

적응적 결함-허용 다단계 상호연결망

김 금호, *김 영만, *배 은호, 윤 성대
부경대학교 전자계산학과, *부경대학교 전산교육학과
전화: 051-620-6398 / 핸드폰 : 016-724-1130

Adaptive Fault-tolerant Multistage Interconnection Network

Gum-Ho Kim, Young-Man Kim, Eun-Ho Bae, Sung-Dae Youn
Dept. of Computer Science, Pukyong National University
E-mail : kim3513@dol.pknu.ac.kr

Abstract

In this paper, we proposed and analyzed a new class of irregular fault-tolerant multistage interconnection network named as Extended-QT(Quad Tree) network. E-QT network is extended QT network. A unique path MIN usually is low hardware complexity and control algorithm. So we proposes a class of multipath MINs which are obtained by adding self-loop auxiliary links at the all stages in QT(Quad Tree) networks so that they can provide more paths between each source-destination pair. The routing of proposed structure is adaptived and is based by a routing tag. Starting with the routing tag for the minimum path between a given source-destination pair, routing algorithm uses a set of rules to select switches and modify routing tag. Trying the self-loop auxiliary link when both of the output links are unavailable. If the trying is failure, the packet discard. In simulation, an index of performance called reliability and cost are introduced to compare different kinds of MINs. As a result, the proposed MINs have better capacity than QT networks.

I. 서론

하드웨어 발달과 더불어 컴퓨터 네트워크와 디지털 데이터통신은 놀라운 속도로 발전해 왔다. 네트워크를 구성하는 요소 중 교환망은 다양한 유형이 있지만 그 중 가장 보편적으로 사용되는 다단계 상호연결망

(MIN:Multistage Interconnection Network)은 끊임없이 연구되어지고 있다[1]. 1980년대 초반부터 관심을 가져온 다단계 상호연결망은 다중 병렬처리 시스템에서 다수의 프로세서와 다수의 프로세서 또는 메모리 모듈간의 상호연결망으로 사용되어지고 있다. 다중 병렬처리 시스템의 성능은 상호연결망의 설계에 의해 좌우되며, 넓은 대역폭과 높은 속도를 요구하는 광대역 종합정보통신망(B-ISDN)의 가장 핵심 기술인 ATM 교환망의 기본구조로 대두되고 있다[2]. 다중 병렬처리 시스템은 수 백 혹은 수천 개의 프로세서를 포함하게 되므로 프로세서 수가 증가함에 따라 결함이 발생할 확률도 증가하게 된다. 그러므로 보다 높은 신뢰성을 유지하기 위해서는 결함이 발생했을 경우에도 동작하는 결함-허용 라우팅 알고리즘 및 하드웨어 체계가 필요하다. 최근에는 cost면에서도 효율성이 있고, 결함-허용이 가능한 불규칙 다단계 상호연결망의 연구가 이루어지고 있다[3,4,5]. 단일 경로 제공 시 스위치 결함이나 블록 상태를 경험하면 패킷이 폐기되므로 다중 경로를 두어 결함이 발생하더라도 이를 허용함으로써 불규칙 다단계 상호연결망의 처리율을 증가시킨다 [3]. 기존 연구로는 Bansal & Saroha가 제안한 DOT를 첨가함으로써 결함-허용 하는 FDOT(fault-tolerant double tree)[4]가 있고 그러나 규칙적인 MIN보다 많은 Cost를 유지하는 단점이 있고, Yingquan와 Yinghua가 제안한 QT network[5]는 마지막 단의 스위치 결함이나 패킷 충돌시 결함-허용을 할 수 없는 단점이 있다.

본 논문에서는 단일 경로가 아닌 다중 경로를 두어 결함을 허용하고, QT network에서 각각의 스위치에 자기 루프 링크를 첨가함으로써 출력단에 결함이 발생하면 다음 사이클 때 다시 출력을 시도하게 하여 기존의 다단계 상호연결망보다 더 효율적인 신뢰도를 제공하고 QT의 단점인 마지막 단의 결함에도 허용할 수 있도록 하였고 일반적인 MINs 보다 더 효율적인 cost

를 제공하는 불규칙 다단계 상호연결망을 제안하고자 한다.

본 논문의 구성은 다음과 같다. 2장에는 QT network의 구조에 대해 설명하고 3장에는 E-QT network에 대한 구조를 소개한다. 4장에는 라우팅 경로 길이와 라우팅 태그에 대해 설명하고 5장은 QT와 E-QT에 대한 신뢰도, Cost 대해 모의 실험 및 결과를 보여주고 6장에는 결론 및 향후과제를 제시한다.

II. QT network 구조

2.1 MDOT 구조

DOT 구조는 불규칙적인 다단계 상호연결망으로서 오른쪽 구조와 왼쪽 구조로 구성되어 있다. 각각의 구조는 이진 트리와 유사하다. DOT network는 $2^n \times 2^n$ 크기를 가진 구조는 $2n-1$ 개의 단을 가지고 있고 $2^{n+1}-3$ 개의 스위치를 가지고 있다. 그림1은 8×8 MDOT network 구조이다.

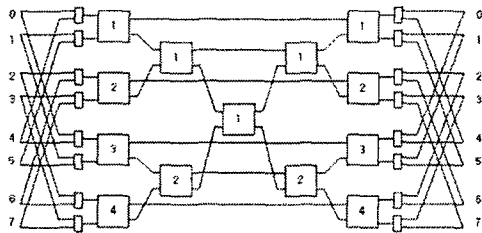


그림1 8×8 MDOT network

2.2 QT(Quad Tree) 구조

QT network는 MDOT network의 확장된 형태로써 $2^n \times 2^n$ 크기를 가진 QT network는 $2^{n-1} \times 2^{n-1}$ 의 MDOT의 두 개로 구성되어 있다. 두 개의 그룹은 소스의 최상위 비트인 MSB비트로 구분한다. 예를 들면 소스가 0에서 7사이의 최상위비트는 0이므로 G^0 로 나타낸다. 하나의 그룹에서 i 단과 $2m-i$ 단은 2^{n-i-1} 개의 스위치를 가지고 있다. 그리고 각 스위치에 1, 2, ..., 2^{n-i-1} ($i=1, 2, \dots, m$ 일때 $m=\log_2 N/2$, $N=2^n$) 숫자를 매긴다. 마지막 단을 제외한 나머지 단의 스위치에는 보조적인 링크를 통해 다른 그룹의 같은 숫자를 가진 스위치로 연결되어 있다. 각각의 소스와 목적지는 두 개의 그룹을 연결하는데 멀티플렉서와 디멀티플렉서로 사용한다. QT network는 $2m-1$ 개의 단을 가지고 있고 $2^{m+2}-6$ 개의 스위치로 구성한다. 그림2는 16×16 의 QT network를 나타낸다.

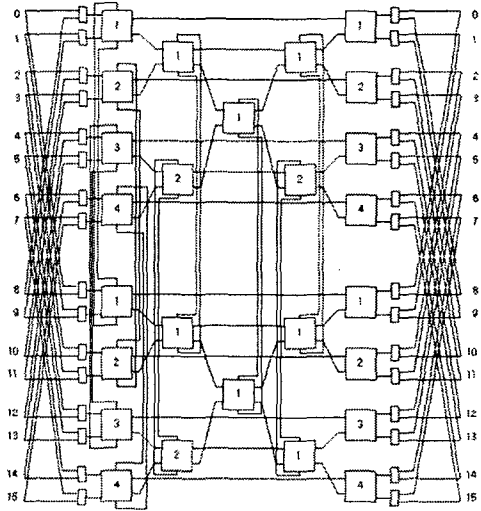


그림 2 16×16 QT network

III. E(Extended)-QT network의 구조

E-QT network는 QT network의 확장된 형태이다. 크기가 $2^n \times 2^n$ 인 E-QT network는 G^{n-1} 의 두 개의 그룹으로 나누어 만들어진다. 각각의 그룹은 $2^{n-1} \times 2^{n-1}$ 의 MDOT network로 구성된다. QT와 E-QT의 차이점은 QT는 그룹의 동일한 위치에 있는 노드끼리 보조적인 링크로 연결되어 있고 마지막 단에는 보조적인 링크가 없다. 그러나 E-QT network는 보조적인 링크가 자기 자신 스위치로 순환하게 만들었고 마지막 단 역시 순환 스위치로 만들어져 있다. E-QT 네트워크의 단의 수는 $2n-3$ 이고 하나의 단에 2^{n-i} 개의 스위치를 가지고 있다. 전체의 스위치 수는 $2^{n+1}-6$ 개의 스위치를 가진다. 그리고 QT의 마지막 단에서 출력단에 블록이 발생할 경우 다시 한번 출력을 시도함으로써 어느 정도의 신뢰도를 증가시킨다. 그림3은 16×16 의 E-QT network를 나타낸다.

IV. 라우팅 알고리즘

라우팅 알고리즘은 라우팅 경로 길이를 구하고 라우팅 태그를 계산하고 라우팅 하는 것은 아래의 절에서 설명한다. 그리고 여러 가지 라우팅 태그가 나옴으로써 그중 경로 길이가 적은 라우팅 태그부터 라우팅을 하는 적응적 라우팅 알고리즘을 소개한다.

4.1 경로 길이 구하는 알고리즘

적응적 결합-허용 다단계 상호연결망

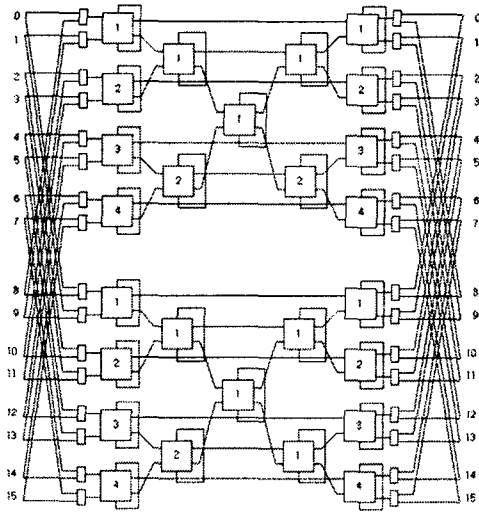


그림3 16×16 E-QT network

라우팅 알고리즘에서 우리는 소스에서 목적 주소까지 최소의 경로에 도달하기를 원할 것이다. 그러기 때문에 최소의 경로를 구하는 알고리즘에 대해 알아보자 알고리즘1은 경로 길이를 구하는 알고리즘이다.

```

if
     $(s_{n-2} \oplus d_{n-2}) + (s_{n-3} \oplus d_{n-3}) + \dots + (s_1 \oplus d_1)$ 
    is zero
then
    minimum_path = 2, 4, \dots, (2n-2), (2n-1)
else
if
     $(s_{n-2} \oplus d_{n-2}) + (s_{n-3} \oplus d_{n-3}) + \dots + (s_2 \oplus d_2)$ 
    is zero
then
    minimum_path = 4, \dots, (2m-1)
else
    .
    .
    .
else
    minimum_path = 2m-1 // 가장 긴 경로 길이
    
```

알고리즘1 경로 길이 알고리즘

4.2 라우팅 태그

E-QT network를 위한 라우팅 알고리즘은 주어진 소스와 목적지 사이의 경로를 만들기 위해 분산 라우

팅 태그를 계산한다.

```

if
     $2 \leq x < 2m-1$ 
then routing tag =
     $s_{n-1} \cdot (1.1.1.)_{(x/2-1)} \cdot 0 \cdot (d_{(x/2-1)} \dots d_0) \cdot d_{n-1}$ 
else
if
     $x=2m-1$ 
then routing tag =
     $s_{n-1} \cdot (1.1.1.)_{x/2} \cdot 0 \cdot (d_{x/2} \dots d_0) \cdot d_{n-1}$ 
else
    no tag is possible
    
```

알고리즘2 라우팅 태그 계산 알고리즘

4.3 라우팅 절차

신뢰도와 성능을 향상시키기 위해 다중 경로를 제공하는 네트워크에 대한 라우팅 알고리즘을 제안한다.

라우팅 알고리즘 절차

[단계1] 알고리즘1에 의해 최소 경로를 구한다

[단계2] [단계1]에서 구한 최소경로를 가지고 알고리즘2)를 통해 라우팅 태그를 계산

[단계3] 경로의 길이가 가장 작은 라우팅 태그를 가지고 첫 번째 멀티플렉서로 연결

[단계4-1] 만약 출력단에서 충돌이 발생하면 다음 사이클에서 다시 출력 시도

[단계4-2] 만약 다른 라우팅 태그가 있으면 다른 라우팅 태그를 가지고 라우팅

[단계4-3] [단계4-1] 반복

[단계5] [단계4]에서 실패로 끝나면 두 번째 멀티플렉서로 연결

[단계6] 출력단에 충돌이 발생하면 단계4) 시도

[단계7] [단계6]도 실패하면 패킷을 폐기

V. 모의실험 및 결과

제안된 구조에 대한 신뢰도와 Cost에 대해 비교한다.

5.1 신뢰도 측정

제안된 네트워크 구조의 종단간 신뢰도를 보이기 위해 네트워크 크기에 따라 모의실험을 수행한다. 모의 실험에서 패킷 생성과 목적지 주소는 랜덤 함수로 발생시켜 얻고, 같은 네트워크 크기에서는 동일한 값을 가지고 QT network와 E-QT network를 비교하였다. E-QT network와 QT network 적응적 라우팅을 하기 때문에 최단 경로를 갖는 라우팅 태그를 가지고 모의 실험을 했다. 소스는 하나의 사이클 때마다 항상 발생

한다.

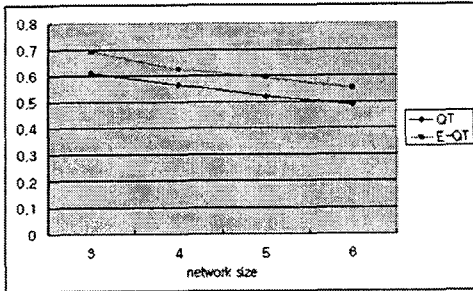


그림4 network size에 따른 QT와 E-QT 신뢰도 비교

결과를 보면 8×8 네트워크 크기에서는 QT는 0.61의 처리율, E-QT는 0.69의 처리율을 가지고, 64×64 네트워크 크기에서는 QT는 0.49 E-QT는 0.55의 처리율을 보인다. E-QT network는 QT network 보다 더 높은 신뢰도를 보이고 있음을 알 수 있다.

5.2 Cost

다단계 상호연결망에서 하드웨어 복잡도는 단의 수, 각 단계 스위치 수와 링크의 수에 결정한다. 이것을 "Cross Point"라는 것이다. Cross Point는 멀티플렉서와 디멀티플렉서의 링크의 총수와 스위치 링크의 총수를 나타내며 예를 들면 $k_1 \times k_2$ 스위치 cross point 수는 $k_1 \times k_2$ 이고 $1 \times k$ 디멀티플렉서나 $k \times 1$ 멀티플렉서는 k 라는 값이 나온다. 우리는 이러한 결과로 Cross Point를 계산한다. 표2)는 하드웨어 복잡도를 나타낸 것이다.

n	ESC	3-rep	INDRA	QT	E-QT
3	128	144	192	102	122
4	288	384	320	258	298
5	640	960	768	570	650
6	1408	2304	1792	1244	1904

표1 각 MIN에 대한 Cost의 비교

VI. 결론 및 향후과제

본 논문에서는 QT network의 다른 그룹의 같은 스위치로 연결된 링크를 자기 순환 링크로 하고 마지막 단계 자기 순환 링크를 첨가하고 E-QT network를 제안했다. 적응적 라우팅 알고리즘을 사용하여 패킷 전송 중 충돌이 발생한 경우 스위치 자신을 순환하고, 재차 충돌이 발생하면 다른 경로를 통해 라우팅 한다. 그리고 다중 경로를 제공함으로써 어떠한 스위치 결합이나 출력단 충돌에서도 결합을 허용하는 불규칙 다단

계 상호연결망을 제안했다.

모의실험을 통해 네트워크의 처리율을 분석하여, 제안된 구조의 종단간 처리율 즉 소스에서 목적지까지 결합이 없는 경로를 발견하는 라우팅 알고리즘의 처리율이 네트워크의 크기가 증가함에 따라 거의 더 신뢰도가 높음을 알 수 있지만 Cost면에서는 일반적인 다단계 상호연결망보다는 Cost가 낮음을 알 수 있지만 QT network 보다는 조금 높음을 알 수 있다.

향후과제로는 본 논문에서 제시되지 않은 통신 지연 시간과 Capacity를 고려한 성능 평가를 본 논문에서 제안한 모델을 적용하는 것이다.

참고문헌

- [1] C. L. Wu, and T. Y. Feng, "On a class of multistage interconnection network", IEEE Trans. Computers, Vol. C-29, pp. 694-702, Aug. 1980
- [2] F. A. Tobagi, "Fast packet switch architectures for broadband integrated service digital network", Proc. IEEE, Vol. 78, No. 1, pp. 133-167, Jan. 1990
- [3] G. B. AdamsIII, D. P. Agrawal and H. J. Siegel, "A survey and comparison of fault-tolerant multistage interconnection networks", IEEE Computers, pp. 14-27, June 1987.
- [4] P. K. Bansal, Kuldip Singh, R. C. Joshi and G. P. Saroha, "Fault tolerant double tree network", proc. of International Conference IEEE INFOCOM91' April, pp. 458-462
- [5] P. K. Bansal, K. Singh and R. C. Jonshi "Quad tree : A cost-effective fault-tolerant multistage interconnection network", IEEE INFOCOM92' pp. 806-866
- [6] Yingquan Zhou and Yinghua Min "Adaptive message routing in a class of fault-tolerant multistage interconnection networks", comp Elect. Vol 23, pp.239-247, 1997.
- [7] C. P. Kruskal and M. Snir, "Performance of Multistage interconnection network for multiprocessors", IEEE Trans. Comp., Dec. 1983
- [8] C. S. Raghavendra and A. Varma, "INDRA : A class of interconnection networks with Redundant paths", 1984 Real Time System symp., Spring, Md., 1984
- [9] G. B. AdamsIII, and H. J. Siegel, "The Extra Stage Cube : A Fault-Tolerant Interconnection Network for Supersystems", IEEE Trans. on Computers, Vol. C-31, No. 5, May 1982.