다.

1. 임의의 출력링크가 패킷 전송에 사용될 확률.
2. 임의의 패킷이 동일한 링크를 요구하는 경쟁 패킷을 만날 확률.
3. 경쟁 패킷들의 충돌에 의한 진행 방향의 조정으로 발생하는 우회 확률.

1) 링크 사용률

임의의 입력링크가 패킷 전송에 사용될 확률은 패킷이 노드로 들어와서 소멸되지 않고 출력링크를 점유할 확률에 영향을 미친다. 가정에 의거하여 정상상태에서 임의의 입력링크의 사용률은 스위치 전체에 대한 평균 입력링크의 사용률로 표현할 수 있다. 또한, 각 입력링크로부터 수신되는 패킷의 출력링크는 무작위적이며 노드 내에서 소멸되는 패킷이 존재하지 않으므로 평균 출력링크의 사용률은 평균 입력링크의 사용률로 대신

할 수 있다.

임의의 시점에 임의의 노드에서 하나의 패킷에 의하여 입력링크 k 가 점유될 확률을 μ_k 라 하면 정상상태에서 노드 내 모든 출력링크의 평균 사용률 μ 는 다음과 같이 표현할 수 있다.

$$\mu = \sum_m^M \sum_n^N \sum_k^K u_k / MNK \quad (4a)$$

여기서 M 및 N 은 스위치의 수평 및 수직의 노드 수이며 K 는 하나의 노드가 가지는 출력링크의 수이다.

어떤 패킷이 다른 패킷을 만남으로써 가고자 하는 방향으로 향하지 못하고 여타의 링크를 선택하여 이동할 확률은 노드 내에 공존하는 패킷의 수 및 그 패킷들의 각 링크에 대한 사용 확률에 종속적이다. 정상상태에서 패킷이 자신이 원하는 출력링크를 차지하지 못하는 상황은 최소한 그 패킷 이외의 패킷이 해당 노드에 하나 이상이 입력되어 처리되고 있는 경우에 발생한다. 그러므로, 하나의 패킷이 경쟁 패킷을 만날 확률은 다음으로 표현된다.

$$\mu^3 + \mu^2(1 - \mu) \quad (4b)$$

여기서 μ^3 항은 주어진 시각에서 세 입력 링크에 패킷이 도착하는 경우, $\mu^2(1 - \mu)$ 항은 두 개의 패킷이 노드에 도착하는 경우의 확률을 나타낸다.

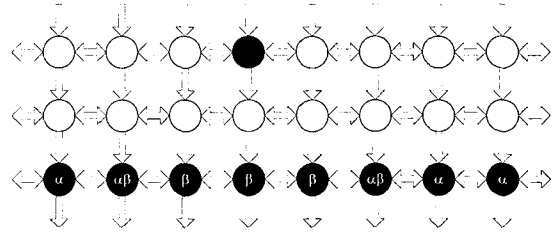


그림 6. 선호도 평가

Fig. 6. Preference evaluation.

본 연구에서는 패킷이 특정 링크를 사용하기를 원하는 성질을 패킷의 링크 선호성으로, 원하는 정도를 선호도로 정의한다. 패킷의 링크 선호도는 노드 내의 공존 패킷의 수와 무관하며 각 패킷의 현재 위치와 목적지 주소, 그리고 링크 선택 전략에만 의존한다. 패킷이 가질 수 있는 좌향 링크, 우향 링크, 하향 링크의 세 개의 출력 링크에 대하여 목적지 방향으로의 두 링크는 다른 하나에 대하여 선호성을 가지며 이들을 각각 수평링크와 수직링크로 명한다. 또 선호성을 갖지 않은

나머지 하나의 링크를 역방링크로 정의한다. 수평링크와 수직링크 사이에도 링크 선택 전략에 따라 서로 다른 정도의 선호도를 가질 수 있으므로 수평링크 및 수직링크의 선호도를 각각 α , β 로, 역방링크로 우회할 확률을 γ 로 정의한다. γ 는 노드 내 경쟁 패킷의 수와 각 패킷의 α 및 β 에 의존적이다.

2) 선호도 평가

스위치 내의 특정 위치에서 임의의 패킷이 가지는 링크에 대한 선호도는 링크 선택 전략에 따라 달라진다. 링크 선택 전략은 패킷이 자신의 목적노드에 도착하기 위하여 현재 위치에서 링크를 선택하기 위한 전략을 말한다. 이에 는 다음의 것들이 있을 수 있다.

- 수평 수직의 공평한 반복법(toggling method)
- 수평 수직의 거리 비에 의한 방법
(distance-measuring method)
- 통계적인 방법(statistical method)

첫째 방법은, 노드 내의 각 패킷을 위하여 수직링크와 수평링크에 반복적으로 선호성을 적용하는 방법으로서 노드의 부담을 최소화하면서 납득할 만 한 성능을 얻을 수 있다는 점에서 깊이 고려해 볼 만 한 방법이다. 수평과 수직의 링크가 공평하게 선택되므로 그 시점에서의 각 링크의 선호도 α , β 는 각각 1/2이 된다. 둘째 방법은, 목적지까지의 거리를 고려한 최적의 경로를 선택하기 위하여 목적지까지의 전체 홉 수에 대한 수평의 홉 수의 비율만큼 수평이동을 적용하는 방법으로서 매우 납득할 만 한 방법이나 확률 계산에 필요한 난수발생기의 요구 등 노드의 하드웨어 부담을 가중시킬 우려가 있다. 또, 셋째 방법은 노드의 각 출력 링크를 통과하는 패킷 수의 변화를 추적하여 주기적인 통계치를 정하고 이들 값을 이용하여 다음에 이용할 링크를 결정하는 방법으로서 노드 내의 모든 상황을 고려하여 링크를 선택하므로 가장 이상적이다. 그러나, 정보의 추출과 분석, 판단 등에 소요되는 노드의 부담이 가중되어 패킷의 전송률을 지킬 수 없는 상황이 발생하므로 현실성이 결여된다. 또한 이 때의 각 링크의 선호도는 통계적으로 표시되므로 미리 확정 지을 수 없다. 본 논문에서는 링크선택을 위하여 두 번째 방안을 고려하였다.

임의의 시점에 임의의 노드에서 임의의 패킷이 가지는 α 와 β 의 값은 해당 패킷이 해당노드로부터 목적노

드까지의 수평 및 수직좌표의 비율로 나타난다. 그러나, 스위치 내의 traffic이 모든 노드에서 균일하게 일어난고 있는 것으로 가정하면 특정의 α 와 β 를 위하여 평균의 α 와 β 를 사용할 수 있다. 평균 α 와 β 를 구하기 위하여 모든 노드에 대한 고려가 요구되며 짝을 지을 수 있는 두 노드간의 거리 비가 스위치의 총체적인 규모에 의하여 결정되므로 평균 α 와 β 는 M과 N의 값에 종속적이다.

$M > N$ 인 모든 스위치 규모에 대하여 평균 α 와 β 는 각각 망 내에서 패킷 선호성 평가의 총 횟수에 대한 평가의 결과가 α 그리고 β 가 되는 횟수의 비로 표현할 수 있다. 즉, 출발노드와 도착노드의 짝을 L(S,D)로 선호성 평가의 결과가 α , β 가 되는 경우를 각각 $\alpha(S,D)$, $\beta(S,D)$ 라고 표현하면

$$\alpha = \{ \text{total \# of } \alpha(S,D) \} / \{ \text{total \# of L(S,D)} \},$$

$$\beta = \{ \text{total \# of } \beta(S,D) \} / \{ \text{total \# of L(S,D)} \}$$

이다. 여기서 목적노드가 되는 D는 망 내의 최하위 N-1행에만 존재한다.

Total # of L(S,D)는 패킷의 목적지가 될 수 있는 최하위 행의 모든 노드들에 대하여 망 내의 모든 노드로부터의 출발노드와 목적노드 짝의 개수이므로 그 값은 M^2N 이 된다. M^2N 번의 선호성 평가에서 결과가 α (즉, $\alpha=1$, $\beta=0$)가 되는 L(S,D)는 $D_x > D_y$ 인 경우이며 β (즉, $\alpha=0$, $\beta=1$)가 되는 L(S,D)는 $D_x < D_y$ 인 경우이다. 또한 $D_x = D_y$ 인 경우에는 $\alpha=1/2$, $\beta=1/2$ 이다. 각각의 경우에 대하여 자세히 살펴보면 다음과 같다.

① $\alpha=1/2$, $\beta=1/2$

$D_x = D_y$ 인 경우이므로 임의의 하나의 출발노드가 정해지면 좌우 두개의 목적노드가 존재한다. 이러한 쌍은 최하위 N-1행을 제외한 모든 노드에 적용되므로 총체적인 값은 다음으로 표현된다.

$$\frac{1}{2} \sum_{n=0}^{N-2} 2M \text{ for } \alpha, \quad \frac{1}{2} \sum_{n=0}^{N-2} 2M \text{ for } \beta \quad (4c)$$

② $\alpha=1$, $\beta=0$

$D_x > D_y$ 인 경우로서 최하위의 N-1행에 존재하는 임의의 노드는 자신을 제외한 모든 노드로의 선호도가 α 이므로 M-1개의 노드가 해당된다. N-2 행에 존재하는 노드들에 대하여는 N-1 행에의 M-3개의 노드가 이

러한 경우에 해당된다. 그러므로 총체적인 값은 다음과 같다.

$$\sum_{n=0}^{N-1} (M-2n-1)M \text{ for } \alpha, \quad 0 \text{ for } \beta \quad (4d)$$

③ $\alpha=0, \beta=1$

$Dx < Dy$ 인 경우이며 최하위 $N-1$ 행에 존재하는 모든 노드에서의 목적노드는 가상적으로 하위에 존재한다. 그러므로 M 개의 노드가 이에 해당한다. $N-2$ 행에 존재하는 임의의 노드에서는 하나의 목적노드만이 이 경우에 포함된다. $N-3$ 행에서의 임의의 노드에게는 세 개의 목적노드가 포함된다. 그러므로 $\sum_{n=0}^{N-2} (2n+1)M$ 개의 노드가 더 포함된다. 따라서 총체적인 값은 다음과 같이 된다.

$$0 \text{ for } \alpha, \quad \sum_{n=0}^{N-2} (2n+1)M + M \text{ for } \beta \quad (4e)$$

그러므로 수평, 수직의 노드 수가 M, N 인 망에서의 평균 α 및 β 는 다음으로 주어진다.

$$\alpha = \frac{\sum_{n=0}^{N-1} (M-2n-1)M + \frac{1}{2} \sum_{n=0}^{N-2} 2M}{M^2N}$$

$$\beta = \frac{\sum_{n=0}^{N-2} (2n+1)M + \frac{1}{2} \sum_{n=0}^{N-2} 2M + M}{M^2N} \quad (4f)$$

3) 경쟁 패킷의 처리

특정 시점에서 하나의 링크에는 오로지 하나의 패킷만이 통과할 수 있으므로 선호도가 큰 링크로 동일한 링크를 선택하는 복수개의 패킷들은 하나의 링크를 두고 서로 충돌하게 된다. 패킷의 수에 따른 충돌 상황을 그림 7과 그림 8처럼 나타낼 수 있다. 그림에서 노드는 원으로 표시되었고 패킷은 흰색 및 회색의 삼각형으로 표시되었다. 여기서, 회색의 삼각형이 현재 고려하고 있는 패킷을 나타내고 있다.

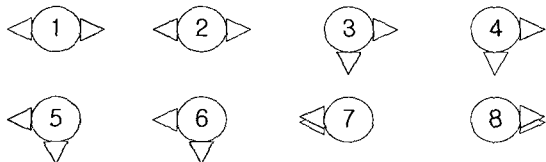


그림 7. 두 패킷이 경쟁하는 경우의 수
Fig. 7. Two competing packets.

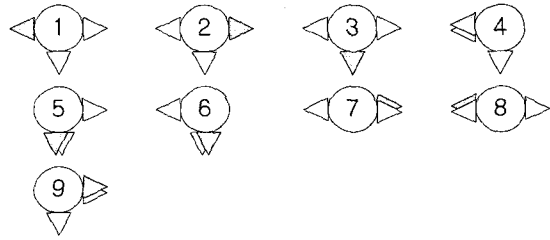


그림 8. 세 패킷이 경쟁하는 경우의 수
Fig. 8. Three competing packets.

(1) 두 패킷 간의 경쟁

두 패킷의 목적지 방향이 일치하는 경우와 그렇지 아니한 경우로 나누어 볼 수 있다. 1, 2, 3, 4, 5, 6번과 같이 최단경로의 방향이 서로 중복되지 않을 경우 충돌에 의한 링크의 경쟁이 발생하지 않는다. 그러나, 7, 8번과 같이 패킷들이 선호하는 링크가 중복되더라도 자신이 선호하는 수직링크가 쉬고 있으므로 우회는 발생하지 않는다.

(2) 세 패킷 간의 경쟁

이미 전송 중인 패킷이 두개일 경우 링크의 경쟁은 기존 패킷과 새로이 전송될 패킷의 선호도가 일치하는 경우에 발생한다. 그림8에서 4, 5, 6, 9번에서 우회가 발생할 수 있으며 나머지 경우에는 우회가 발생하지 않는다. 또한 4와9, 5와 6번은 서로 같은 경우이므로 그 중 4번과 5번 또는 6번과 9번만이 고려대상이 된다. 4번에서 실험패킷은 오른쪽 방향으로 전송될 수 밖에 없다. 그러므로 이러한 사건을 전제한 우회확률은 1이며 실제의 우회확률은 $\beta^2\beta\mu^3$ 로 표현할 수 있다. 5번에서는 실험패킷이 왼쪽으로 갈 수 밖에 없다. 이러한 사건을 전제하고 그 실험패킷에 대하여 왼쪽링크를 선택하는 것이 우회가 될 확률은 1/2이다. 그러므로 실제의 우회확률은 $\alpha\beta^2\mu^3/2$ 로 표현된다. 전체적인 우회확률은 이 두 값을 합한 것이므로 우회확률은

$$r = (\alpha^2\beta + \alpha\beta^2/2)\mu^3 \quad (4g)$$

이 된다.

V. 모의 실험 결과

모델의 적합성을 검증하기 위하여 실험적 모델을 구성하고 이 두 모델의 결과를 비교하였다. 실험적 모델로는 스위치 구조와 성능의 검증을 위하여 준비한 모

의 실험을 통하여 이루어졌다. 패킷을 연속적으로 스위치에 특정 패턴으로 포류시켜 패킷들이 종착노드까지 도달하는 동안 경험하는 우회확률을 분석함으로써 패킷의 지연상황을 유도해 낼 수 있다. 모델은 스위치 내의 패킷들의 경로이동 상황을 검출하기 위한 각종의 파라미터를 제공한다. 이러한 파라미터들은 노드를 위한 것과 패킷을 위한 것으로 나누어진다. 노드를 위한 파라미터에는 각 노드를 경유한 패킷의 수나 우회 발생 횟수, 소멸한 패킷의 수와 최상위 행 노드들의 입력단에서 생성된 패킷의 수, 최하위 행 노드들의 출력단에서 수신한 패킷의 수 등이 포함된다. 또 패킷을 위한 파라미터로는 패킷이 진행하면서 경험한 경유 노드의 수 등이 포함된다.

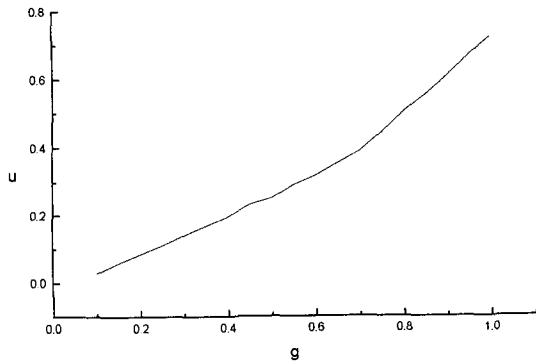


그림 9. 패킷 생성률 대 링크 사용률
Fig. 9. Link utilization to packet generation rate.

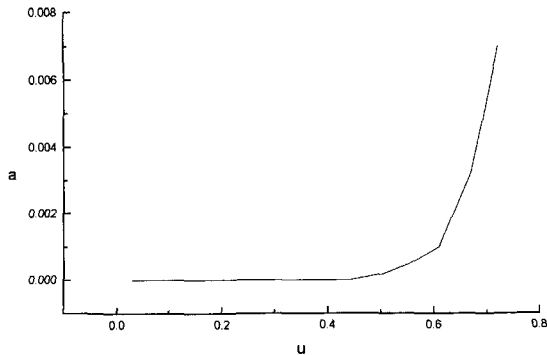


그림 10. 링크 사용률 대 패킷 유실률
Fig. 10. Packet loss ratio to link utilization.

실험적 모델에서의 링크의 사용률은 임의로 제어할 수가 없으므로 최상위 노드에서 PE로부터의 패킷 생성률을 점진적으로 증가시킴으로써 링크 사용률의 변화를 꾀하였다. 패킷 생성률의 변화에 따라 나타나는 링

크 사용률과 그에 따른 패킷 유실률을 그림9 및 그림10에 나타내었다. 그림에서 보는 바와 같이 패킷 생성률이 40%를 넘어서면서 링크 사용률의 기울기가 커짐을 볼 수 있다. 이 때의 링크 사용률은 40%이며 그림10에서 링크 사용률이 40%인 지점에서 패킷의 유실이 시작됨을 알 수 있다. 링크 사용률이 60%를 넘어서면서 패킷의 유실률은 급격히 커짐을 볼 수 있다. 이에 따라 스위치 내에서의 링크 사용률이 40%를 넘지 않도록 하기 위하여 패킷의 생성률을 70%로 제한할 것이 요구됨을 알 수 있다.

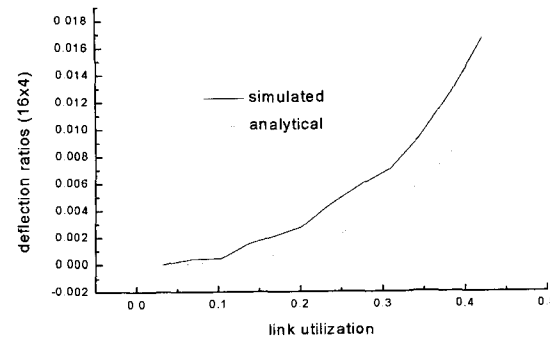
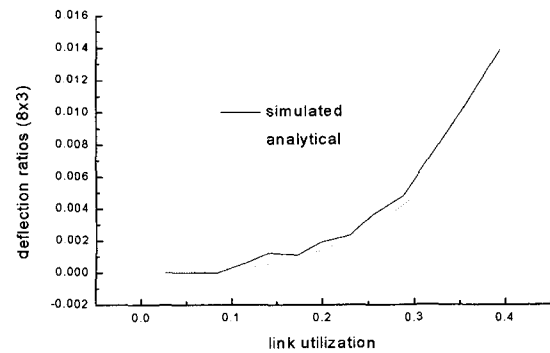


그림 11. 우회율
Fig. 11. Deflection ratio.

실험적으로 얻은 우회 확률과 수식 모델에 의한 우회 확률을 비교하기 위하여 링크의 사용률에 대한 각각의 우회확률을 8x3망과 16x4망에 대하여 각각 그래프로 나타내면 그림11과 같다. 그림에서 실험적인 결과는 실선으로, 모델에 의한 결과는 파선으로 표시하였다. 보는 바와 같이 수식 모델이 실험 모델에 잘 접근하고 있음을 볼 수 있다.

VI. 결 론

본 논문은 멀티미디어 서버 PIMODS에서 비디오 파

일의 선호도에 의한 부하편중 현상을 해결하기 위하여 요구되는 스위치를 소개하고 그 스위치 내에서의 트래픽 특성 분석을 위한 모델을 제안한다. 스위치는 그물망 형태이며 스위치의 경로 설정을 위하여는 우회 방안을 사용한다. 스위치 내의 트래픽이 무거울수록 우회의 발생확률이 높아지며 우회 발생확률이 높아짐에 따라 패킷의 전송지연이 증가한다.

본 논문에서는 스위치 내의 트래픽 부하에 의한 우회확률을 예측하기 위하여 단순한 확률적 모델을 정의하고, 확률적 모델에 의거하여 트래픽 부하에 따르는 우회확률을 추적하였으며 그 결과를 실험적 모델의 결과와 비교함으로써 확률적 모델이 실험적 결과와 근사함을 검증하였다. 그 결과, 멀티미디어 서버를 위한 스위치 내에서의 우회확률을 예측할 수 있게 됨으로써 이를 스위치 내의 트래픽을 제어하기 위한 척도로 이용하거나 전송지연의 측정에 매우 중요한 자료로 이용할 수가 있다.

참 고 문 헌

- [1] 신준호, "분산구조 주문형 멀티미디어 서버를 위한 단위처리기 설계," *Ajou University SG-MOD Technical Paper*, 1996
- [2] Kangbin Yim, Kihyun Jung, "Design Issues on A Popularity-Independent Video-On-Demand Server," *IEEE Transactions on Consumer Electronics*, Vol. 45, No. 2, May 1999.
- [3] Xiaola Lin, Philip K. McKinley, and Abdol-Hossein Esfahanian, "Adaptive Multicast Wormhole Routing in 2D Mesh Multicomputers," *Proceeding of PARLE '93*, Munich, Germany, June 1993.
- [4] Chengchang Huang, Eric P. Kasten, and Philip K. McKinley, "Design and Implementation of Multicast Operations for ATM-Based High Performance Computing," *Proceeding of Supercomputing '94*, Washington D.C., November 1994.
- [5] David F. Robinson, Philip K. McKinley, and Betty H. C. Cheng, "Optimal Multicast Communication in Wormhole-Routed Torus Networks," *Proceeding of International Conference on Parallel Processing '94*, St. Charles, Illinois, August 1994.
- [6] Dan Picker and Ronald D. Fellman, "Successive Superposition: A Technique for the Exact Modeling of Deterministic Packet Queuing Networks," *IEEE Transactions in Parallel and Distributed Systems*, Vol. 7, No 10, October 1996.
- [7] Albert G. Greenberg and Jonathan Goodman, "Sharp Approximate Models of Deflection Routing in Mesh Networks," *Ieee Transactions on Communications*, Vol. 41, No.1, January 1993.
- [8] Reza Rooholamini and Vladimir "ATM-based Multimedia Servers," *IEEE Multimedia Journal*, Spring 1995.

— 저 자 소 개 —

任 綱 彬(正會員)

1969년 4월 27일생. 1992년 아주대학교 전자공학과 학사. 1994년 아주대학교 전자공학과 석사. 1994년~현재 아주대학교 전자공학부 박사과정. 1997년~현재 아주대학교 공학연구소 전문연구요원. 1999년~현재 (미) 아리조나 대학교 연구원. 관심분야: 멀티미디어 컴퓨터 구조, 실시간 시스템, 운영체제 등

申 俊 浩(正會員)

1972년 8월 2일생. 1995년 순천대학교 전자공학과 학사. 1997년 아주대학교 전자공학과 석사. 1997년~현재 아주대학교 전자공학과 박사과정. 관심분야: 디지털 영상 처리, 통신 시스템, 운영체제 등

金 庠 中(正會員)

1949년 5월 15일생. 1977년 한양대학교 전자공학과 학사. 1980년 연세대학교 전자공학과 석사. 1997년 아주대학교 전자공학과 박사. 1977년~1998년 6월 한국전자통신연구소 뉴미디어시스템연구실장. 1998년 8월~현재 계명문화대학 사무자동화과 교수. 관심분야: 영상처리, 멀티미디어 데이터 처리, 멀티미디어 통신

柳 文 幹(正會員)

1971년 3월 20일생. 1990년 3월~1994년 2월 아주대학교 전자공학과 학사. 1997년 8월~현재 아주대학교 전자공학과 대학원. 관심분야: 컴퓨터 구조, 디지털 통신 등

崔 景 熙(正會員)

1953년 8월 7일생. 1982년 Paul Sabatier대(프) 공학박사. 1982년~현재 아주대학교 정보및컴퓨터공학부 교수. 관심분야: 운영체제, 분산 처리, 실시간 시스템 등

鄭 己 鉉(正會員)

1958년 10월 21일생. 1990년 (미) 퍼듀대학교 전자공학과 박사. 1990년~1992년 현대전자 수석연구원. 1992년~현재 아주대학교 전자공학부 교수. 관심분야: 멀티미디어 컴퓨터 구조, 실시간 시스템, 운영체제 등