3-Disjoint gamma interconnection networks

Ching-Wen Chen [a,*], Neng-Pin Lu [b], Chung-Ping Chung [c]

[a] *Department of Information and Communication Engineering, Chaoyang University of Technology, Wufeng 413, Taichung County, Taiwan, ROC*
[b] *Department of Information Management, Chang Gung University, Kwei-Shan, Tao-Yuan 333, Taiwan, ROC*
[c] *Department of Computer Science and Information Engineering, National Chiao Tung University, Hsinchu 300, Taiwan, ROC*

Abstract

In this paper, we propose a new multistage interconnection network, called 3-disjoint gamma interconnection network (3DGIN). The 3DGIN is a modified gamma interconnection network that provides 3-disjoint paths to tolerate two switch or link faults between any source and destination pairs. The 3DGIN has lower hardware cost than GIN; furthermore, the routing and rerouting tags to generate 3-disjoint paths can be obtained in $O(\log N)$ time. To show the advantage features of 3DGIN, we also make a comparison between the gamma-related networks, the GIN, enhanced IADM, and 3DGIN.
© 2002 Elsevier Science Inc. All rights reserved.

1. Introduction

Interconnection networks are critical to parallel systems because their performances are closely related to the latency and throughput of the network. Multistage interconnection networks are well suitable for communication systems in terms of their tightly coupled components and can offer a good balance between cost and performance. For a complex system, assuring its high reliability is a challenging task. Therefore, in regard to the large-scale multiprocessor systems, fault tolerance is of crucial importance in term of fulfilling the communication needs for MINs (Adams et al., 1987).

To enhance the capability of fault tolerance, gamma interconnection network (GIN) (Parker and Raghavendra, 1984) provides multiple paths between any source and destination pairs except that if the source and destination are identical. Furthermore, to improve the fault-tolerant capability of GIN, several schemes have been introduced, such as extra stage gamma network (Yoon and Hegazy, 1988), CGIN (Chuang, 1996), composite banyan (Seo and Feng, 1995), PCGIN, FCGIN (Chen et al., 2000), B-network (Lee and Yoon, 1990) and enhanced IADM (McMillen and Siegel, 1982). Among them, extra stage gamma, CGIN, composite banyan network, FCGIN and PCGIN can provide at least 2-disjoint paths. On the other hand, FCGIN and B-network use dynamic rerouting to tolerate faults, but B-network cannot guarantee one-fault tolerant whereas FCGIN is one-fault tolerant. Although enhanced IADM has the capability to tolerate two-faults, it needs higher hardware cost, one stage lookahead technique and dynamic rerouting (McMillen and Siegel, 1982).

In this paper, we propose a new 3-disjoint paths network; namely, 3-disjoint GIN (3DGIN). This 3DGIN can also tolerate two links or switch faults, and its hardware cost is almost equal to GIN. The remainder of this paper is organized as follows. In Section 2, we shall introduce the GIN, and routing scheme in GIN. In Section 3, we shall present the 3DGIN, and the routing algorithm. In Section 4, the comparison between gamma-related networks and how to use disjoint paths to reduce hot spots are discussed. Finally, Section 5 concludes this paper.

---

* Corresponding author.
E-mail addresses: chingwen@mail.cyut.edu.tw (C.-W. Chen), nplu@mail.cgu.edu.tw (N.-P. Lu), cpchung@csie.nctu.edu.tw (C.-P. Chung).

2. Gamma interconnection network and enhanced IADM

2.1. Topology of gamma interconnection network

A GIN of size $N = 2^n$ has $n + 1$ stages being labeled from 0 to n and each stage involves N switches (Parker and Raghavendra, 1984). Basically, switches of sizes 1×3 and 3×1 are coupled with the first and last stage respectively. Moreover, each switch located at intermediate stages is a 3×3 crossbar. And each switch number j at stage i has three output links connecting to switches at stage $(i + 1)$ according to the plus-minus-2^i function. In other words, the jth switch at stage i has three output links to switches $[(j - 2^i) \bmod N]$, j, and $[(j + 2^i) \bmod N]$ at each consecutive stage. Fig. 1 illustrates a GIN network with size 8.

In GIN, an n-digit tag determines the path connecting the source to its destination. Each tag digit can be 1, 0, or $\bar{1}$. An n-digit tag T represents the difference between D and S, i.e., $T = (D - S) \bmod N$. Digit t_i is used at stage i in such a way that the lower connection is selected when t_i is equal to 1, and the straight connection is selected when t_i is 0, where the distance $T = t_0 t_1 \ldots t_{n-1}$. Moreover, a non-zero tag T has multiple representations, that is, there are multiple paths between source S and destination D if $S \neq D$. For example, in the case when $N = 8$, source node S is 2, and destination node D is 0, the tag T can be $0\,1\,1$, $0\,1\,\bar{1}$ or $0\,\bar{1}\,0$ as shown in Fig. 1.
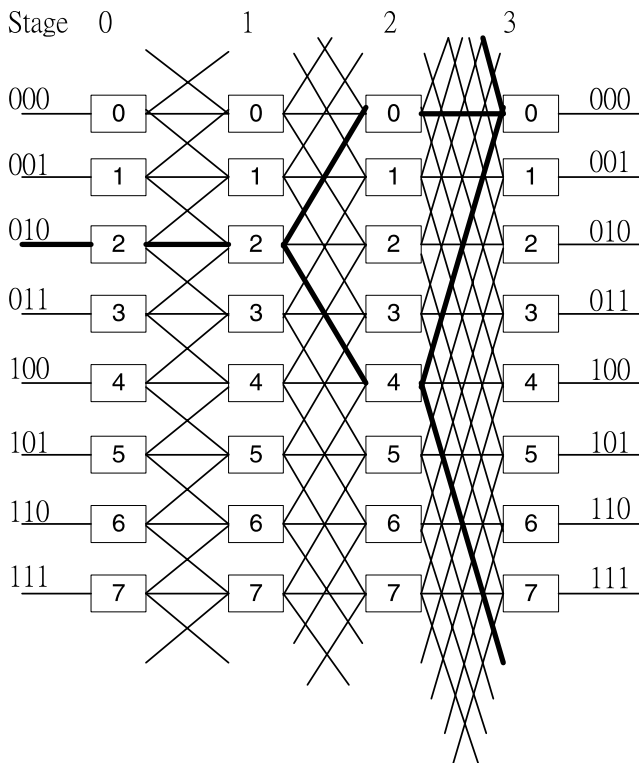
2.2. Topology of enhanced IADM

The enhanced IADM (McMillen and Siegel, 1982) is a revised version of IADM with topology equivalent to GIN. There are two frameworks of enhanced IADM. One of them is designed to provide redundant straight links that allows fault links to be avoided by using the second straight link, but not the switch faults. The other arrangement is to add half links to each stage from stage 1 to $n - 1$. Half links are used to connect a switch m at stage i with switch $(m + 2^{i-1}) \bmod N$ and $(m - 2^{i-1}) \bmod N$ at stage $i + 1$ as shown in Fig. 2. Adding half links provides single-fault tolerance to any intermediate switch or link fault since there exist at least two links for distinct switches to connect to each other at the successive stage, and each one can be used to satisfy the routing requirement. However, the network cannot be two-fault tolerant unless a single-stage look-ahead technology and dynamic routing are engaged. For example, in the present case when $S = 6$ and $D = 6$, and the routing path may be $6 \rightarrow 6 \rightarrow 6 \rightarrow 6$, however, as the straight link from stage 1 to stage 2 is faulty, the path can be $6 \rightarrow 6 \rightarrow 4 \rightarrow 6$. If the link from switch 4 to switch 6 is faulty (the second fault), the routing algorithm must use the look-ahead technique to overlook one stage further to get the information from switch 4 at stage 2. Consequently, the algorithm will take the link to switch 0 at stage 2. As a result, the path will change to $6 \rightarrow 6 \rightarrow 0 \rightarrow 6$. This example is illustrated in Fig. 2.

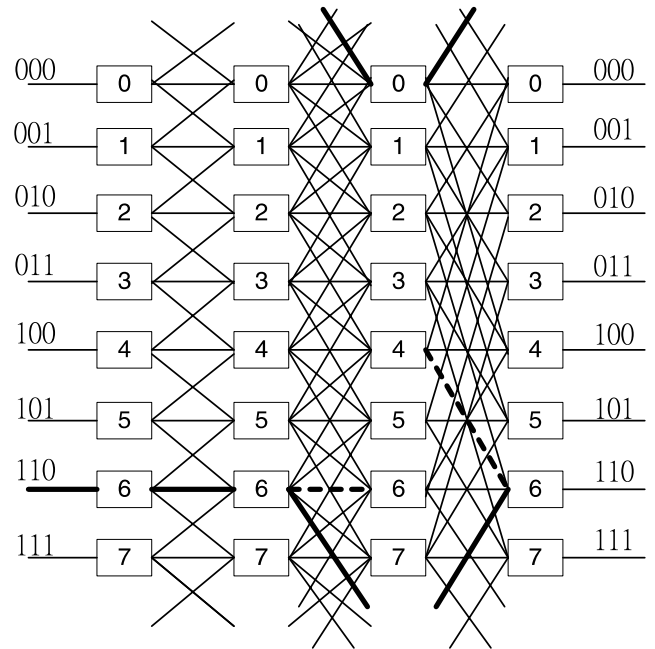Fig. 1. GIN with $N = 8$ and its three paths between nodes 2 and 0.

Fig. 2. The routing condition in enhanced IADM with $S = 6$, $D = 6$ and two links faults. The dash lines mean faulty links.

3. 3-Disjoint paths gamma interconnection network

Enhanced IADM, although it is two-fault tolerant, needs higher hardware cost, one stage look-ahead technique, and dynamic rerouting capability to complete the operation of networking. In this section we demonstrate the proposed 3-disjoint paths GIN (3DGIN), which is comparable to GIN in terms of the lower hardware cost and two-fault tolerant capability.

3.1. Topology

A 3DGIN, 3-disjoint paths GIN of size $N = 2^n$, is similar to GIN in many respects except that switches $2i$ and $2i + 1$ are combined into one 2×4 switch at stage 0 with the rest of stages being the same as GIN. Fig. 4 shows the 3DGIN with $N = 8$. The naming scheme at stage 0 is described as follows. The four associated links are named to be $0\,0$, $0\,1$, $1\,0$, and $1\,1$ as shown in Fig. 3. And the links at other stages are denoted as GIN.

Theorem 1. *There are 3-disjoint paths between any source and destination pair in 3DGIN.*

Proof. There are two sections in this proof: one of them is $(D - S) \bmod 2 = 0$, and the other one is $(D - S) \bmod 2 = 1$.

(a) In the part of $(D - S) \bmod 2 = 0$, let us consider a path which automatically takes the straight link from switch S at stage 0 to switch S at stage 1 in GIN because $(D - S) \bmod 2 = 0$.

1. If $|D - S| \neq N/2$, there is at least one path P in GIN routing through the straight link between final two stages since $|D - S|$ or $(N - |S - D|)$ is less than $N/2$. We apply this path P to the proposed 3DGIN. In this regard, the remaining 2-disjoint paths will traverse the switch $[(i + 1) \bmod n]$ and $[(i - 1) \bmod n]$ if the path P traverse the switch i between stage 1 to stage $n - 1$, whereas the first and final stages are (plus and minus)/(minus and plus) one as shown in Fig. 4.
2. If $|D - S| = N/2$ and S is odd/even; where the slash indicates a state of either the former or the latter within
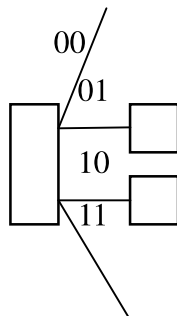

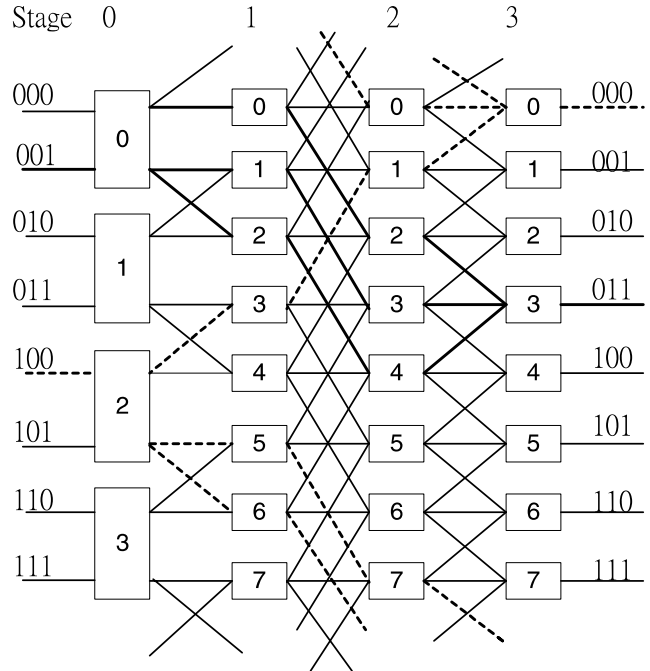Fig. 3. The naming scheme of links at stage 0 in 3DGIN.

Fig. 4. 3-Disjoint paths GIN with size 8. (Two routing conditions in 3DGIN. Solid lines mean $(D - S) \bmod 2 \neq N/2$, and dash lines mean $(D - S) \bmod 2 = N/2$.)

a time domain. Let P_1, P_2 and P_3 be the 3-disjoint paths taking the links $(0\,0,\ 0\,1,\ 1\,1)/(0\,0,\ 1\,0,\ 1\,1)$ at stage 0. From stage 1 to destination, $(P_2, P_3)/(P_1, P_2)$ always take non-straight upward and downward links respectively. However, P_1/P_2 is paralleling P_2/P_3 except for the stage $n - 1$ to stage n by straight link. P_2/P_1 and P_3/P_2 go through different switches because the sum of the vertical length to stage $n - 1$ is less than $N - 2$. Finally, two non-straight links are taken to destination as shown in Fig. 4.

(b) When $(D - S) \bmod 2 = 1$ and S is even/odd, the 3-disjoint paths are the same as those of $(S + 1)/(S - 1)$ to D, because the input of S and $(S + 1)/(S - 1)$ are combined in the same switches.   □

There are 3-disjoint paths in 3DGIN between any source and destination pair.

3.2. Routing and rerouting tags

In this section, we shall demonstrate the algorithm to compute routing and rerouting tags. The algorithm needs source and destination tags as input to produce three routing tags. Because there are four links from stage 0 to stage 1, the binary representation of the routing tags, Ta, Tb, and Tc, occupy $(n + 1)$ bits. Also, the notation of the first two bits is T_{x0} and T_{x1} and of other bits is $T_{x2}, T_{x3}, \ldots, T_{xn}$. Let us discuss the conditions for generating routing and rerouting tags.

1. Input source S and destination D.
2. Calculate the distance T. Because the distance $T = (D - S)$ must be less than $N/2$.
   (a) The distance $T = (D - S) \bmod N$ if $((D - S) \bmod N) < N/2$.
   (b) The distance $T = N - ((D - S) \bmod N)$ if $((D - S) \bmod N) > N/2$.
3. Find the first routing tag Ta under T is not equal to $N/2$ and we do not care stage 0 to 1 because the link needs 2 bits.
   (a) $T = N/2 - 1$,
      (i) If S is odd, the path goes up-direction or straight links only. We can find the path from $(S - 1) \bmod N$ to D and the bits of tag Ta are composed of 0 or $\bar{1}$ only.
      (ii) If S is even, the path goes down-direction or straight links only. We can find the path from $(S + 1) \bmod N$ to D and the bits of tag Ta are composed of 0 or 1 only.
   (b) $T < N/2 - 1$,
      (i) If T is even and the direction is down-ward/up-ward, the Ta is calculated by T only by 0 or $1/\bar{1}$.
      (ii) If T is odd, we calculate the distance $(S - 1)/(S + 1)$ to D if S is odd/even.
   (c) From the result of procedure 3(a) and 3(b), the first and last links take the straight, that is, the first two bits of Ta are 0 1 or 1 0 and the last bit of Ta is 0.
4. Find two more rerouting tags Tb and Tc
   (a) The bits of Tb and Tc are the same as Ta between bit 1 to bit $n - 1$, and the first two bits are 0 0/0 1 and 1 0/1 1 if the first two bits of Ta is 0 1/1 0, and last bits are $\bar{1}$ and 1.
5. Find Ta, Tb and Tc under $T = N/2$
   (a) If ($T = N/2 - 1$ and S is odd and direction is down), we let $S = S - 1$ because the inputs of S and $S - 1$ connect the same switch at stage 0.
   (b) If ($T = N/2 - 1$ and S is even and direction is up), we let $S = S + 1$ because the inputs of S and $S + 1$ connect the same switch at stage 0.
   (c) If $T = N/2$ (the condition is the same as previous two after procedure (d) and (e) are processed), Ta and Tb could be $00\bar{1}\bar{1}\ldots\bar{1}/01\bar{1}\bar{1}\ldots\bar{1}$ and $10111\ldots1/11111\ldots1$ respectively. And the third tag Tc is $1111\ldots110/00\bar{1}\bar{1}\bar{1}\ldots0$.

To be specific, Algorithm 1 describes the details of this routing algorithm. Some examples are given below to provide the methods of calculation. Three examples of 3DGIN are illustrated for comparison; and their expressions are listed as follows:

1. $(S - D) \bmod N < N/2$,
2. $(S - D) \bmod N = N/2$,
3. $(S - D) \bmod N > N/2$.

The first and third condition are similar, we just present the instance for condition 1. For example, if $S = 2$, $D = 3/4$, and $N = 8$, the routing tags can be derived as $0110/1000$, $0011/0101$, and $101\bar{1}/110\bar{1}$ after the distance $T = 2/1$ is calculated, as shown in Fig. 5. In the second example: if $S = 4$, $D = 0/1$, and $N = 8$; the routing tags can be derived as $00\bar{1}\bar{1}/00\bar{1}0$, $1011/01\bar{1}$, and $1110/1111$ after the distance $T = 4/5$ is calculated, as shown in Fig. 5.

Algorithm 1. Calculate the routing and rerouting tags of 3DGIN

Input: Source tag $S = s_0 s_1 s_2 \ldots s_{n-1}$, destination node $D = d_0 d_1 d_2 \ldots d_{n-1}$

Output: Routing tags $\text{Ta} = \text{ta}_{00}\text{ta}_{01}\text{ta}_1\text{ta}_2 \ldots \text{ta}_{n-1}$, $\text{Tb} = \text{tb}_{00}\text{tb}_{01}\text{tb}_1\text{tb}_2 \ldots \text{tb}_{n-1}$, $\text{Tc} = \text{tc}_{00}\text{tc}_{01}\text{tc}_1\text{tc}_2 \ldots \text{tc}_{n-1}$

```
Begin
    Up = False;
    T = (D − S) mod N;
    If T > N/2 then
        T = N − T;
        Up = True;
    End If
    If T < N/2 − 1 or (T = N/2 − 1 and S is odd and
    Up = True) or (T = N/2 − 1 and S is even and
    Up = False)
        then
        If (T is even and S is even) or (T is odd and S is
        odd) then
```
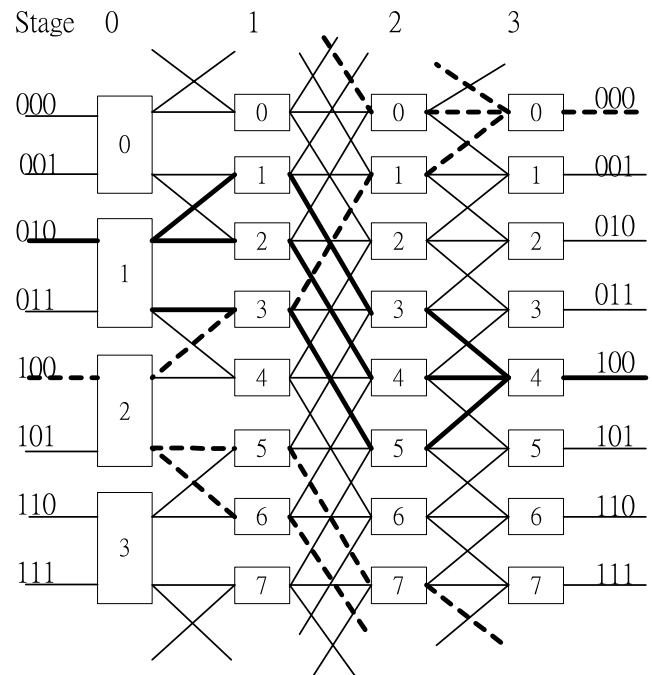

Fig. 5. The 3-disjoint paths in 3DGIN with $S = 2$ and $D = 4$ for the $D - S < N/2$ case with bold line and $S = 4$ and $D = 0$ for the $D - S < N/2$ case with bold dash line.

$Ta_{00} = 01$; $Tb_{00} = 00$; $Tc_{00} = 10$;
Else
$\quad$ $Ta_{00} = 10$; $Tb_{00} = 01$; $Tc_{00} = 11$;
End If
$T = T/2$; /* take the integer part */
For $i = 1$ to $n - 2$
$\quad$ $Ta_i = T\%2$; /*get the remainder*/
$\quad$ If Up = True then $Ta_i = \overline{Ta_i}$;
$\quad$ $T = T/2$;
$\quad$ $Tb_i = Ta_i$;
$\quad$ $Tc_i = Ta_i$;
End for
$Ta_{n-1} = 0$; $Tb_{n-1} = 1$; $Tc_{n-1} = \bar{1}$
Else
$\quad$ If ($T = N/2$ and S is even) or ($T = N/2 - 1$ and S is odd) then
$\quad$ $\quad$ $Ta_{00} = 00$; $Tb_{00} = 10$; $Tc_{00} = 11$;
$\quad$ $\quad$ For $i = 1$ to $n - 2$
$\quad$ $\quad$ $\quad$ $Ta_i = \bar{1}$; $Tb_i = 1$; $Tc_i = 1$;
$\quad$ $\quad$ End For
$\quad$ $\quad$ $Ta_{n-1} = \bar{1}$; $Tb_{n-1} = 1$; $Tc_{n-1} = 0$;
$\quad$ Else
$\quad$ $\quad$ $Ta_{00} = 00$; $Tb_{00} = 01$; $Tc_{00} = 11$;
$\quad$ $\quad$ For $i = 1$ to $n - 2$
$\quad$ $\quad$ $\quad$ $Ta_i = \bar{1}$; $Tb_i = 1$; $Tc_i = 1$;
$\quad$ $\quad$ End For
$\quad$ $\quad$ $Ta_{n-1} = \bar{1}$; $Tb_{n-1} = 1$; $Tc_{n-1} = 0$;
$\quad$ End If
End If
End

4. Discussion

3DGIN provides 3-disjoint paths for two-fault tolerant and the same network latency $O(\log_2 N)$ in light traffic under no fault occurring. When a fault occurs, a packet backtracks to source node and takes another path to destination by rerouting tag. However, the network latency will be

$$\log_2 N + \frac{1}{3n} \sum_{i=0}^{n-1} 2i = \frac{n-1}{3} + \log_2 N,$$

but GIN will be $\log_2 N$ or ∞ because GIN has no alternative path when straight link is fault shown in Table 1.

Although 3DGIN has the characteristics of 3-disjoint paths and the same network latency $O(\log_2 N)$ in light traffic under no fault occurs and lower hardware cost than GIN, its feature of redundant links is lost. However, the redundant links in GIN can help a packet go either upward or downward as it encounters a non-straight link (Rau et al., 1992). For example, $S = 2$, $D = 0$, and $N = 8$, the packet can go up-direction, $2 \rightarrow 2 \rightarrow 0 \rightarrow 0$, or go down-direction, $2 \rightarrow 2 \rightarrow 4 \rightarrow 0$, because of the redundant properties of GIN as shown in Fig. 1. If the property of redundancy is needed, a plus/minus-2^{n-1} link can be added to final stage and the hardware cost will increase. Yet this brings the following additional advantage: Alternative path when a non-straight link is taken. In our design, we choose not to have the feature though.

Additionally, more disjoint paths can reduce the effects of hot spots in a MIN (Wang et al., 1995; Chuang and Tu, 1999). The reducing hot spot scheme is described as follows:

1. The packets can be synchronous or asynchronous and with hot spots or not.
2. We find 3-disjoint paths by the preceding procedure.
3. When tree saturation occurs, the routing path is selected according to the request pattern (synchronous or asynchronous) and the current network traffic (with or without hot spots).
(A) If the pending request is a synchronous request, we set the routing to use the upper routing path.
(B) For an asynchronous request, the routing path should be chosen following the current network traffic.
(I) If the destination of the pending asynchronous request is not a hot spot, the lower or middle path is chosen as the routing path and the upper path is released for the after pending synchronous request.
(II) If it is a hot spot, the upper path is chosen as the routing path and the middle and lower paths are reserved for the other pending asynchronous requests. (If the asynchronous

Table 1
The comparison of GIN, 3DGIN, and enhanced IADM

Network	Fault-tolerance method	Fault tolerant ability	Routing method	Hardware cost (total switch's crossing points)	Fault penalty
GIN	Multiple paths	Faults robust	Distance tag	$9N \log N - 3N$	0 or ∞
3DGIN	Disjoint paths	3-Disjoint paths	Distance tag and rerouting tags	$9N \log N - 6N$	$(n-1)/3$
Enhanced IADM	Dynamic rerouting and look-ahead one stage method	3-Disjoint paths	Distance tag with look-ahead one stage information	$25N \log N - 27N$	NA

request still traverses the network by the middle or lower path, the path will be blocked since the target port is a hot spot.)

As a result, if the asynchronous transmission constitutes the majority of communication in the network traffic, such a static routing scheme will be able to reduce average delay time.

5. Conclusion

A new multipath multistage interconnection network called 3DGIN is proposed in this paper. The 3DGIN can provide 3-disjoint paths between any source and destination pairs by converting two switches into one at stage 0. In addition, without one stage look-ahead technique, 3DGIN still offers the capabilities of two-faults tolerance, less hardware cost, and 3-disjoint paths, as compared with enhanced IADM. Table 1 summarizes the comparison of GIN, 3DGIN and enhanced IADM. In conclusion, 3DGIN distinct from other schemes has been demonstrated in this work with innovative features, including low hardware cost, 3-disjoint paths capability, and an $O(\log N)$ algorithm of getting routing and rerouting tags.

References

Adams III, G.B., Agrawal, D.P., Siegel, H.J., 1987. A survey and comparison of fault-tolerant multistage interconnection networks. IEEE Transactions on Computer 20 (6), 14–27.

Chen, C.W., Lu, N.P., Chen, T.F., Chung, C.P., 2000. Fault-tolerant gamma interconnection networks by chaining. IEE Proceedings of Computers and Digital Techniques 147 (2), 75–81.

Chuang, P.J., 1996. CGINs: A fault tolerant modified gamma interconnection network. IEEE Transactions on Parallel and Distributed Systems 7 (12), 1301–1306.

Chuang, P.J., Tu, H.Y., 1999. Dynamic scheme for reducing hot-spot effects in multipath networks. IEE Proceedings of Computers and Digital Techniques 146 (4), 179–184.

Lee, K.Y., Yoon, H., 1990. The B-network: A multistage interconnection network with backward links. IEEE Transactions on Computers 39 (7), 966–969.

McMillen, R.J., Siegel, H.J., 1982. Performance and fault tolerance improvements in the inverse augmented data manipulator network. In: 9th Symposium on Computer Architecture, pp. 63–72.

Parker, D.S., Raghavendra, C.S., 1984. The gamma network. IEEE Transactions on Computers C-33, 367–373.

Rau, D., Fortes, J.A.B., Siegel, H.J., 1992. Destination tag routing techniques based on a state model for the IADM network. IEEE Transactions on Computers 41 (3), 274–285.

Seo, S.W., Feng, T.Y., 1995. The composite banyan network. IEEE Transactions on Parallel and Distributed Systems 6 (10), 1043–1054.

Wang, M.C., Siegel, H.J., Nichols, M.A., Abraham, S., 1995. Using a multipath network for reducing the effects of hot spots. IEEE Transaction on Parallel and Distributed Systems 6 (3), 252–268.

Yoon, K., Hegazy, W., 1988. The extra stage gamma network. IEEE Transactions on Computers 37 (11), 1445–1450.

Ching-Wen Chen is an assistant professor of Department of Information and Communication Engineering, Chaoyang University of Technology, Wufeng, Taichung County, Taiwan, Republic of China. He received the B.E. degree in information engineering and computer science from Feng Chia University, Taiwan, Republic of China in 1993, and the M.E. degree in Department of Computer Science from National Tsing Hwa University, Taiwan, Republic of China in 1995 and the Ph.D. degree in Department of Computer Science and Information Engineering from Chiao Tung University, Taiwan, Republic of China in 2002. His research interests include computer architecture, interconnection network and parallel processing.

Neng-Pin Lu is an assistant professor of Department of Information Management, Chang Gung University, Kwei-Shan, Tao-Yuan, Taiwan, Republic of China. Before joined Chang Gung University, he has been an assistant processor of Chihlee Institute of Commerce, Panchiao, Taipei, Taiwan, Republic of China. He received the B.E., M.E., and Ph.D. degrees of Computer Science and Information Engineering from the National Chiao Tung University, Hsinchu, Taiwan, Republic of China in 1989, 1991, and 2000, respectively. His research interests include computer architecture, interconnection network, parallel processing, and cluster computing.

Chung-Ping Chung received the B.E. degree from the National Cheng-Kung University, Taiwan, Republic of China in 1976, and the M.E. and Ph.D. degrees from the Texas A&M University in 1981 and 1986, respectively, all in electrical engineering. He was a lecturer in electrical engineering at the Texas A&M University while working towards the Ph.D. degree. Since 1986 he was been with computer science and information engineering at the National Chiao Tung University, Taiwan, Republic of China, where he is a professor. From 1991 to 1992, he was a visiting associate professor of computer science at Michigan State University. Currently, he is on leave and served as the director of Advanced Technology Center, Computer and Communications Research Laboratories, Industrial Technology Research Institute (CCL, ITRI), ROC, and then the consultant of CCL, ITRI. His research interests include computer architecture, parallel processing, and parallel compiler design.