

# 다중프로세서 시스템을 위한 상호결합 네트워크의 성능분석

## Performance Analysis of Interconnection Network for Multiprocessor Systems

金 元 燮\* · 吳 在 喆\*\*  
(Won-Sob Kim · Jae-Chul Oh)

### 요 약

VLSI기술의 발달로 고도의 병렬처리 시스템이 실현가능하게 되었다. 수 많은 프로세서와 메모리 모듈을 가진 다중프로세서 시스템은 복잡한 데이터 버스가 필요하게 된다. 다중프로세서들과 메모리 모듈 사이에 프로그램 가능한 데이터 경로를 제공하는 유용한 네트워크 방식이 다단상호 결합네트워크(MINS)이다. 본 논문에서는 Star네트워크에 MINS성능을 분석하여 기존의 Delta Network, Generalized Shuffle Network, Crossbar Network의 성능과 비교검토하였다.

**Abstract-**Advances in VLSI technology have made it possible to have a larger number of processing elements to be included in highly parallel processor system. A system with a large number of processing elements and memory requires a complex data path. Multistage Interconnection Networks(MINS) are useful in providing programmable data path between processing elements and memory modules in multiprocessor system. In this thesis, the performance of MINS for the star network has been analyzed and compared with other networks, such as generalized shuffle network, delta network, and referenced crossbar network.

### 1. 서 론

VLSI 기술의 급격한 발달로 다중프로세서 시스템의 이용이 크게 증가되고, 병렬처리 시스템이 실현 가능하게 되므로 다수의 프로세서와 메모리 모듈 사이를 연결해주는 복잡한 통신로가 필요하게 되었으며 고도의 병렬처리 시스템은 컴퓨터의 성능을 향상 시키기 위한 시스템으로서 상호결합 네트워크가 이 시스템의 연구에서 중요한 부분을 차지하고 있다. 이에 따라 단일프로세서 시스템에서는 프로

세서와 메모리 모듈 사이에 간단한 데이터 버스로서 충분했으나, 수많은 프로세서와 메모리 모듈을 가진 다중프로세서 시스템은 복잡한 데이터 버스가 필요하게 되었다.

다중프로세서 시스템에서 프로세서들과 메모리 모듈 사이에 프로그램 가능한 데이터 경로(path)를 제공하는 유용한 네트워크 방식이 다단상호결합 네트워크(Multistage Interconnection Networks:MINS)이다. MINS에서는 결합 특성에 대하여 많은 연구가<sup>1)2)3)</sup> 진행되었으며 2x2스위칭 소자가 개발되어서 상호결합 네트워크의 구성요소로 이용되었다. 그러므로 2x2스위칭 소자는 상호결합 네트워크의 구성 및 성능에 중대한 영향을 끼쳤고, MINS에 관한 연구는 스위칭 방식, 제어 방식, 통신 프로토콜에 따

\* 正 會 員 : 全 北 大 教 授 · 工 博  
\*\* 正 會 員 : 順 天 大 學 電 子 計 算 學 科 專 任 講 師 · 工 博  
接受日字: 1987年 7月 8日  
1次修正: 1988年 6月 11日

라 분류하면 Baseline,<sup>3)</sup> Data manipulator<sup>4)</sup> SW banyan,<sup>5)</sup> Cube omega<sup>6)</sup>, Flip<sup>7)</sup>, Inbirect binary<sup>8)</sup>, Delta<sup>9)</sup>, Gamma<sup>10)</sup>네트워크가 있다.

MINS는 대규모 병렬 처리 시스템에서 버스구조나 Crossbar 네트워크에 비하면 여러가지 장점이 있기때문에 MINS로 구성된 시스템의 성능분석에 관한 많은 연구<sup>11)~16)</sup>가 되었다. 따라서 네트워크 시스템의 중요한 특성을 나타내는 파라미터(parameter)를 변수로 네트워크의 성능을 해석하는데는 patel<sup>9)</sup>이 최초로 MINS성능을 분석하였다.

Star 네트워크<sup>17)</sup>는 회로스위칭(Circuit Switching) MINS을 이용한 다중프로세서 시스템이며 주어진 위상(Topology)으로 설계된 병렬처리 알고리즘과 부합시킬수있는 재구성력을 가진 시스템으로서 Star 네트워크 관한 연구<sup>17)~19)</sup>가 있었다. Star네트워크의 성능은 상호결합 네트워크와 호스트 프로세서(Host processor)의 특성에 따라 좌우된다. 앞으로 전개될 논문에서는 상호결합 네트워크의 성능을 분석하기 위해서 Baseline 네트워크를 이용하여 Star네트워크를 구성하며 Cluster가 형성되었을때 해석적인 모델로 Cluster mode의 시스템 성능을 해석한다. 그리고 patel의 이론을 이용하여 네트워크의 연결request가 발생하였을때 request의 받아들여질 기대치(Expected number of request acceptance) B W(Bandwidth)와 request가 받아들여질 확률  $P_A$ 를 구하여 기존의 Delta 네트워크, Generalized Shuffle네트워크, Crossbar 네트워크들의 성능과 비교 검토하기로 한다.

## 2. Star 네트워크

### 2.1 개념

Star 네트워크는 다단 상호결합 네트워크를 이용한 다중 프로세서 시스템이며 주어진 위상(Topology)에 따라 설계된 병렬처리 알고리즘과 여러가지 위상이 부합 될수있도록 재구성한 시스템이다. 이러한 Star네트워크는 다음과 같은 세가지 특성이 있다. 첫번째가 MINS의 성능 향상이다. 두번째는 Star네트워크의 재구성력이다. 세번째는 Star네트워크의 다용도성(Versatility)이다. 그림1에서와 같이 Star네트워크의 하드웨어 모델은 N개의 프로세서 노드(processor node)와 통신망으로 구성되어있다.

각 프로세서의 노드 인터페이스 장치를 통하여 b개의 Baseline 네트워크<sup>3)</sup>와 결합 되어있다. Baseline 네트워크 각각의 스위칭 소자의 상태를 적정하게 세

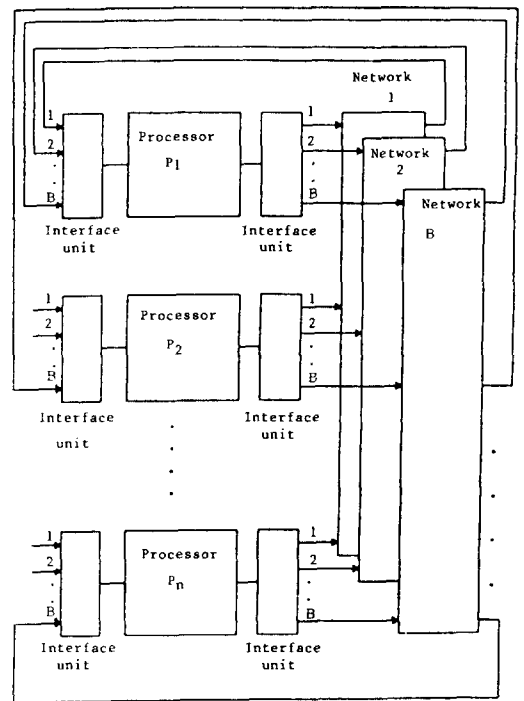


그림 1 Star 네트워크의 하드웨어 모델  
Fig. 1 Hardware model of star network.

트(set) 시키므로서 회로 경로는 프로세서간에 통신을 할수있도록 한다.

N개의 프로세서로 구성된 각 Baseline 네트워크는  $\log_2 N$  단(stage)으로 이루어졌고 각단은  $N/2$ 스위칭 소자로 구성되었다.

각 스위칭 소자는 두개의 입력단과 두개의 출력단을 가지고있다. 그러한 스위칭 소자를  $2 \times 2$ 스위칭 소자라하며  $4 \times 4$  혹은 그이상의 스위칭 소자는 이론적으로 이용되나 실제적으로 회로의 스위칭 소자에 복잡성 때문에 실현에 어려움이 많다.

### 2.2 Baseline 네트워크

Star네트워크는 한개 이상의 Baseline네트워크로 구성되어있다. Baseline네트워크는 많은 MINS중의 하나이다. 이것은 구조적으로 Omega, SW banyan, N-cube, Delta네트워크와 일치한다.<sup>3)</sup> 주어진 MINS의 한쪽 혹은 양쪽에 순열(permutation)을 적용시키면 하나의 네트워크가 다른 MINS로 변환된다.

16개의 입력 노드와 16개의 출력 노드를 가진 Baseline네트워크가 그림2와 같다.

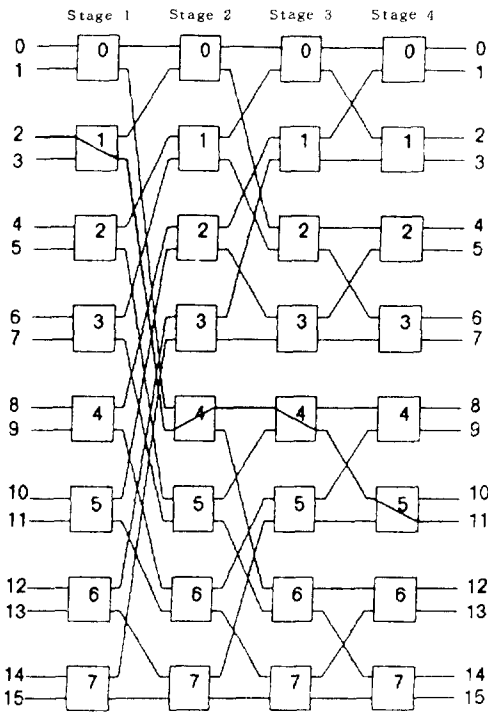


그림 2 16×16 Baseline 네트워크 모델  
Fig. 2 16×16 baseline network model.

Baseline네트워크는 다른 MINS와 비교할때 경로 제어(Routing)가 쉽다. 도착점 어드레스(Destination address)의 2진 부호 표기가 경로제어 태그(Routing tag)방법으로 사용된다. 경로제어 방법은 다음과 같다. 출발점 노드 어드레스(Source node address)를 C라 하고 도착점 어드레스를 D라 한다.

이 경우에  $S = s_{n-1}s_{n-2}\dots s_0$ 이고  $D = d_{n-1}d_{n-2}\dots d_0$ 이다. S와 연결된 첫번째 스위칭 소자는  $d_{n-1}=0$ 이면 상위링크(Upper link)에 연결되고  $d_{n-1}=1$ 이면 하위링크(lower link)에 연결된다. 두번째 단계에서  $d_{n-2}=0$ 이면 상위링크가 이용되고  $d_{n-2}=1$ 이면 하위링크가 이용된다. 이러한 방법으로  $d_0$ 까지 반복한다. 그림2에서 제시한 바와 같이  $S=2$ 이면  $D=11(d_3d_2d_1d_0=1011)$ 이다.

### 2.3 운영단계

Star네트워크는 회로교환(Circuit Switching)이론을 기초로한 시스템이다. 이것은 실제적인 데이터 전송전에 경로가 형성된다.

결과적으로 운영단계는 다섯가지로 구성되는데 1

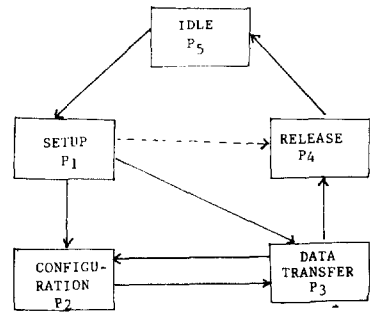


그림 3 StarNet의 운영단계  
Fig.3 Phase of the starnet operation.

단계는 Set up, 2단계는 path Configuration, 3단계는 Data transfer, 4단계는 Release, 5단계는 Idle이다. 각 단계는 네트워크의 상태, 프로세서 스위칭 소자의 상태를 나타내며 네트워크 자체에서는 각 단계를 지시해주는 하드웨어적인 레지스터가 없다. 이들 상태는 인터페이스 장치 언어 혹은 주어진 프로세서에서 수행되는 task에 따라 결정된다. 각단계의 관계를 나타내는 것이 그림3이다.

그림3에서 Set up단계( $P_1$ )는 프로세서가 Star네트워크에 연결request를 발생하여 경로가 형성되도록 하며 이단계는 경로가 형성되었거나 혹은 경로 형성이 포기되었을때 종료된다. path configuration 단계( $P_2$ )는 새로운 경로가 형성되었을때 데이터 전송방향과 경로의 Ownership을 결정한다.

Data transfer 단계( $P_3$ )는 경로 방향과 Ownership이 결정되면 프로세서간에 비동기로 데이터를 전송하게 된다. 또한 이 단계는 데이터 전송 방향을 변경시킬때는 다시 path configuration 단계로 수행된다.

Release단계( $P_4$ )는 데이터 전송이 종료되어 경로를 해제할 경우에 사용된다. Idle단계( $P_5$ )는 프로세서가 어떠한 task도 수행하지 않는 상태를 나타낸다. Cluster mode는  $P_1-P_3-P_4-P_5$ 의 운영 단계가 구성되며 회로교환 방법과 같으며 message mode는  $P_1-P_2-P_3-P_4-P_5$ 의 단계로 구성되며 패킷 교환 방법과 유사하며 store and forward시스템이다.

경로 상태의 시간은 Star네트워크의 총체적인 동작에 매우 중요한 역할을 한다. 경로가 프로세스 기간 전체에 계속되어 있으면 경로의 주요 목적은 메시지를 전송하기 보다는 대부분의 프로세서 상태를 Cluster 상태로 한다. 다중 경로가 형성되고 병렬처리 과제(task)동안 지속되면 경로의 위상은 수행되

고 있는 알고리즘의 위상과 부합시키기 위해서 만들어질수있다. 이러한 것을 Cluster mode라 한다. 반면에 경로는 짧은 시간에 이용된다. 이러한 경우에 경로는 메시지의 끝에서 해제(Release)되고, 링크도 역시 해제되어서 다음 동작에 이용된다. Cluster 개념이 동작방식에서는 약하다. 이러한 상태를 Message mode라 한다. 그림4는 Cluster mode와 Message mode를 나타낸다.

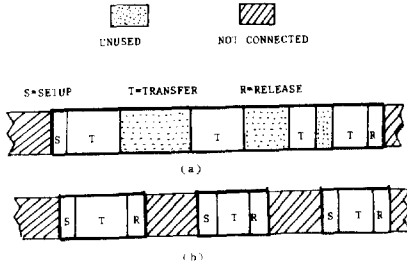


그림 4 Starnet의 운용 모드  
(a) Cluster 모드 (b) Message 모드

Fig.4 Modes of starnet operation.  
(a) Cluster mode (b) Message mode

### 3. Cluster mode의 해석모델

성능 평가는 시스템의 중요한 특징을 나타내는 파라미터의 결지에서, 컴퓨터 시스템의 가능한 능력을 측정하는 것이다. Star네트워크의 성능은 상호결합 네트워크의 성능과 호스트 프로세서의 능력에 좌우된다. Star네트워크는 특정한 프로세서로 설계된 것이 아니므로 본 논문에서는 상호결합 네트워크에 대하여 고찰한다.

운영 단계에서 결합 상태는 Cluster mode에서 주어진 임무를 수행하는 동안 반영구적이다. 경로는 항상 데이터 전송에만 사용되지는 않는다. 경로가 Set up되는 시간을 절약하기 위해서 수행중에 request의 패턴은 변화시키지 않는다. Cluster의 예와 등가 request결합이 그림5에 나타나 있다.

하나의 노드에 하나 이상의 request가 주어졌을때 Star네트워크에서는 네트워크의 성능을 향상시키기 위해서 한개 이상의 Baseline네트워크가 이용된다. 이중(Duplication)의 Baseline네트워크가 존재할때 그러한 Cluster가 실현 가능하다. 그러나 주어진 노드의 수에서 몇개의 네트워크가 최적인지는 알려지지 않았다. 해석하고자 하는 네트워크는 b개의 Ba-

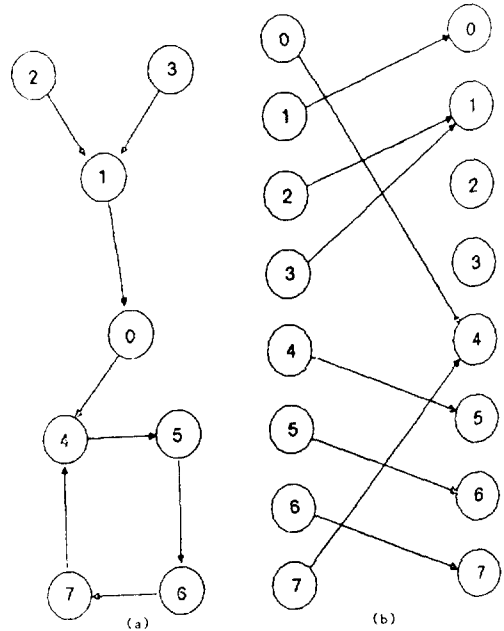


그림 5 샘플 Cluster와 등가 request형태  
(a) 프로그램 Cluster  
(b) 등가결합

Fig.5 Sample cluster and its equivalent request-pattern.

(a) program cluster  
(b) equivalent connection

aseline네트워크로 이루어졌다. 각 Baseline네트워크는 N개의 입력노드와 N개의 출력노드가 연결되어 있다. 성능을 비교하기 위해서는 부 네트워크(Sub-network)인  $N \times N$  Baseline 네트워크를 구성하였다. 네트워크의 효율성은 주어진 Cluster를 만들기 위해서 네트워크를 결합시키는 방법을 검토하므로써 측정할수 있다.

Clustering의 확률을  $P_c$ 로 정의한다.

해석 모델을 설명하기 위해서는 다음과 같은 세가지 가정을 기본 이론으로 정리한다.

가정(1): Task의 처음에서 각 프로세서는 무작위로 독립된(Random and Independant) Request를 확률 r로 네트워크 1에 발생시킨다. Request는 일정(uniform)하게 N출력(Receivers)에 분산된다.

가정(2): 그림1에서 다수의 네트워크(1, 2, ..., j, j+1, ...)설계가 가능하면 j번째 네트워크에서 저지된 request는 j+1번째 네트워크에 request가 다시 주어지게 될 것이다.

가정(3): 새로운 request는 j번째 네트워크에서 저지된 request와는 서로 다르다. 새로운 request의 도착점은 저지된 request와는 무관하게 형성된다.

동일 표준의 무작위 표본 가정(1)은 다른 분산 방법이 없기 때문에 이용되었다. 그래서 해석 모델을 간소화 시킨다. 가정(2)에서 이용된 방법보다 네트워크에 request를 받게하는 더 좋은 방법이 존재할 수도 있다. 그러나 이 방법이 아주 적합하다면 실제 시스템에 좋은 방법이다. 가정(3)은 같은 Cluster가 다른 request로 이루어질수 없기 때문에 비현실적이다. 그러나 이 가정은 모델을 설정하는데 쉽다.

#### 4. Baseline 네트워크 해석

하나 이상의 request가 동일한 도착점 노드에 주어졌을때 하나의 경로 결합이 이루어진다. 네트워크 성능을 향상 시키기 위해서 한개 이상의 Baseline 네트워크가 이용된다. 그래서 이러한 시스템의 성능을 이해하는 것이 중요하다. 다음에 주어진 가정이 Baseline model을 정의 하는데 이용된다.

$P\{A_{ij}\}$ : 입력단 i번 노드가 출력단 j번 노드에 request를 발생하여 그request가 받아들여질 확률.

$P\{S_{ij}\}$ : 입력단 i번 노드는 출력단 j번 노드에 request를 발생할 확률.

$P\{B_{ijk}\}$ : 입력단 i번 노드와 나머지 k-1개를 노드는 출력단 j번 노드에 request를 발생할 확률.

앞에서 언급한 가정에서 k개의 노드는 노드j에게 request를 발생시키면 k개의 request중 하나가 받아들여질 확률은  $b/k$ 이다.  $k < b$ 이면 모든 request는 받아들여진다.

따라서

$$P\{A_{ij}\} = \sum_{k=1}^{b-1} P\{B_{ijk}\} + \sum_{k=b}^N (b/k) P\{B_{ijk}\} \quad (1)$$

(b는 Baseline 네트워크의 수이다)

$b=1$ 이면 식(1)에 의하여 다음과 같다.

$$P\{A_{ij}\} = (1/k) P\{B_{ijk}\} \quad (2)$$

$b=1$  경우에는 patel<sup>9</sup>에 의해서 연구되었다.

$q(i)$ 를  $N \times N$  Baseline 네트워크에서 i개의 request가 발생할 확률이라 한다.

$$q(i) = \binom{N}{i} r^i (1-r)^{N-i} \quad (3)$$

(r은 하나의 프로세서가 한 사이클 동안 한개의 request가 발생할 확률)

$E(i)$ 는 i개의 request가 주어졌을때  $N \times N$  Baseline 네트워크가 받아들여질 request의 기대치라 한다. E

(i)를 계산하기 위해서 i개의 무작위 request가 N개의 도착점(Destination)에 주어질 방법은  $N^i$ 이다. 어느 특정한 도착점에 request가 도달하지 않았다고 가정하면 i개의 request가 나머지  $(N-1)$ 개의 노드에 주어질 방법은  $(N-1)^i$ 이다. 그래서  $N^i - (N-1)^i$ 는 어느 도착점 노드가 항상 request를 받을 횟수이다. 어떤 하나의 도착점에 request가 통과할 확률은  $[N^i - (N-1)^i] / N^i$ 이다.

주어진 i개의 request가 통과될 기대치는 식(4)와 같다.

$$E(i) = \frac{N^i - (N-1)^i}{N^i} N = \left[ 1 - \left( \frac{N-1}{N} \right)^i \right] N \quad (4)$$

Bandwidth(BW)는 상호결합 네트워크에서 정보소통 능력을 나타내는 것으로서 단위 시간당 request가 받아들여질 기대치인데  $N \times N$  Baseline 네트워크에서 i개의 request가 발생할 확률인데 i개 request가 받아들여질 기대치의 적(product)으로 표시된다.

$$BW = \sum_{i=0, i \leq N} E(i) q(i)$$

간단히 표현하면

$$BW = N - N(1-r/N)^N \quad (5)$$

단일 request가 받아들여질 확률은

$$P_A(b=1) = \frac{BW}{N \cdot r} = \frac{1}{r} \left[ 1 - \left( 1 - \frac{r}{N} \right)^N \right] \quad (6)$$

$A_{ij}$ 와  $S_{ij}$ 의 정의에 따라 request가 받아들여질 확률은 식(7)과 같다.

$$P_A = P\{A_{ij} : S_{ij}\} = P\{A_{ij}\} / P\{S_{ij}\} \quad (7)$$

$$P\{S_{ij}\} = r(1/N) = r/N \quad (8)$$

따라서

$$P\{A_{ij}\}_{b=1} = P\{S_{ij}\} P_{A,b=1} = (1/N) \left[ 1 - \left( 1 - \frac{r}{N} \right)^N \right] \quad (9)$$

$b=1$ 일때 받아들여질 확률은 식(1)과 식(9)로부터 얻을수있다.  $P\{B_{ij1}\}$ 은 하나의 request가 노드j에게 주어질 확률이고 식(10)과 같이 나타낼수있다.

$$\begin{aligned} P\{A_{ij}\}_{b=2} &= P\{B_{ij1}\} + \sum_{k=2}^N (2/k) P\{B_{ijk}\} \\ &= 2 \sum_{k=1}^N (1/k) P\{B_{ijk}\} - P\{B_{ij1}\} \\ &= 2P\{A_{ij}\}_{b=1} - P\{B_{ij1}\} \\ &= \frac{2}{N} \left[ 1 - (1-r/N)^N \right] - \frac{r}{N} \left( 1 - \frac{r}{N} \right)^{N-1} \quad (10) \end{aligned}$$

$$\begin{aligned} P_{A,b=1} &= (N/r) P\{A_{ij}\}_{b=2} \\ &= \frac{2}{r} \left[ 1 - \left( 1 - \frac{r}{N} \right)^N \right] - \left( 1 - \frac{r}{N} \right)^{N-1} \quad (11) \end{aligned}$$

$b=3$  경우에는 식(1)로부터

$$\begin{aligned}
 P\{A_{ij}\}_{b,3} &= P\{B_{ij}\} + P\{B_{ij_2}\} + \sum_{k=3}^N (3/k) P\{B_{ijk}\} \\
 &= 3 \sum_{k=1}^N (1/k) P\{B_{ijk}\} - 2P\{B_{ij_1}\} \\
 &\quad - (1/2) P\{B_{ij_2}\} \\
 &= \frac{3}{N} \left[ 1 - \left(1 - \frac{r}{N}\right)^N \right] - 2 \left(\frac{r}{N}\right) \left(1 - \frac{r}{N}\right)^{N-1} \\
 &\quad - \frac{1}{2} \left(\frac{r}{N}\right)^2 \left(\frac{N-1}{1}\right) \left(1 - \frac{r}{N}\right)^{N-1} \quad (12)
 \end{aligned}$$

$$\begin{aligned}
 P_{A,b,3} &= (N/r) P\{A_{ij}\}_{b,3} \\
 &= \frac{3}{r} \left[ 1 - \left(1 - \frac{r}{N}\right)^N \right] - 2 \left(1 - \frac{r}{N}\right)^{N-1} \\
 &\quad - \frac{1}{2} \left(\frac{r}{N}\right) \left(1 - \frac{r}{N}\right)^{N-2} (N-1) \quad (13)
 \end{aligned}$$

그러므로 식(1)에 의하여 b=1일때 식(9), b=2일때 식(10), b=3일때 식(12)와 같은 방법으로 일반식을 표현하면 다음과 같다.

$$\begin{aligned}
 P\{A_{ij}\} &= \sum_{k=1}^{b-1} P\{B_{ijk}\} + \sum_{k=b}^N (b/k) P\{B_{ijk}\} \\
 &= b \sum_{k=1}^N (1/k) P\{B_{ijk}\} - \sum_{k=1}^{b-1} (b/k - 1) P\{B_{ijk}\} \\
 &= \frac{r}{N} \left[ 1 - \left(1 - \frac{r}{N}\right)^N \right] \\
 &\quad - \sum_{k=1}^{b-1} \left(\frac{b}{k} - 1\right) \left(\frac{r}{N}\right)^k \left(1 - \frac{r}{N}\right)^{N-k} \left(\frac{N-1}{k-1}\right) \quad (14)
 \end{aligned}$$

그리고

$$P_A = (N/r) P\{A_{ij}\} \quad (15)$$

하나의 프로세서가 하나의 request를 발생할 확률이 r이고 출력단에서 하나의 request가 받아들여질 확률이 P<sub>A</sub>이므로 N개의 출력단에서 request를 발생하면 Nr request가 출력단에서 받아들여져 경로가 형성되어 Clustering 확률은 식(16)과 같다.

$$P_C = P_A^{Nr} \quad (16)$$

Cluster mode를 이용한 MINS의 해석 모델은 실제적인 시스템이 request를 발생시켰을때 각네트워크의 효율성을 비교하는데 유용하다. 그러나 MINS의 응용에 있어서 네트워크를 효율적으로 이용하기 위해서는 정교한 알고리즘을 이용되어야 한다. 이러한 알고리즘을 이용하여 최적적인 request를 발생시키므로서 MINS에서 발생하는 request들의 충돌(conflict)에 의한 문제점들이 줄어들수있다. 이러한 알고리즘은 Cluster가 오랜 시간 동안 형성되기 때문에 Cluster mode의 응용에서는 가능하다. 그러므로 Clustering 확률은 식(16)을 이용하여 r=1.0일때 네트워크의 수를 변수로 나타내리면 그림6에서와 같다. 그림6에서는 네트워크의 수 b=1이던 Clustering 확률 P<sub>C</sub>는 프로세서의 수 변화에 따라 확률1에 접근

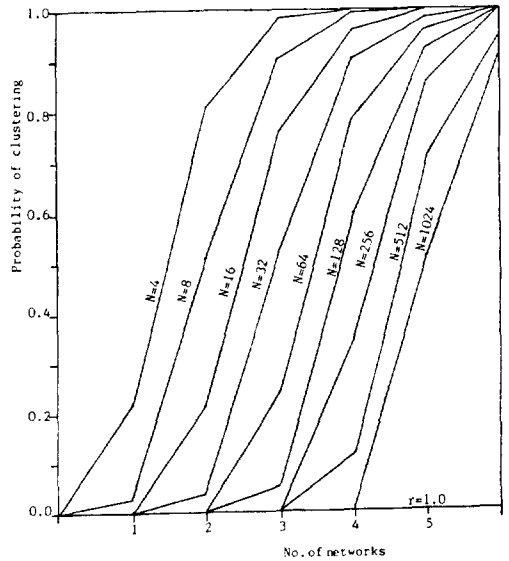


그림 6 Baseline 네트워크의 Clustering 확률  
Fig.6 Probability of clustering in baseline network.

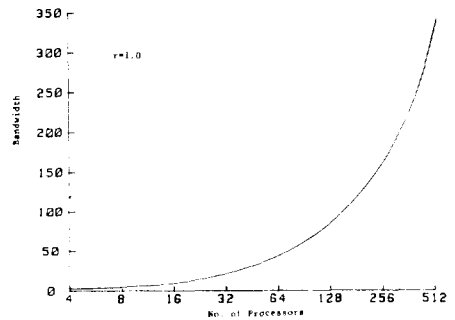


그림 7 N x N 네트워크의 Bandwidth  
Fig.7 Bandwidth of N x N network.

함을 알수있다. 즉 그림에서 얻을수있는 중요한 사실은, 만약 대부분의 프로세서가 request를 일정하게 도착점에 발생시키면 하나의 네트워크 시스템을 운용하는데 부족함을 알수있다. 그림7에서 Bandwidth는 r=1.0일때 프로세서 수를 변수로 나타냈으며 네트워크의 크기N을 증가시키면 Bandwidth가 증가됨을 알수있다.

### 5. MINS의 성능분석

시스템의 성능은 호스트 프로세서와 MINS의 특성

에 따라 좌우된다. 여기에서는 다중 프로세서 시스템을 구성하는 중요한 요소인 MINS, Generalized Shuffle network, Delta network, Star network의 성능과 비교 검토한다.

5.1 Delta 네트워크 분석

Delta 네트워크를 분석하기 위해서  $a \times b$ -크로스바 모듈로 구성된  $a^n \times b^n$  네트워크로 가정한다. 따라서  $a \times b$  모듈에서 request는 출력단  $b$ 에 독립적으로 균일하게 주어진다. 그래서 식(5)를  $a \times b$  모듈 Delta 네트워크에 적용할수있다. 각 입력단에서 request율이  $r$ 로 주어질때 일정한 시간에 request가 통과할 기대치  $E(a)$ 는

$$E(a) = b - b(1 - r/b)^a \tag{17}$$

식(17)에 출력단 request율을  $r_{out}$ 이고 입력단의 request율을 변수로 나타내면 식(19)와 같다.

식(17)을 출력단의 수  $b$ 로 나누면

$$E(r) = 1 - (1 - r/b)^a \tag{18}$$

$$r_{out} = 1 - (1 - r_{in}/b)^a \tag{19}$$

어떤 출력단의 request율은 다음단의 입력 request율이다. 이러한 과정을 처음단에서 마지막단까지 반복하므로써 Delta 네트워크의 Bandwidth를 결정할수 있다.

$a \times b$  Delta 네트워크의 Bandwidth  $B(a^n, b^n)$ 는 주어진 request율  $r$ 에서

$$B(a^n, b^n) = b^n r_i \tag{20}$$

$r_i$ 는  $i$ 단의 출력 request율이다. ( $i = 1, 2, 3 \dots n$ )

$$r_i = 1 - (1 - (r_{i-1}/b))^a \quad \text{단, } r_0 = r$$

그래서 하나의 request가 받아들여질 확률  $P_A$ 는

$$P_A = \frac{b^n r_i}{a^n r} = \frac{b^n \left[ 1 - \left( 1 - \frac{r_{i-1}}{b} \right)^a \right]}{a^n r} \tag{21}$$

5.2 Shuffle 네트워크 분석

Bandwidth와 request가 받아들여질 확률  $P_A$ 는 Pease<sup>8)</sup>에 의하여 다음과 같다.

Delta 네트워크에서  $m \times n$  모듈을 하였으나 Shuffle 네트워크에서  $m \times n$  모듈을 이용하였다.

$$BW = n - n \left( 1 - \frac{r}{n} \right)^m \tag{22}$$

$$P_A = \frac{n}{r \cdot m} = \frac{n}{r \cdot m} \left( 1 - \frac{r}{n} \right)^m \tag{23}$$

BW를  $n$ 로 나누면  $m \times n$  크로스바 모듈의  $n$ 출력단에서 하나의 request율을 나타낸다.

$$r_{out} = 1 - \left( 1 - \frac{r_{in}}{n} \right)^m \tag{24}$$

Shuffle네트워크에서  $i$ 번째단의 출력 request율은 역시  $(i+1)$ 번째단의 입력 request율이다. 그래서  $r$ 번째단의 출력 request율은 Shuffle네트워크의 BW를 결정해준다.

$i$ 번째단의 출력 request율을  $r_i$ 라하면

$$r_i = 1 - \left[ 1 - \left( \frac{r_{i-1}}{n} \right)^m \right] \tag{25}$$

( $r_0$ 는 하나의 프로세서가 발생하는 request율이다.)

$$BW = n \times r_i \tag{25}$$

그러므로 Shuffle네트워크에서 request가 받아들여질 확률은

$$P_A = \frac{n \cdot r_i}{m \cdot r_0} \tag{26}$$

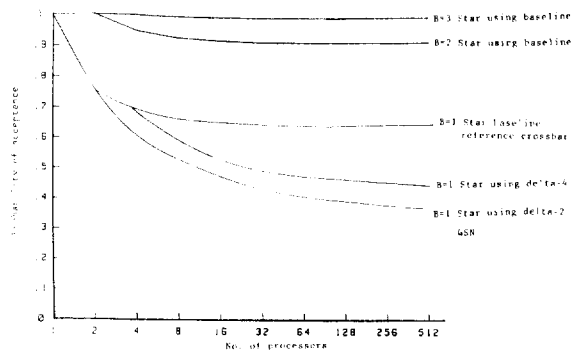


그림 8  $N \times N$  네트워크의  $P_A$ .  
Fig.8  $P_A$  of  $N \times N$  network.

그림8에서 Shuffle네트워크에서  $P_A$ 는 크로스바 네트워크와 같이  $N > 8$  이상에서 거의 일정한 확률을 가진다. 그래서 Shuffle네트워크는 MINS설계에 좋은 특성을 나타낸다.

위에서 언급한 네트워크와 Star네트워크 성능을 비교하면  $b=2$ 일때에 확률1에 접근하는 일정한 확률을 나타낸다. 그러므로 고도의 성능을 가진 다중 프로세서 시스템을 설계하기 위해서는 네트워크 수 2개 이상이 적정함을 보여준다.

6. 결 론

본 논문에서는 Star네트워크에서 이용되는 MINS의 성능을 분석하고, 기존의 네트워크 Generalized

Shuffle network, Delta network Crossbar network의 성능과 비교 검토하였다.

1) Star네트워크의 프로세서의 수를 증가시키면Bandwidth가 증가됨을 알수있다.

2) Generalized Shuffle network에서 request가 많을수록 확률  $P_{AN}$ 은 크로스바네트워크와 비슷하게  $N > 8$ 에서 변화가 없음을 알 수 있다. 그래서 MINS단계에서 최적적인 Shuffle의 수를 제시해준다.

4) Star네트워크의 확률  $P_{AN}$ 은 네트워크의 수를 증가시킬수록 확률1에 접근함을 알수있다.

위에지 보여준 결과에 따라 코드의 성능을 가진 Star네트워크를 설계하기 위해서는 2개 이상의 Baseline네트워크가 적정함을 나타냈다.

참 고 문 헌

- 1) T. Feng, "A Survey of Interconnection Networks", Computer, vol. 14, no. 12, Dec. 1981, pp. 12-27.
- 2) H. S. Stone, "Parallel Processing with the Perfect Shuffle,"IEEE Trans. Computers, vol. C-20, no. 2, Feb. 1971, pp. 153-161.
- 3) C. Wu and T. Feng, "On a Class of Multistage Interconnection Networks,"IEEE Trans. Computers, vol. C-29, no.8, Aug. 1980, pp. 694-702.
- 4) T. Feng, "Data Manipulating Functions in Parallel Processors and Their Implementations", IEEE Trans. Computers, vol. C-23, no. 12, Mar. 1974, pp. 309-318.
- 5) L. R. Goke and G. J. Lipovski, "Banyan Networks for Partititoning Multiprocessing Systems," First Annual Computer Architecture Conference, Dec. 1973, pp. 2128.
- 6) D. H. Lawrie, "Access and Alignment of Data in an Array Processor,"IEEE Trans. Computers, vol. C-24, no. 12, Dec. 1975, pp. 1145-1155.
- 7) K. E. Batcher, "The Flip Network in STARAN", 1976 International Conference on Parallel Processing, Aug. 1976, pp. 65-71.
- 8) M. C. Pease, "The Indirect Binary n-Cube Microprocessor Array,"IEEE Trans. Computers, vol. C-26, no. 5, May 1977, pp. 548-573.
- 9) J. H. Patel, "Performance of Processor Memory Interconnections for Multiprocessors", IEEE Trans. Computers, vol. C-30, no. 10, Oct. 1981, pp. 771-780.
- 10) D. S. Parker and C. S. Raghavendra, "The Gamma Network", IEEE Trans. Computers, vol. C-33, no. 4, April 1984, pp. 367-373.
- 11) S. Thanawastien and P. Nelson, "Interference Analysis of Shuffle / Exchange Network", IEEE Trans. Computers, vol. C-30, no. 8, Aug. 1981, pp. 546-556.
- 12) D. M. Dias and J. R. Jump, "Analysis and Simulation of Buffered Delta network", IEEE Trans. Computers, vol. C-30, no. 4, Apr. 1981, pp. 546-556.
- 13) L. Bhuyan and D. P. Agrawal, "Design and Performance of Generalized Interconnection Networks", IEEE Trans. Computers, vol. C-32, no. 12, pp. 1081-1090.
- 14) S. Cheemalavagu and M. Malek, "Analysis and Simulation of Banyan Interconnection Networks with 2x2, 4x4, and 8x8 Switching Elements,"P-roceedings Real-Time Systems Symposium, 1982, pp. 83-89.
- 15) C. D. kruskal and M. Snir, "The Performance of Multistage Interconnection Networks for Multiprocessors," IEEE Trans. Computers, vol. C-32, no. 12, April 1982, pp. 83-89.
- 16) E. D. Brokks, III, "Performance of the Butterfly Processor Memory Interconnection in a Vector Environment," 1985International Conference on Parallel Processing, 1985, pp. 21-24.
- 17) C. Wu and T. Feng, and M. Lin, "Star : A Local Network Systems for Real-Time Management of Imagery Data", IEEE Trans. Computers, vol. C-31, no. 10, Oct 1982, pp. 923-933.10)
- 18) M. Lee et al., "Network Facility for a Reconfigurable Computer Architechure", Fifth International Conference on Distributed Computing, 1985, pp. 264-271.
- 19) C. Wu et al, "Prototype of Star Architecture A Status Report", Proceedings of the National Computer Conference, 1985, pp. 191-201.