

AMD 고장감내 다단계 상호 연결망의 설계 및 분석

김 정 선[†]

요 약

고성능 병렬처리 시스템에서 다단계 상호 연결망(Multistage Interconnection Network(MIN))은 프로세서간에 또는 프로세서와 메모리 모듈간에 가격대 효율 면에서 효과적으로 고대역폭의 통신을 제공한다. 본 논문에서는 고장감내 기능, 모듈 구조, 우수한 성능 등으로 인하여 실시간 병렬/분산 처리 시스템에 유용한 복수 경로를 갖는 다단계 상호 연결망-AMD(Augmented Modified Delta) 연결망-의 설계 방법을 제시하였고 그 성능과 신뢰성을 분석하였다. 한편으로 AMD 연결망은 기존 Kappa 연결망의 우수한 성능 및 고장감내 특성을 유지하면서, 다른 한편으로 Kappa 연결망과는 달리 임의의 Delta 연결망으로부터 일관성 있고 쉽게 설계될 수 있다. 또한 Delta 연결망의 경우와 마찬가지로, 모든 AMD 연결망들은 위상학적으로 서로 동가인 특성을 갖는다.

Design and Analysis of a Class of Fault Tolerant Multistage Interconnection Networks: the Augmented Modified Delta (AMD) Network

Jungsun Kim[†]

ABSTRACT

Multistage interconnection networks(MINs) provide a high-bandwidth communication between processors and/or memory modules in a cost-effective way. In this paper, we propose a class of multipath MINs, called the Augmented Modified Delta(AMD) network, and analyze its performance and reliability. The salient features of the AMD network include *fault-tolerant capability*, *modular structure*, and *high performance*, which are essential for real-time parallel/distributed processing environments. The class of the AMD network retains well-known characteristics of the Kappa network, but its design procedure is more systematic. Like Delta networks, all the AMD networks are topologically equivalent with each other.

1. Introduction

Parallel computers with multiple processors are opening the door to the high performance computing to

meet the increasing demand of computational power. There exists a wide spectrum of parallel computer architectures today, including symmetric multiprocessors, distributed shared memory computers, message passing multicomputers, and networks of workstations (NOWs) [1, 2, 3, 4]. In these modern parallel computers, interconnection networks (INs) play an important role in the overall system performance be-

※본 연구는 96 교내 연구비에 의하여 수행되었음.

† 정 회 원: 한양대학교 전자계산학과 전임강사

논문접수: 1996년 10월 4일, 심사완료: 1997년 7월 29일

cause a considerable amount of time may be spent just for interprocessor communications and/or memory accesses. Thus the construction of a flexible high bandwidth interconnection network at a reasonable cost is a critical problem, and has been the subject of extensive study over the past decade [4, 5, 6, 7, 8, 9, 10]. Multistage interconnection networks (MINs) have been introduced as one of the solutions to meet the above objectives in a cost-effective way as a balance between the two extreme interconnection schemes [6]: a *crossbar* which may be cost prohibitive and a *shared bus* which may be inefficient and unreliable.

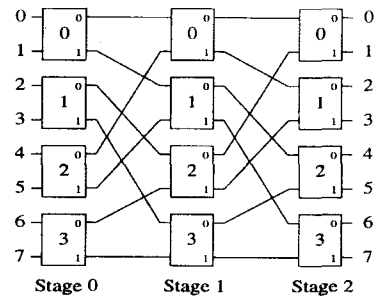
In this paper, we discuss the design of a class of fault tolerant MINs, called the Augmented Modified Delta (AMD) network. The topology and control strategy of the AMD network are explained, and its performance and reliability analysis is presented. The results of the analysis are tabulated for various sized networks. The paper is organized as follows. Section 2 provides background information about MINs and motivation of the research. Section 3 explains the design procedure and the control strategy for the AMD network. Performance analysis and reliability analysis are presented in Section 4 and 5, respectively. Finally, Section 6 provides concluding remarks.

2. Backgrounds

A MIN consists of several stages of switching elements (SEs) with interconnections between adjacent stages. A MIN normally connects N inputs (sources) to N outputs (destinations) and is referred to as an $N \times N$ MIN. The parameter N is called the size of the network, and only networks of size $N (= 2^n)$ will be considered for the sake of simplicity.

The topology of a MIN is formally defined as the pair (S, Λ) where $S = (S_0, S_1, \dots, S_{n-1})$, n is the number of the stages, $S_i = p_i \times q_i$ denotes the size of an SE at stage i , and Λ is the interconnection function for interstage links [5]. The function Λ is specified using the format $\Lambda(SE_i^j, \delta)$ where the first parameter

indicates an SE numbered j in stage i , and the second indicates the output labels of the SE. The SE in stage $(i + 1)$ that is connected via the δ^{th} output of SE_i^j is represented by the function application $\Lambda(SE_i^j, \delta)$. An example of $2^3 \times 2^3$ MIN is shown in (Fig. 1).



(Fig. 1) A $2^3 \times 2^3$ delta network

A spectrum of MINs with $n = \log_2 N$ stages of 2×2 SEs has been introduced and analyzed in the literature, including the Baseline network [10], the Omega network [11], the Generalized n -cube network [12], the regular SW-Banyan (Spread = Fanout = 2) network [13], and etc. However, all these networks are topologically equivalent with each other [5, 10]. Thus we call these networks collectively as Delta networks. An $N \times N$ Delta network has the following characteristics [5].

- **Unique path property (UPP):** existence of a unique path between every source destination pair.
- **Block structure property:** partitioning of SEs into equal-sized disjoint subblocks, each of which is a self-contained UPP subnetwork, through successive stages.
- **Destination tag algorithm (DTA):** setting of the states of each SE on the path is digit controlled such that an SE connects its input to one of its outputs depending on a single digit (base 2) taken from the destination address.

The block structure of a network is a fundamental concept which provides insights into the partitionability, control strategy, fault-tolerance, and the modular

design for the VLSI implementation of the network. The DTA is well suited to the parallel processing environment and improves the overall speed of the system.

In contrast to these desirable features, Delta networks have two significant drawbacks as a direct consequence of the unique path property: *low performance* and *devoid of fault-tolerance*. As a solution to these problems, a number of MINs with multiple paths from source to destination have been devised and proposed. Such MINs are collectively called *multipath MINs*.

In this paper, we propose a class of multipath MINs called the Augmented Modified Delta (AMD) network. The AMD networks are characterized by the following features.

- They subsume the Delta network and can, therefore, achieve at least all the permutations as the relevant Delta network.
- They can tolerate any single SE fault and many combinations of multiple SE faults. The fault-tolerant capability of the AMD network is given in section 5.
- They have redundant paths from each source to all destinations and they are balanced. By the term "balanced", we mean the number of alternate paths between any source-destination pair is the same.
- They employ dynamic rerouting strategy that is a simple modification of a DTA.
- They have a higher bandwidth than the Delta network.
- The link permutations between successive stages are consistent and straightforward so that they retain the block structure property.

The design of the AMD network is motivated from the analysis of the Kappa network [15]. The Kappa network was designed to improve the performance and reliability of the Gamma network [16] by adding additional links at the block level so that the link pattern is symmetric without losing the permutation capability of the Gamma network. Thus, in order to

use the DTA [5], the Kappa network requires a relabeling scheme for the outputs of several SEs after establishing the entire network. This process is awkward and error-prone.

In contrast, AMD networks can be designed consistently and systematically starting from any Delta network, and no relabeling scheme is necessary. We can simply choose a specific Delta network of our primary interest as a base network. The design philosophy of the AMD network emphasizes the modularity and the regularity in the network design procedure. Moreover, all the AMD networks are topologically equivalent. Topologically equivalent networks can be reconfigured to simulate each other by logical relabeling of SEs, or with the hardwired link permutations at inputs and outputs [5, 10]. Therefore, the AMD network can be thought of as a class of fault-tolerant MINs, rather than simply refers to a specific MIN.

3. Design of the AMD networks

3.1 Preliminaries

Definition 1 [17] A switching element SE_i^p is a *predecessor* of an SE_{i+1}^q if an output link of the SE_i^p is connected to an input link of the SE_{i+1}^q . The SE_{i+1}^q is then a *successor* of the SE_i^p . An SE_{i+1}^q is the $O(1)$ -*successor* if the output $O(1)$ -link of an SE_i^p is connected to the SE_{i+1}^q .

Definition 2 [17] Two switching elements, SE_i^p and SE_i^q , are said to be *conjugates* of each other if they have the same pair of successors, and denoted by $SE_i^q = Conj(SE_i^p)$ and $SE_i^p = Conj(SE_i^q)$. Then the pair of successors are called *inverse conjugates* of each other.

Definition 3 [18] If the link connection patterns between successive stages of an $N \times N$ UPP network with 2×2 SEs are arranged such that for every switching element SE_i^p , there exists another switch which is a conjugate of the SE_i^p , and two pairs of conjugate SEs constitute two inverse conjugate SEs except the first stage, then the network is said to be *strict*.

Delta networks fall into the strict UPP network

category.

Theorem 1 [5] An $N \times N$ strict UPP network with 2×2 SEs is destination tag controlled if (1) it is block structured, and (2) the set of 0-successors and the set of 1-successors of SEs form disjoint subblocks through successive stages until subblocks containing a single SE are reached at the output stage.

Lemma 1 Let SE_i^p and SE_i^s be a pair of conjugates in an $N \times N$ strict UPP network with destination tag controlled. And let SE_{i+1}^r and SE_{i+1}^s be a 0-successor and a 1-successor of the conjugates SE_i^p and SE_i^s , respectively. Then $Conj(SE_{i+1}^r)$ belongs to the same subblock as SE_{i+1}^r and $Conj(SE_{i+1}^s)$ belongs to the same subblock as SE_{i+1}^s .

Proof: The Lemma is a direct consequence of Definition 3 and Theorem 1. □

The fault tolerant scheme of the AMD network is based on Lemma 1.

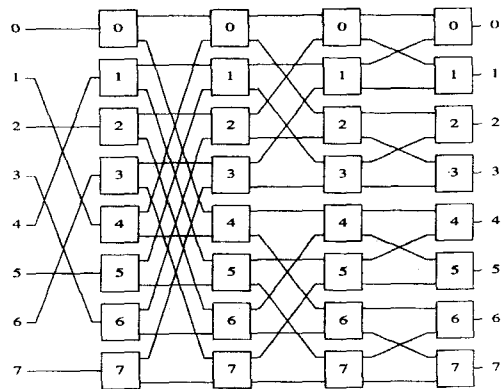
3.2 Design Procedure for the AMD Networks

The multipath MINs proposed here are constructed by adding redundant links at the block level to the Modified Delta Network (MDN) [5] in a consistent and systematic way. An $N \times N$ ($N = 2^n$) MDN consists of $n + 1$ stages of $N, 2 \times 2$ SEs except the first and the last stages. Formally, the topology of an $N \times N$ MDN can be defined as (S, A) where $S = S_0, S_1, \dots, S_n, S_0 = 1 \times 2, S_i = 2 \times 2$ ($0 < i < n$), $S_n = 2 \times 1$.

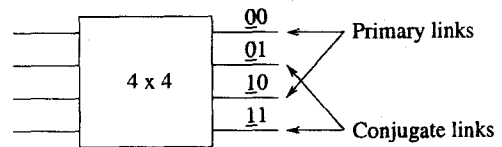
The A is the same as that of the original Delta network. For example, an MDN derived from the Generalized cube network [15] is shown in (Fig. 2).

In AMD networks, every 2×2 SE in the MDN is replaced by a 4×4 SE as shown in (Fig. 3). In addition, every source and destination is attached to a 1×4 demultiplexer SE and a 4×1 multiplexer SE, respectively. The 00 and 10 links are called *primary links* and are connected exactly as in the original MDN. The 01 and 11 links are named as *conjugate links* and are connected as follows.

- In an MDN, let SE_{i+1}^p be a 0-successor of an SE_i^j , and let SE_{i+1}^s be a 1-successor of the SE_i^j where



(Fig. 2) The $2^3 \times 2^3$ modified generalized cube network

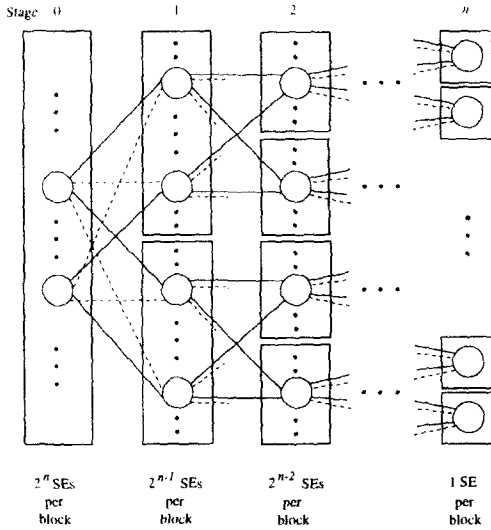


(Fig. 3) A 4×4 SE for the AMD network

$0 \leq i < n$. Then in the corresponding AMD network, the 01 and 11 output links of an SE_i^j are connected to $Conj(SE_{i+1}^p)$ and $Conj(SE_{i+1}^s)$, respectively.

- Note: Connection rule for the conjugate links can be described by the A function: $A(SE_i^j, \underline{00}) = SE_{i+1}^p$, $A(SE_i^j, \underline{10}) = SE_{i+1}^s$, $A(SE_i^j, \underline{01}) = Conj(SE_{i+1}^p)$, and $A(SE_i^j, \underline{11}) = Conj(SE_{i+1}^s)$.

The interconnection pattern of the AMD network and its validity can be easily understood by referring to (Fig. 4). In (Fig. 4), each level corresponds to the stage number, rectangular boxes denote blocks at a given level, circles stand for SEs, solid edges represent primary links, and dotted edges depict conjugate links in an AMD network. Different types of AMD networks are described below. A detailed explanation of each original Delta network is omitted here, and only the interconnection function A is given. By AMD- X network, we mean an AMD network based on the Delta network X . Note that all the AMD networks are based on a Delta network, and thus it can be easily proved, as in [18], that all the AMD networks are topologically equivalent with each other.



(Fig. 4) The interconnection pattern of the AMD network

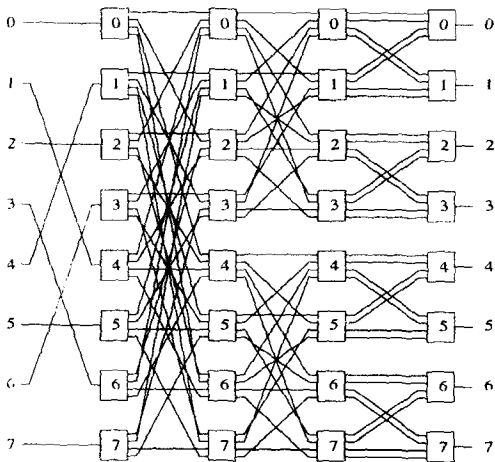
3.2.1 AMD - Generalized Cube Network

Let SE_i be represented by a sequence of binary digits $(p_n p_{n-1} \dots p_1)_i$.

- Two conjugate SEs have labels which differ only in the $(n-i-1)th$ bit position.
- connection rule

$$\Lambda((p_n \dots p_1)_i, \underline{00}) = (p_n \dots p_{n-i-1} \underline{0} p_{n-i-1} \dots p_1)_{i+1}$$

$$\Lambda((p_n \dots p_1)_i, \underline{01}) = (p_n \dots p_{n-i-1} \underline{0} \overline{p_{n-i-1}} \dots p_1)_{i+1}$$



(Fig. 5) Topology of the $2^3 \times 2^3$ AMD-generalized cube network

$$\Lambda((p_n \dots p_1)_i, \underline{10}) = (p_n \dots p_{n-i+1} \underline{1} p_{n-i-1} \dots p_1)_{i+1}$$

$$\Lambda((p_n \dots p_1)_i, \underline{11}) = (p_n \dots p_{n-i+1} \underline{1} \overline{p_{n-i-1}} \dots p_1)_{i+1}$$

- The topology of the AMD-Generalized cube network is shown in (Fig. 5).

3.2.2 AMD - Omega Network

- Two conjugate SEs have labels which differ only in the leftmost bit position.
- connection rule

$$\Lambda((p_n p_{n-1} \dots p_1)_i, \underline{00}) = (p_{n-1} p_{n-2} \dots p_1 \underline{0})_{i+1}$$

$$\Lambda((p_n p_{n-1} \dots p_1)_i, \underline{01}) = (\overline{p_{n-1}} p_{n-2} \dots p_1 \underline{0})_{i+1}$$

$$\Lambda((p_n p_{n-1} \dots p_1)_i, \underline{10}) = (p_{n-1} p_{n-2} \dots p_1 \underline{1})_{i+1}$$

$$\Lambda((p_n p_{n-1} \dots p_1)_i, \underline{11}) = (\overline{p_{n-1}} p_{n-2} \dots p_1 \underline{1})_{i+1}$$

3.2.3 AMD - Baseline Network

- Two conjugate SEs have labels which differ only in the rightmost bit position.

- connection rule

$$\Lambda((p_n \dots p_1)_i, \underline{00}) = (p_n \dots p_{n-i+1} \underline{0} p_{n-i} \dots p_2)_{i+1}$$

$$\Lambda((p_n \dots p_1)_i, \underline{01}) = (p_n \dots p_{n-i+1} \underline{0} p_{n-i} \dots \overline{p_2})_{i+1}$$

$$\Lambda((p_n \dots p_1)_i, \underline{10}) = (p_n \dots p_{n-i+1} \underline{1} p_{n-i} \dots p_2)_{i+1}$$

$$\Lambda((p_n \dots p_1)_i, \underline{11}) = (p_n \dots p_{n-i+1} \underline{1} p_{n-i} \dots \overline{p_2})_{i+1}$$

3.3 Fault-tolerant Destination Tag Algorithm (FDTA)

It is assumed that an SE can determine the fault state of the successor SEs. Consider a source $S = s_0 s_1 \dots s_{n-1}$ and destination $D = d_0 d_1 \dots d_{n-1}$. The FDTA [5] works the same way as the DTA. It uses the i -th digit d_i of the destination tag to set the SE at the stage i in the routing path. If there is no conflict in an SE in stage i and no faults detected in an SE and/or at an input link in stage $i+1$, switch to $d_i 0$ else switch to $d_i 1$. Note that the routing algorithm that is adopted by the AMD network is exactly the same as that of the Kappa network. However, unlike the Kappa network, an awkward relabeling scheme is not required in the AMD network.

The fault-tolerant capability of an AMD network can be described as follows. Let the sequence $S, SE_0^i, SE_1^i, \dots, SE_n^i, D$ denote a normal path along the pri-

mary links between a given source-destination pair (S, D), and the routing tag for the path is given by $d_0 d_1 \dots d_{n-1} (=D)$. Then it is clear from Theorem 1 and Lemma 1 that for any pair of SEs (SE_i^j, SE_{i+2}^j) on the path, $0 \leq i < n-1$, the following relation holds:

$$\Lambda(\Lambda(SE_i^j, \underline{d}_i 0), \underline{d}_{i+1} 0) = \Lambda(\text{Conj}(\Lambda(SE_i^j, \underline{d}_i 0)), \underline{d}_{i+1} 0).$$

This relation is based on the fact that application of the destination tag algorithm is a continuous process of identifying the block which contains the target SE to which the desired destination is connected. Since the two SEs, $\Lambda(SE_i^j, \underline{d}_i 0)$ and $\text{Conj}(\Lambda(SE_i^j, \underline{d}_i 0))$, belong to the same block, the concept of conjugate SEs ensures that a request will be routed to the same SE at stage $i+2$ depending on d_{i+1} .

4. Performance Analysis

In this section, the performance of the AMD network is given under the fault-free condition in a random access environment. The measure of merit commonly used in comparing the performance of diverse networks is the probability of acceptance (PA). The (PA) is the probability that a request will reach its destination without blockage. Detailed performance analysis of the AMD network in a (multiple) packet switching environment can be found in [19].

In our analysis, the following assumptions are made.

- Each source submits a request to the network every network cycle synchronously. The requests generated by the sources in each cycle are random and uniformly distributed among all destinations.
- The network is fault-free.
- The routing algorithm assigns alternate links with equal probability to a request having a corresponding destination tag digit.
- Conflict resolution in SEs is unbiased. If two or more requests compete for the same output, a conflict is said to occur. In case of conflicts, up to two requests are randomly chosen with equal probability.

- Requests that are not accepted are discarded. The requests generated in the subsequent cycle are independent of those submitted in the previous cycle.

The final assumption is unrealistic, but it has been used by several researchers for the sake of simplicity and feasibility. Judging from the resemblance of their topologies and hardware complexities, there is every reason to believe that the performance of the AMD network is the same as that of the Kappa network. This is verified by comparing the analytical result of the Kappa network with the simulation result of the AMD network. Both results were found to be in close agreement (within 1 percent). Therefore the details of the analytical analysis for the AMD network are omitted, and only the simulation result is provided in (Table 1).

(Table 1) PA of a request

Network Size	UPP Analysis	Kappa Analysis	AMD simulation	Gamma Analysis	Hybrid Analysis	ACN Analysis
8	0.5165	0.6564	0.6598	0.6422	0.5968	0.6486
16	0.4498	0.6423	0.6361	0.6083	0.5399	0.6220
32	0.3393	0.6325	0.6280	0.5783	0.4981	0.5990
64	0.3594	0.6244	0.6233	0.5511	0.4655	0.5786
128	0.3271	0.6169	0.6152	0.5263	0.4390	0.5600
256	0.3004	0.6098	0.6112	0.5036	0.4168	0.5431
512	0.2778	0.6030	NA	0.4828	0.3978	0.5276
1024	0.2585	0.5964	NA	0.4636	0.3814	0.5134

The simulation result for the AMD network was obtained as follows. Each simulation cycle corresponds to a network cycle. Each simulation was run for 2000 cycles. Tests indicate that extending the run time had little effect on the simulation results. Lastly, each simulation was repeated ten times and results were averaged in order to reduce the statistical variances. The entry marked "NA" in (Table 1) means that the simulation result is not available for the corresponding network size.

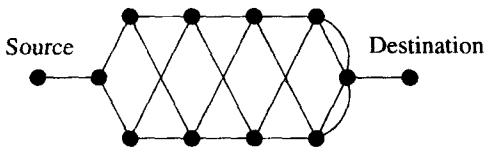
5. Reliability Analysis

As measures of the robustness of a MIN, the relia-

bility analysis can be done in terms of (1) the number of arbitrary SE faults the network can tolerate without losing the full access capability¹⁾, (2) the MTTF (Mean Time To Failure), and (3) the terminal reliability.

A network is said to be i fault-tolerant if it can function despite arbitrary set of i SE faults [7]. The MTTF of a MIN is the average time the network will operate before some source is disconnected from some destination. The terminal reliability [14] is the probability that at least one path is operational between a given source-destination pair, given that each SE has a certain reliability. We use a switch fault model [20] for the reliability analysis of the AMD network. In the switch fault model, an SE is totally unusable if it becomes faulty.

A *redundancy graph* [14] of a multipath MIN is used to describe all the available paths between a given source-destination pair. A node on the path may be the source or the destination or an SE. A redundancy graph for a $2^5 \times 2^5$ AMD network is shown in (Fig. 6).



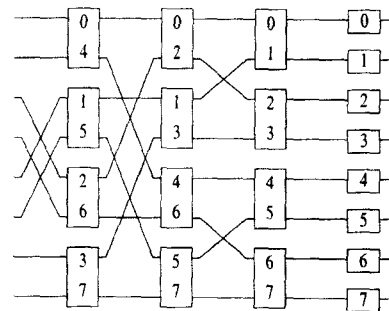
(Fig. 6) A redundancy graph for a $2^5 \times 2^5$ AMD network

Note that the redundancy graph in (Fig. 6) is exactly the same as that of the Kappa network because it shows a set of different output-input connections of SEs between adjacent stages as a single edge. This fact gives us an intuition that many properties regarding the performance and the reliability are shared between these two networks.

From (Fig. 6), we see that any one of the conjugate SEs at stage i ($1 \leq i \leq n-1$) can be used to provide a path to the destination. Therefore it is straightfor-

ward to identify that the AMD network is a *single fault tolerant MIN* if we assume that the input stage and the output stage are robust and fault-free. This assumption has been used in [15], and will be used throughout our analysis. It is also apparent that there are 2^n ($n = \log_2 N$) paths between every source-destination pair.

Our analysis of the MTTF for the AMD networks is based on the simplified switch fault model and thus provides an upperbound of the average time that has passed before the network fails in view of full access capability. We assume that only failures on any conjugate pair of SEs in the interior stages lead to a network failure. An 8×8 AMD network is redrawn in (Fig. 7) to show that the MTTF can be derived in the same way as the Delta network. In (Fig. 7), a rectangular box denotes a pair of conjugate SEs, and an edge from a box in stage i to a box in stage $i+1$ symbolizes the existence of a link from each SE in stage i to all SEs in stage $i+1$, one along the primary link and the other along the conjugate link. Then under the switch fault model, we can treat the AMD network as if it were a UPP MIN each of whose SEs is replaced by a pair of conjugate SEs. A pair of conjugate SEs can be modeled as a parallel-redundant system.



(Fig. 7) Redrawn 8×8 AMD network for MTTF analysis

1) The capability of a network to provide a connection from any of its inputs to any one of its outputs.

Let $R(t)$ be a reliability function of a system, representing the probability that a system will survive in the time interval $(0, t)$. Let the SE failure occur in the Poisson process with the rate of λ . Then the reliability of an SE, $R_{SE}(t)$, is given by $R_{SE}(t) = e^{-\lambda t}$. Then we have a reliability of a pair of conjugate SEs from $R_{CSE}(t) = 1 - (1 - e^{-\lambda t})^2$. Since there are $\frac{N}{2} \cdot (\log_2 N - 1)$ pairs of relevant conjugate SEs, we can get the reliability of the AMD network by $R_{AMD}(t) = [1 - (1 - e^{-\lambda t})^2]^P$ where $P = \frac{N}{2} \cdot (\log_2 N - 1)$. The MTTF of the AMD network can be obtained by $MTTF = \int_0^\infty [1 - (1 - e^{-\lambda t})^2]^P dt$. <Table 2> lists the MTTF of the AMD network and other MINs with differing network sizes.

<Table 2> Mean time to failure (MTTF)

Network Size	UPP	AMD	Hybrid	ACN
8	$0.0833\lambda^{-1}$	$0.3808\lambda^{-1}$	$0.5821\lambda^{-1}$	$0.4627\lambda^{-1}$
16	$0.0313\lambda^{-1}$	$0.2027\lambda^{-1}$	$0.3002\lambda^{-1}$	$0.2545\lambda^{-1}$
32	$0.0125\lambda^{-1}$	$0.1188\lambda^{-1}$	$0.1729\lambda^{-1}$	$0.1531\lambda^{-1}$
64	$0.0052\lambda^{-1}$	$0.0732\lambda^{-1}$	$0.1055\lambda^{-1}$	$0.0958\lambda^{-1}$
128	$0.0022\lambda^{-1}$	$0.0465\lambda^{-1}$	$0.0666\lambda^{-1}$	$0.0615\lambda^{-1}$
256	$0.0010\lambda^{-1}$	$0.0302\lambda^{-1}$	$0.0430\lambda^{-1}$	$0.0402\lambda^{-1}$
512	$0.0004\lambda^{-1}$	$0.0198\lambda^{-1}$	$0.0282\lambda^{-1}$	$0.0265\lambda^{-1}$
1024	$0.0002\lambda^{-1}$	$0.0132\lambda^{-1}$	$0.0187\lambda^{-1}$	$0.0177\lambda^{-1}$

The terminal reliability of a given source-destination pair (S, D) of an $N \times N$ MIN with n stages can be obtained recursively using the following notation.

$p(s)$: the probability that an SE in the MIN is not faulty.

$R(i)$: the probability that at least one fault-free path is available from an SE_{i-1} at stage $i-1$, to the destination D , given that SE_{i-1} is fault-free.

Then the terminal reliability is given by $R(1)$. As is the usual case, the output stage is assumed to be fault-free, so that $R(n) = 1$.

Again from (Fig. 6), we have the following recurrence relations for the AMD network: $R(i) = p(s)R(i+1) + (1 - p(s))p(s)R(i+1)$.

The probability $R(1)$ denotes the terminal reliability of the AMD network. Terminal reliability of the AMD network and other representative networks are listed in <Table 3>.

<Table 3> Terminal reliability ($p(s) = 0.9$)

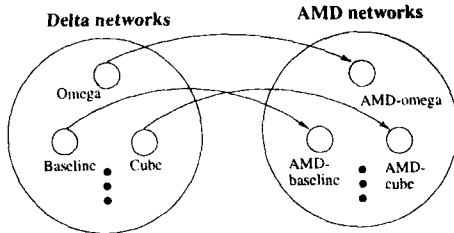
Network Size	Delta	Gamma	AMD	Hybrid	ACN
16	0.656	0.859	0.970	0.953	0.961
32	0.590	0.844	0.961	0.935	0.951
64	0.531	0.830	0.951	0.917	0.942
128	0.478	0.815	0.942	0.899	0.932
256	0.430	0.801	0.932	0.882	0.923
512	0.387	0.787	0.923	0.866	0.914
1024	0.345	0.773	0.914	0.849	0.904
2048	0.314	0.759	0.904	0.833	0.895
4096	0.282	0.746	0.895	0.817	0.886

6. Conclusion

In this paper, we proposed a class of multipath MINs, called the Augmented Modified Delta (AMD) network, which is well suited to real-time, large-scale parallel/distributed processing environments because of its fault-tolerant capability, simple distributed control scheme, modular structure, and high performance characteristics. The design philosophy of the AMD network emphasizes the modularity and regularity in the network design procedure. Thus the design of the AMD network is very straightforward and consistent as compared with that of the Kappa network. The AMD network maintains the superior performance and reliability of the Kappa network. Moreover, all the AMD networks are topologically equivalent, and thus they can simulate the behavior of the other AMD networks with the same method used in the Delta network. As shown in (Fig. 8), the AMD network can be thought of a natural fault-tolerant extension to the Delta network.

The same link connection pattern as the AMD network is adopted by a network, called the ACN (Augmented C-Network) [17]. But there is a significant

ant difference between these two networks in terms of hardware complexity, and the performance of the AMD network is better than that of the ACN.



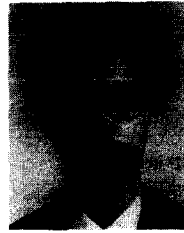
(Fig. 8) Relationship between Delta networks and AMD networks

References

- [1] M. Flynn, "Computer architecture: Pipelined and parallel processor design," Jones and Bartlett Pubs., 1995.
- [2] V. Kumar, A. Grama, A. Gupta and G. Karypis, "Introduction to Parallel Computing: Design and analysis of algorithm," Benjamin/Cummings Pub., 1994.
- [3] T. von Eicken, D. E. Culler, S. C. Goldstein and K. E. Schauer, "Active messages: A mechanism for integrated communication and computation," *Proc. of the 19th Int'l Symposium on Computer Architecture*, May 1992.
- [4] J. Duato, S. Yalamanchili and L. Ni, "Interconnection networks: An engineering approach," *book draft*, Jan. 1997.
- [5] S. C. Kothari, "Multistage interconnection networks for multiprocessor systems," *Advances in Computers*, Academic Press: New York, 1986.
- [6] L. N. Bhuyan, "Interconnection networks for parallel and distributed processing," *Computer*, vol. 20, no. 6, pp. 9-12, June 1987.
- [7] G. B. Adams, III, D. P. Agrawal and H. J. Siegel, "A survey and comparison of fault-tolerant multistage interconnection networks," *Computer*, vol. 20, no. 6, pp. 14-27, June 1987.
- [8] H. J. Siegel, "Interconnection networks for large-scale parallel processing," 2nd ed., McGraw-Hill, 1990.
- [9] T. Hanawa, H. Adamo, and Y. Fujikawa, "Multi-stage interconnection networks with multiple outlets," *Proc. of the 1994 Int'l Conf. on Parallel Processing*, vol. 1, pp. 1-8, Aug., 1994.
- [10] C. Wu and T. Feng, "On a class of multistage interconnection networks," *IEEE Trans. Computer*, vol. C-29, no. 8, pp. 694-702, Aug. 1980.
- [11] D. H. Lawrie, "Access and alignment of data in an array processor," *IEEE Trans. on Computer*, vol. C-24, no. 12, pp. 1145-1155, Dec. 1975.
- [12] H. J. Siegel and R. J. McMillen, "The multistage cube: a versatile interconnection network," *Computer*, vol. 14, no. 12, pp. 65-76, Dec. 1981.
- [13] L. R. Goke and G. J. Lipovski, "Banyan networks for partitioning multiprocessor systems," *1st Int'l Symp. Computer Architecture*, pp. 21-28, Dec. 1973.
- [14] C. S. Raghavendra and D. S. Parker, "Reliability analysis of an interconnection network," *4th Int'l Conf. Distributed Computing Systems*, pp. 461-471, May 1984.
- [15] S. C. Kothari, G. M. Prabhu, and R. Roberts, "The Kappa network with fault-tolerant destination tag algorithm," *IEEE Trans. on Computer*, vol. 37, no. 5, pp. 612-617, May 1988.
- [16] D. S. Parker and C. S. Raghavendra, "The Gamma network: a multiprocessor interconnection network with redundant paths," *9th Int'l Symp. Computer Architecture*, pp. 73-80, May 1982.
- [17] S. M. Reddy and V. P. Kumar, "On fault-tolerant interconnection networks," *Int'l Conf. Parallel Processing*, pp. 155-164, Aug. 1984.
- [18] D. P. Agrawal, "Graph theoretical analysis and design of multistage interconnection networks," *IEEE Trans. Computer*, vol. C-32, no. 7, pp. 637-648, July 1983.
- [19] J. Kim, "Performance study of packet switching multistage interconnection networks," *ETRI*

Journal, vol. 16, no. 3, pp. 27-41, Oct. 1994.

- [20] V. P. Kumar and S. M. Reddy, "Design and analysis of fault-tolerant multistage interconnection networks with low link complexity," *12th Int'l Symp. Computer Architecture*, pp. 376-386, June 1985.



김 정 선

1986년 서울대학교 컴퓨터공학과 졸업(학사)

1988년 Iowa State University 전기 및 컴퓨터공학과 졸업(공학석사)

1994년 Iowa State University 전기 및 컴퓨터공학과 졸업(공학박사)

1994년~1996년 한국전자통신연구원 (ETRI) 선임 연구원

1996년~현재 한양대학교 전자계산학과 전임강사

관심분야: Interconnection networks, Parallel programming environments, Distributed object-oriented systems