

ATM 교환을 위한 비용 효율적인 동적 결함내성 bitonic sorting network

이 재 동[†] · 김 재 홍^{††} · 최 흥 인^{†††}

요 약

본 논문에서는 Batcher-Banyan 네트워크에 기반한 ATM 스위치를 설계하는데 사용되는 bitonic sorting 네트워크의 새로운 결함내성 기법을 제안한다. 본 논문의 최종 목표는 비용 효율적인 결함내성 bitonic sorting 네트워크의 설계에 있다. 여기서는 결함을 처리하기 위하여 추가적인 CE와 링크들이 사용된다. 동적 결함내성 bitonic sorting 네트워크는 동적 결함내성 네트워크에 기반하여 몇 가지 다른 변화에 의해 구축될 수 있다. 본 논문에서 제안한 결함내성 네트워크는 높은 결함내성, 적은 시간 지연, Cell sequence의 유지, 단순한 자기전달 및 규칙성회 모듈성을 제공한다.

A Cost-Effective Dynamic Redundant Bitonic Sorting Network for ATM Switching

Jae-Dong Lee[†] · Jae-Hong Kim^{††} · Hong-In Choi^{†††}

ABSTRACT

This paper proposes a new fault-tolerant technique for bitonic sorting networks which can be used for designing ATM switches based on Batcher-Banyan network. The main goal in this paper is to design a cost-effective fault-tolerant bitonic sorting network. In order to recover a fault, additional comparison elements and additional links are used. A Dynamic Redundant Bitonic Sorting (DRBS) network is based on the Dynamic Redundant network and can be constructed with several different variations. The proposed fault-tolerant sorting network offers high fault-tolerance; low time delays, maintenance of cell sequence, simple routing; and regularity and modularity.

1. Introduction

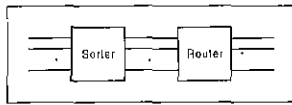
B-ISDN(Broadband Integrated Services Digital Network) as a common network for supporting a wide range of services such as voice, data and image have received much attention in the literature. Asynchronous Transfer Mode(ATM) is the transport technique

for the broadband ISDN recommended by CCITT[1]. Many switching networks have been proposed to accommodate the ATM, which requires fast packet switching [6]. Self-routing switches which apply a banyan network to the basic interconnections structures are suitable for fast switching because of their parallel processing ability [4-7].

As shown in (figure 1), a sorting network is used to sort the cells which are transported to the output ports through a router in ATM switching systems. A bitonic

† 정 회 원 단국대학교 전기계산학과 교수
†† 손 회 원 영동대학교 컴퓨터공학과 교수
††† 정 회 원 아코넥(주) 연구팀장
논문접수 : 1999년 7월 2일, 심사완료 : 1999년 9월 17일

sorting network which supports fast sorting capabilities and solves communication problems in ATM switching systems was introduced in 1968 [4]. Many ATM switches such as Starlite, Moonshine, SXmin and Sunshine use the bitonic sorting network which can be applied to simplify the design of an arbitration circuit structure and a router since the internal blocking is prevented [6, 7]



(Figure 1) A Sorter-Router ATM Switching System

Due to the large number of comparators(or CE) and links in Batcher-Banyan networks, failure of these components is common. A single fault in these networks can destroy their functionality. Therefore, many researchers have studied fault-tolerant sorting circuits, networks, and algorithms in various models. In 1985, Yao[12] has proposed a fault-tolerant sorting network in which a faulty comparator simply outputs its two inputs without comparison and use canonical redundancy method to fix the mis-routed data. Rudolph[8] has suggested that a fault tolerant sorting network can be achieved by reducing the crucial comparators and increasing the number of necessary pass in a recirculating shuffle-exchange block. Schimmer and Starko[9] have shown asymptotically optimal number of additional comparators and additional delay stages for single half(passive)-fault correction of arbitrary sorting network. Assaf and Upfal[3] have described a general method for converting any sorting circuit into a x (reversal)-fault-tolerant or destructive-fault-tolerant sorting network by replicating $\Theta(\log N)$ copies of each item. All these fault-tolerant sorting models require $\Omega(\log^2 N)$ depth and/or $\Omega(N \log^2 N)$ comparison elements to sort N items. Since many users share the same network, high availability in Batcher-Banyan networks is desirable. Therefore, fault-tolerance capability in these networks is an important consideration when designing

ATM switches.

Fault diagnosis and recovery techniques are crucial for verifying the chips and maintaining the reliability of the network. Fault diagnosis and recovery techniques for blocking networks have been well researched. However, direct application of these techniques to the design of ATM switch based on Batcher-Banyan network will not work since paths are influenced by the other inputs. Fault detection and recovery algorithms for the bitonic sorting network have also been studied recently. Not much study has been achieved on the design of fault-tolerant bitonic sorting networks. Therefore, a new fault recovery scheme for the bitonic sorting network is proposed in this paper.

The remaining of this paper is organized as follows. Section 2 reviews the bitonic sorting network. Section 3 presents a new fault tolerant bitonic sorting network, the Dynamic Redundant Bitonic Sorting Network (DRBSN), and its variations. Section 4 discusses an analysis of the hardware increase and section 5 states some conclusions.

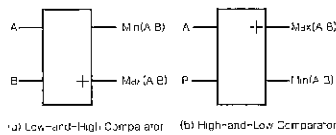
2 Bitonic Sorting Networks

In this section, the basic definitions which are used in this paper are described.

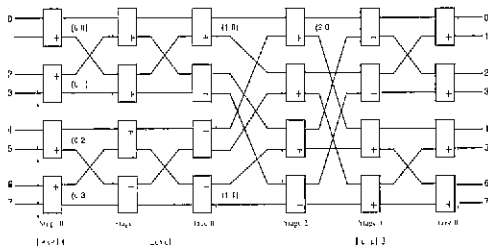
Batcher introduced a parallel sorting network which consists of recursively connected bitonic merge sorters in 1968. The fast sorting capability of this network allows its use in solving some problems where large sets of data must be manipulated. Some of these applications are a switching network with buffer, a multistage blocking network with a non-blocking feature, and a multi-access memory.

The key component of the bitonic sorting network is a comparison element (or comparator, for short) as shown in (figure 2). (Figure 3) shows an 8×8 bitonic sorting network. level 0 has four bitonic sorters, (0,0), (0,1), (0,2) and (0,3), where each bitonic sorter handles two elements; level 1 has two bitonic sorters, (1,0) and (1,1), where each bitonic sorter handles four elements; and level 2 has a single bitonic sorter which can handle

all eight elements. As shown in (figure 3), each bitonic sorter merges two sorted sequences into a full-size sorted sequence. Therefore, a bitonic parallel sorting network can be used to sort any sequence of elements by successively merging larger and larger bitonic sequences. Given $N(=2^k)$ unsorted elements, a network with $k(k+1)/2$ steps suffices. Each stage contains $N/2(=2^{k-1})$ comparators. Hence the total number of comparators is $2^{k-2} * k(k+1)$.



(Figure 2) Basic Comparison Element



(Figure 3) An 8x8 Bitonic Sorting Network

3 Fault Recovery In Bitonic Sorting networks

One drawback of the bitonic sorting networks is that they are not fault-tolerant in the sense that a single fault in the CE's or links destroys the functionality of the network. In 1990, Amano[2] has introduced the Fault Tolerant Batcher (FTB) network which can work even if an element in every bitonic sorter is faulted. In this scheme, the mis-routed output of each bitonic sorter can be corrected using a fault recovery mechanism which consists of multiplexers and inserters. However, the FTB network has the following problems. First, the increase of hardware is large - about 84%~42% depending on the size of input. Second, the performance of the

network is degraded since the error output of each bitonic sorter has to be fixed by the additional hardware. Finally, the FTB has no capability to tolerate a fault in the additional hardware.

In this section a new fault tolerant bitonic sorting network called the Dynamic Redundant Bitonic Sorting Network (DRBSN), is presented. Furthermore, some variations of this network will be discussed.

3.1 Design of a Dynamic Redundant Bitonic Sorting Network

The design of the DRBSN is based on the Dynamic Redundant (DR) Network which is introduced by Jeng and Siegel[10] in 1986. A CE that participates in the sorting task is called a functioning CE; otherwise it is called a spare CE. A spare CE is used when a faulty CE is detected and isolated. The structure of the Dynamic Redundant Bitonic Sorter is as follows:

- (1) Each comparison element (CE) is combined with input and output selectors (Figure 3).
- (2) Each stage of the bitonic sorter of $N(=2^n)$ inputs contains $N/2$ functioning comparison elements (CE) and δ indicates the spare rows of CEs. The number of actual extra CEs will be $\delta(i+1)$. For simplicity the value of δ will be $2^k(0 \leq k < 2^{n-2})$.
- (3) The levels of stages are labeled in a sequence from $n-1$ to 0 with 0 for the output level and $n-1$ for the input level.
- (4) A spare CE will become functioning when a faulty functioning CE is detected and isolated.
- (5) The number of links per CE and the connection scheme between stages are defined by the following:
 - The number of links per CE between stage k and $k-1$ ($0 < k < n$) is three. The three output links of a CE(j) ($0 \leq j < N/2 + \delta$) in stage k are connected to the input switch of the following CEs in stage $k-1$, where

$$C_k(j) = \begin{cases} j - (2^{k-1} \div \lfloor \frac{\delta}{2} \rfloor) \bmod (2^k + \delta) \\ (j + (2^{k-1} \div \lfloor \frac{\delta}{2} \rfloor)) \bmod (2^k + \delta) \end{cases}$$

The above connection scheme is applied recursively from stage $n-1$ through stage 1, where the value k represents the stage number.

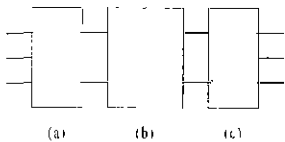
$$R_k(C_i(j), \delta) =$$

$$\left\{ \begin{array}{l} \text{if } \delta > 1 \\ \text{if } 0 \leq j < (2^{k-1} + \frac{\delta}{2}) \text{ and } k > 1 \\ \quad R_{k-1}(C_{i-1}(j), \frac{\delta}{2}) \\ \text{if } (2^{k-1} + \frac{\delta}{2}) \leq j < (2^k + \delta) \text{ and } k > 1 \\ \quad R_{k-1}(C_{i-1}(j - (2^{k-1} + \frac{\delta}{2})), \frac{\delta}{2}) + 2^{k-1} + (\frac{\delta}{2}) \\ \text{if } \delta \leq 1 \\ \text{if } 0 \leq j \leq 2^{k-1} \text{ and } k > 1 \\ \quad R_{k-1}(C_{i-1}(j), 0) \\ \text{if } 2^{k-1} \leq j \leq 2^k \text{ and } k > 1 \\ \quad R_{k-1}(C_{i-1}(j - 2^{k-1}), 0) + 2^{k-1} \end{array} \right.$$

$$\begin{aligned} q &= (2^k + \delta) / \delta \\ n &= (p \text{ div } q) * (q - 1) \\ m &= p \text{ mod } q \\ r &= j \text{ mod } q \\ l &= \begin{cases} m + n & \text{if } m \leq r \\ m + m - 1 & \text{if } m > r \end{cases} \end{aligned}$$

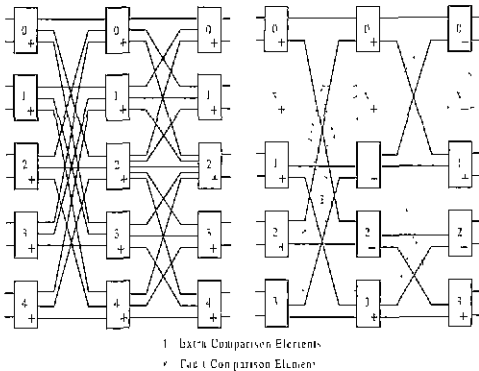
Here p and l represent the physical and logical numbering of each CE, respectively.

As shown in (figure 5) and from the above equations, CE (2^{k-1}) can be used as the upper or lower part of the cube connection depending on the location of the faulty CE. Therefore the maximum number of links per CE will be five between stage k and $k-1$ ($0 \leq k \leq n-2$).



(a) Input Selector (3 by 2)
(b) Original Comparison Element
(c) Output Selector (2 by 3)

(Figure 4) New Comparison Element



(Figure 5) Reconfiguration of 8x8 DRBS, $\delta = 1$

The control of the bitonic sorter is different from the DR network. As shown in (figure 5), if CE(j) or a link attached to CE(j) is found to be faulty, the system is reconfigured so that the physical number of CEs are re-numbered in the following way:

3.2 Implementation of a perfect shuffle connection in the DRBSN

As explained earlier, the bitonic sorting network is constructed by recursively merging two lower, ($i-1, 2j$) and ($j-1, 2j+1$), bitonic sorters into an upper, (i, j) bitonic sorter. This merging process can be viewed as a shuffle operation. By the definition of the DRBS, any number of additional CEs between 1 and 2^{l-1} can be used by each bitonic sorter. But, it is impossible to use the direct layout of a perfect shuffle connection for the DRBSN since the reconfiguration of any one of the bitonic sorter affects the other bitonic sorter. Therefore, the independent fault-tolerant capability by each bitonic sorter cannot be achieved.

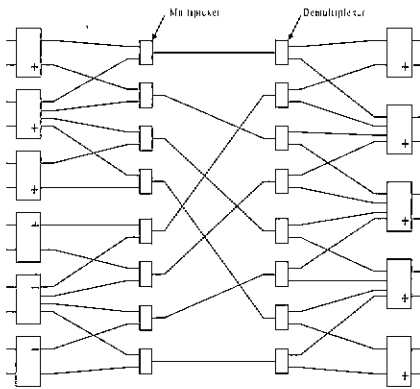
(Figure 6) shows how the perfect shuffle connection[11] is implemented in the DRBS Network. For simplicity the multiplexers and demultiplexers can be used to connect the bitonic sorters so that each of them can tolerate a single fault independently. However, the DRBS network implemented by the above technique can not tolerate a fault on the multiplexer, demultiplexers, or the perfect shuffle links. To solve this problem, multiplexers and demultiplexers are eliminated but more extra links are used instead. The connection scheme for the perfect shuffle stage, e.g., between lower bitonic sorters and upper one, is defined by the following :

- Let $\delta_{i-1, 2j}$, $\delta_{j-1, 2j+1}$, and δ_j be the number of the

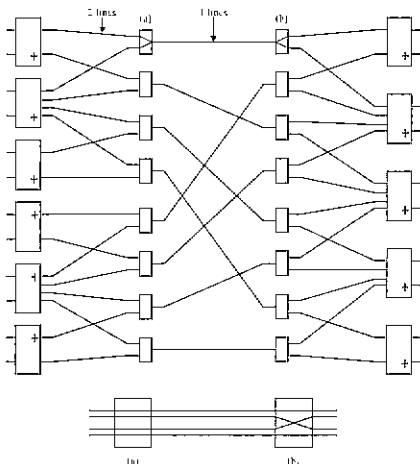
extra CEs for $(i-1, 2j)$, $(i-1, 2j+1)$, and (i, j) bitonic sorter, respectively. Here k represents the output port number of a bitonic sorter ($0 \leq k < (2^{i+1} + 2 * \delta_{i-1, 2j})$).

$$C_{2^i(k)} = \begin{cases} \text{if } 0 \leq k < 2^{i+1} \\ \text{shuffle}(k), \text{shuffle}(k) \div 2 * \delta_{ij} \\ \text{if } 2 * \delta_{i-1, 2j} \leq k < 2^{i+1} + 2 * \delta_{i-1, 2j} \\ \text{shuffle}(k - 2 * \delta_{i-1, 2j}), \\ \text{shuffle}(k - 2 * \delta_{i-1, 2j}) + 2 * \delta_{ij} \end{cases}$$

The connection for $C_{2^i+1}(i, 2j+1)$ bitonic sorter, is the same as C_{2^i} . As described in above equation, the number of links per CE is eight if the number of CE



(Figure 6) Perfect Shuffle Connection using MUXs and DEMUXs in the DRBSN

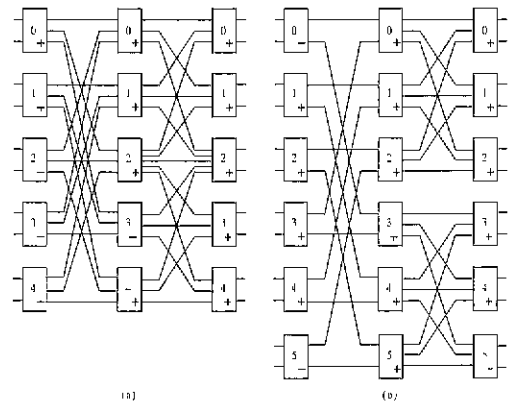


(Figure 7) Perfect Shuffle Connection using extra links in the DRBSN

is between $\delta_{i-1, 2j}$ and $2^i - 1$. Otherwise the number of link per CE is four. (Figure 7) shows the hardware of this perfect shuffle connection using extra links for the Dynamic Redundant Bitonic Sorting Network.

3.3 Variations of the Dynamic Redundant Bitonic Sorting network

In this section some variations of this Dynamic Redundant Bitonic Sorting network are discussed. Depending upon the size of δ_{ij} for each bitonic sorter several different DRBS networks can be constructed. The value of δ_{ij} represents the number of the extra CEs in a bitonic sorters. For convenience, the value of δ_{ij} will be power of two in this paper. The number of the extra CEs in (i, j) bitonic sorter is equal to $\delta_{ij} * (i + 1)$. The more extra CEs are used in a bitonic sorter, the less links per each CE will be used. (Figure 8) shows the connection of bitonic sorter which has different size of δ_{ij} . (Figure 9) shows the Dynamic Redundant Bitonic Sorting Network which can tolerate fault(s) by using additional CEs.



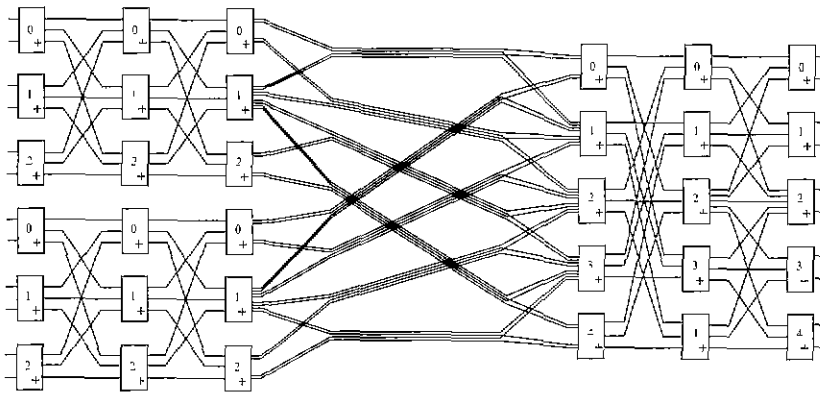
(Figure 8) Connection of bitonic sorter
(a) $\delta_{ij} = 1$ (b) $\delta_{ij} = 2i-1$

The DRBSN can even work even if an element in any bitonic sorter is faulted. However the upper level bitonic sorters which contain more comparison elements have much higher probability of a fault than the lower level one. For an example, a $(2, j)$ and $(9, j)$ bitonic sorter

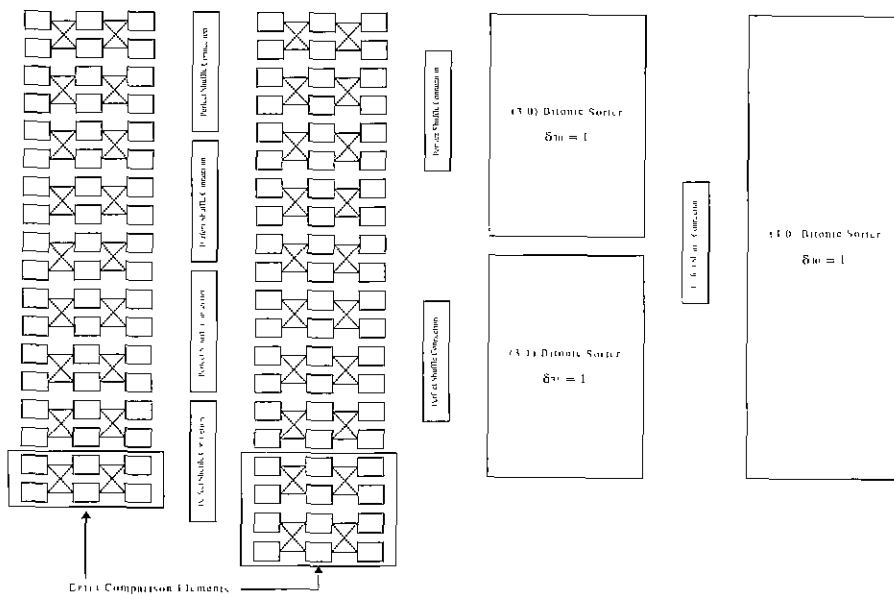
($\delta = 2^{l-1}$) can be constructed with 18 CEs and 7680 CEs, respectively. The DRBSN ($N=2^{10}$) based on this technique can tolerate maximum 128 faults in 2^{nd} level, but only one fault in 9^{th} level even if there are significantly more CEs in 9^{th} level. Therefore, in the number of bitonic sorters which can tolerate a single fault in lower level or the number of faults which can tolerate in upper level bitonic sorters has to be carefully considered based upon the trade-offs between the amount of hardware increase and the reliability of the system.

Two different schemes are presented in this section: one technique is to use less hardware and tolerate less faults especially in lower levels, the other one is to use more hardware and tolerate more faults especially in higher levels

(Scheme 1) The first one is based upon the idea in which it requires only six extra CEs to tolerate single fault in 0^{th} and 1^{st} level for any size of DRBS networks without using additional links within these bitonic sorters, 12 more CEs in 2^{nd} level, and so on. As shown



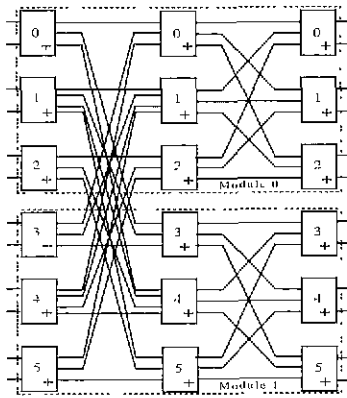
(Figure 9) An 8x8 New Fault Tolerant Bitonic Sorting Network, where $\delta_{ij} = 1$



(Figure 10) 32x32 Dynamic Redundant Bitonic Sorting Network using Scheme 1

in (figure 10), the network contains two extra bitonic sorters (CEs) in 0^{th} level, one extra bitonic sorter in 1^{st} and 2^{nd} level. However, for an upper level, the bitonic sorter with $\delta_y = 1$ could be used in order to reduce the number of extra CEs

(Scheme 2) The other method is based on the idea in which the more CEs and links are used, the more faults can be tolerated in the bitonic sorter. In order to tolerate more faults in the DRBS network with $\delta_y = 2^{i-1}$, additional links per CE can be used. The (i, j) bitonic sorter ($i \geq 2$) for the DRBSN with $\delta_y = 2^{i-1}$ can be divided into 2^{i-1} modules and each modules can tolerate a single fault (Figure 11), e.g., (i, j) bitonic sorter can tolerate maximum 2^{i-1} faults.



(Figure 11) Dynamic Redundant Bitonic Sorter using scheme 2 each module can tolerate a fault

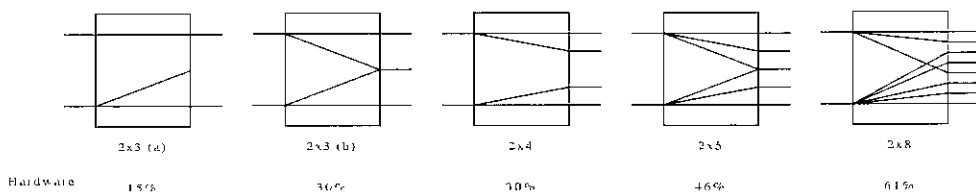
4. Analysis of Hardware Cost

As explained in the previous section, the new

comparison element contains input and output selectors so that the number of links per CE effects the hardware increase. The hardware increase of the new comparison element is shown in (figure 12). The original comparison element can be constructed by 13 NORs [4]. Therefore, the hardware increase of a CE with 2×3 input and output selector will be approximately 30% since each selector needs four switches and one inverter for the purpose of controlling the path of data (Figure 12 (b)).

<Table 1> shows the increase of the hardware of two different DRBS networks depending upon the size of δ_y . And <table 2> shows the hardware increase of two variations of the DRBS network (Scheme 1 and 2). Here, III represents the percent of hardware increase compared with the original network.

Although some of the DRBS networks can be constructed with relatively less hardware increase than the previous method, careful consideration is required in order for the design of them in VLSI. First of all the analysis of system reliability has to be performed to find out the most liable network among these DRBS networks. Consequently, the most reliable parameters such as the number of faults are required in each levels, the number of the extra CEs(= δ_y), and the number of the extra links. The chip area is the most important factor in the design of a VLSI chip. The value shown in (figure 12), <table 1>, and <table 2> for the hardware increase simply represents the increase of the control logic, not actual cost of the hardware increase in VLSI chip. Second, the proper layout of these network has to be considered. Usually a large bitonic sorter can be built from the standard modules of convenient size of a smaller one to save design and testing costs.



(Figure 12) Hardware increase of new CEs with each input and output selectors

<Table 1> The hardware increase of the Dynamic Redundant Bitonic Sorting network : $\delta_v=1$ and $\delta_w=2^i-1$ for all $0 < i \leq n-1$ and $0 \leq j < 2^{n-i}$.

Input Size	Original Network		DRBS Network ($\delta_v = 1$)			DRBS Network ($\delta_w = 2^{i-1}$)		
	CEs	Links	CEs	Links	HI(%)	CEs	Links	HI(%)
8	24	40	33	100	58.9	36	98	66.7
16	80	144	102	353	52.4	120	344	66.0
32	240	448	289	1060	46.8	360	1032	64.7
64	672	1280	776	2923	42.5	1008	2848	63.4
128	1792	3456	2007	7642	39.1	2688	7456	62.2
256	4608	8960	5046	19257	36.4	6912	18816	61.2
512	11520	22528	12405	47224	34.2	17280	49208	60.3
1024	28160	55296	29940	113399	32.5	42240	111104	59.5

<Table 2> The hardware increase of the two variations of DRBS network

Input Size	DRBS network(scheme 1)			DRBS network(scheme 2)		
	CEs	Links	HI(%)	CEs	Links	HI(%)
8	33	88	51.7	36	106	69.7
16	96	321	41.5	120	392	71.4
32	271	940	34.3	360	1224	71.8
64	722	2659	31.3	1008	3488	71.8
128	1881	6810	27.3	2688	9376	71.7
256	4744	17521	26.3	6912	24192	71.5
512	11751	42240	23.6	17280	60544	71.3
1024	28502	103231	23.4	42240	147968	71.1

5 Conclusions

This paper has focused on a new fault-tolerant technique for bitonic sorting networks which can be used for designing ATM switches based on Batchier-Banyan network. To overcome the single path limitation of the bitonic sorting network, a fault-tolerant sorting network(DRBSN) based on bitonic sorting has been presented. Additional CEs and links are used in order to recover a fault. Some variations of the proposed network can even tolerate several faults in specially upper level bitonic sorters which has considerably more comparison elements than lower level bitonic sorter. Unlike the other fault tolerant sorting networks this proposed network(DRBS) even works with faults on the extra hardware without any time delays in performance. The proposed network has high reliability, maintains cell

sequences, simple routing, regularity and modularity for VLSI implementation. This paper has also presented how the perfect shuffle connection can be implemented on the proposed DRBSN. The concept of this DRBS network can be applied to other sorting networks with slight modifications such as k-way sorting networks, etc.

References

- [1] "Broadband aspects of ISDN." CCITT recommendation, I.121, Nov. 1988.
- [2] H. Amano, "A Fault Tolerant Batchier Network," in Proceeding of the 1990 International Conference on Parallel Processing, Vol.I, pp.441-444, 1990.
- [3] S. Assaf and E. Upfal, "Fault-tolerant sorting network," in Proceedings of the 31st Annual IEEE Symposium, on Foundations of Computer Science, pp.275-284, Oct. 1990.
- [4] KE Batchier, "Sorting networks and their applications." 1968 Spring Joint Computer Conference, AFIPS Proc., Vol.32, pp.307-314
- [5] Imagawa, H., Urushidani, S. and Hagishima, K., "A new self-routing switch driven with input-output address difference," in Proc. GLOBECOM '88, pp.1607-1611, Dec. 1988.
- [6] Giacomelli et al. J., "Sunshine : A High-performance self-routing Broadband packet switch Architecture," IEEE J. of selected Areas in Communications. Vol.9, No.8, pp.1289-1298, Oct. 1991.
- [7] Pattavina, Achille, "Non-blocking Architectures for ATM Switching." IEEE Communications Magazine, pp.38-48, Feb. 1993.
- [8] L. Rudolph, "A Robust sorting network," IEEE Transactions on Computers Vol.C-43, pp.344-354, 1985
- [9] M. Schummler and C. Starke, "A Correction network for N-sorters," SIAM J. Comput. Vol.18, No.6, pp.1179-1187, Dec. 1989.
- [10] H.J. Siegel, Interconnection networks for Large-Scale Parallel Processing. Theory and Case Studies, Second Edition, McGraw-Hill, 1990.
- [11] H.S. Stone, "Parallel processing with the perfect

shuffle," IEEE Transactions on Computers. Vol.C-20, pp.153-161, 1971

[12] A.C. Yao and F.F. Yao, "On fault-tolerant networks for sorting," SIAM J Comput Vol.14, No.1, pp.120-128, Feb. 1985.



이재동

e-mail : jdlcc@cs.dankook.ac.kr

1985년 인하대학교 전자계산학과 (B.S)

1987년~1988년 대우 중공업 정보관리센터

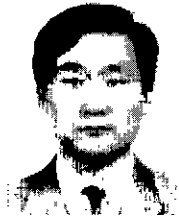
1991년 미국 Cleveland State University, Dept of computer and Information science (MS)

1996년 미국 Kent State University, Dept of computer science (Ph.D)

1996년~1997년 (주)두루넷, 기술기획팀

1997년~현재 단국대학교 전자계산학과 조교수

관심분야 : 분산/병렬처리, ATM networks, Internet Applications



김재홍

e-mail : jhong@kachi.yit.ac.kr

1985년 인하대학교 전자계산학과 (B.S)

1990년 인하대학교 전자계산학과 석사

1994년 인하대학교 전자계산학과 박사

1995년~현재 영동대학교 컴퓨터공학과 부교수

관심분야 : Multimedia DBMS, GIS, ATM networks



최홍인

e-mail : kchoi@widelink.co.kr

1982년 한성대학교 영어영문학과 (B.S)

1991년 미국 Kent State University, Dept. of CS (MS)

1997년 미국 Kent State University, Dept. of CS (Ph.D)

1998년~현재 단국대학교 강사겸 아코텍(주) 연구팀장
관심분야 : Sorting Networks, ATM networks