Survivability Performance Evaluation of an Optical Switch

Mohsen Guizani and A. Memon
University of West Florida
Email: mguizani@cs.uwf.edu

Abstract: An optical switch is designed and its survivability performance is presented in this paper. Its performance/reliability analysis is carried out and compared with two major fault-tolerant networks, Itoh's and Benes'. The results show that without redundant switches, better network survivability is achieved. In addition, much less hardware is needed. The preliminary results of the analysis show that the proposed switch has comparable performance to other fault-tolerant networks whereas it outperforms them in other aspects such as number of extra switches/links, thus providing better cost-effectiveness.

I. INTRODUCTION

There is a need for switches that are capable of routing unprocessed data to the next stage of an interconnection network even in case of failure. The next stage's regular/normal switch(es) should be able to process this data. This eliminates the need for extra switches. Moreover, no extra links are used to route the unprocessed data to the next stage.

Several switch implementations of fault-tolerant interconnection networks are found in the literature [1,2,3]. Almost all of these introduce redundancy in the network in terms of adding extra links and switches. These solutions are expensive since in most cases the number of links as well as switches are increased [2].

These types of switches are required to exchange large volumes of data to be executed on multiprocessor systems. This requires a robust interconnection network with good switching capabilities [4,5]. These could either be in the form of frequent short bursts such as ATM applications [2,7,8] or large continuous streams [9,10,11]

II. FAULT-TOLERANT NETWORKS

There are several possible techniques that can improve the fault-tolerance of a switching network. A fault-tolerant switch can be realized either by adding an extra stage of switches and/or links (ports), or varying the switch size.

The multipath omega network [12] was among the first of the proposed fault-tolerant networks. The original omega network does not have redundant paths whereas this network has multiple paths since it has redundant switching stages. The network can tolerate more faulty components as the number of switches increases. However, this causes a rapid increase in the number of undesirable crosspoints.

The augmented C-network [13], derived from its parent the C-network, is a fault-tolerant network that doubles both the input and output switching ports. The redundancy introduced by increased complexity of switches provides multiple paths between any input-output pair. Although the network is more fault-tolerant, its routing schemes become far more complicated than its parent network. In addition, extra computation is needed to determine routing paths and thus the advantages of self-routing are lost.

The extra stage cube network [14], derived from the generalized cube network, adds an extra stage to the input and output sides of its parent network along with multiplexers and demultiplexers. This network is single fault-tolerant and robust in the presence of multiple faults. However, extra logic and extra computation time are required to generate the new tags.

Tzeng [3] proposed a simple scheme to enhance the fault-tolerance of multistage interconnection networks (such as Banyan networks, omega networks, etc.) which only have a unique path between each input-output pair. The approach followed is one of creating multiple paths between each input-output pair by introducing extra links between switching elements (SE's) in the same stage. This scheme requires a simple routing algorithm and allows a low level of fault-tolerance with reasonable cost. However, the switching throughput of the network degrades quickly when the number of faulty components increases.

The augmented shuffle-exchange network [15], is constructed by the addition of certain links to the shuffle-exchange multistage interconnection network. Its advantage is that the input and output stages need not be fail-proof. However, the switching throughput of this network degrades rapidly when the number of faulty components increases.

Itoh's self-routing fault-tolerant ATM switching network [2] provides multiple paths by adding many subswitches between switching stages of a Banyan network. It has a large number of redundant paths. It maintains high throughput with acceptable switching delay even when element failures occur. However, substantial amount of redundant SEs/links are required and the fault-tolerance decreases drastically on switch failure in the first and last stages of the network.

III. PROPOSED ARCHITECTURE

The proposed 2X2 switch architecture is an extension of existing switches that achieve straight and exchange connections. Fault-tolerance is increased by introducing buffers in the switch that act as queues and store the packets that cause contention, hence preventing their loss. Additional fault-tolerance is provided by augmenting the

switch with two circuits. These are responsible for detecting faults in the network and correcting them. They can also route data to other switches in case the main module of the switch fails.
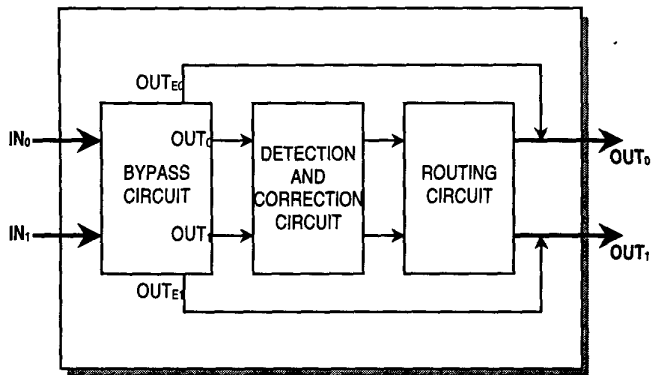

Fig. 1. Block Diagram of the 2X2 Switch

The top level breakdown of the complete switch is shown in Fig.1. It consists of three circuits. The first is the bypass circuit that passes the data to the next stage when the current switch fails. The second is the error detection and correction circuit responsible for detecting errors in the data and removing them before being processed by the next routing switch. The routing switch examines the bits in the address field of the data and routes it to the appropriate output. In case of contention it may loop the data within the same switch.

The first two circuits are used only for handling faults in the network. If the routing switch of stage i fails to respond correctly, its bypass circuit routes the data to the next stage $i+1$ where routing decisions are taken. Details of the subcomponent are described in [4].

The bypass circuit takes over the operations whenever the routing circuit (within the same switch) fails. Due to its simple design, it does not perform complex operations on the incoming data. The main purpose of this circuit is to channel incoming data to the next stage of the network. Incoming packets are channeled to all possible outputs (in this case two) of the switch. This duplication temporarily creates extra traffic in the network. It is up to the switch in the next stage to determine whether the incoming data packet is indeed destined to it.

To enable successive stages to correctly detect and remove data duplication in the network, an extra field of *error bits* is introduced by the bypass circuit. To minimize

complexity in the bypass circuit hardware, 0 (1) is appended to the incoming packet depending on whether it is routed to the upper (lower) link.

Here, we explain the functionality of the error detection and correction circuit using an example. Let F_i represent a faulty switch and W_i represent a properly working switch. In Fig. 2, the bypass circuit of F_1 (faulty switch in the first stage) sends the data packet to both output links connected to W_1 and F_2 which are properly working and faulty switches in the next stage, respectively. Note that the error field of the packet contains only one bit since it encountered only one faulty switch. At the second stage, the extra traffic created due to F_1 needs to be detected and removed. On receiving both data packets, W_1 strips off the error field bits and an equal number of routing bits and compare them. Equivalence of these bits validates presence of the data packet at that switch, otherwise the data packet is rejected. Consequently, in Fig. 2, W_1 rejects the first packet and accepts the second. Since W_1 is working properly, the data is routed to the correct destination despite the presence of a faulty switch in the network. This process continues with the rest of the stages. This type of detection and correction can be achieved even in the presence of $n-1$ faulty switches, where n is the number of stages in the network.

IV. RELIABILITY ANALYSIS

The assumptions made in carrying out this analysis are: First, the failure of one switch can in no way affect the reliability of any other switch in the network. Second, the network is said to have failed if at least one connection between input and output ports can not be realized. Most of other reliability analyses given in the literature assume the first and last stages of the network are fully operational under all conditions. This is not a realistic assumption since all the switches in the network are equally likely to fail. Therefore, in this analysis, all switches have equal probability of failure, that is including all switches of the first and last stages.

The above assumptions are used to compute the survivable probability, $Q(k)$, for both Itoh's and Benes' networks (see fig. 3 and 4). Fig. 5 shows the survival probability, $Q(k)$, of the proposed network. It was found that it performs better (using the proposed network) for the same range of faults in addition to having lower slope of the corresponding curves. Note that in the proposed network, no additional switches/links are required. The total number of switches in the network is $n2^{n-1}$. Any of these switches can fail with equal probability.
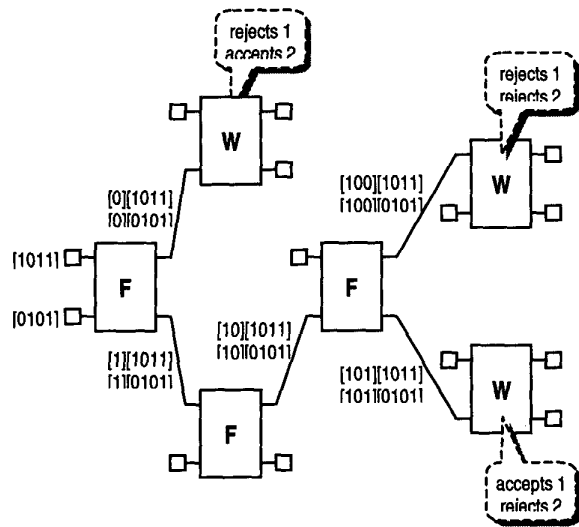
Fig. 2. Example of One Faulty Switch, Redundant Data and Error Recovery
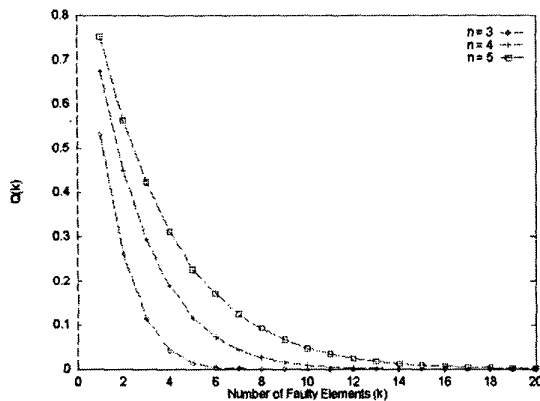

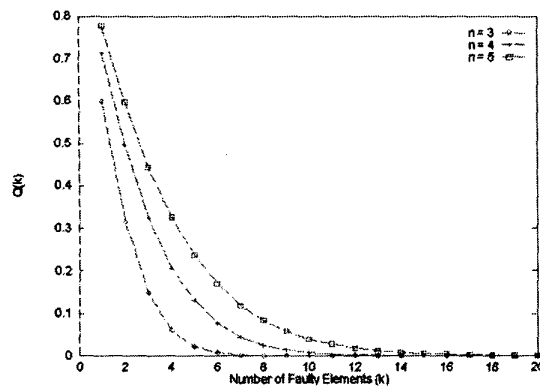Fig. 3. Network Survivability of Itoh's Network for Different Number of Faults

Fig. 4. Network Survivability of Benes Network for Different Number of Faults
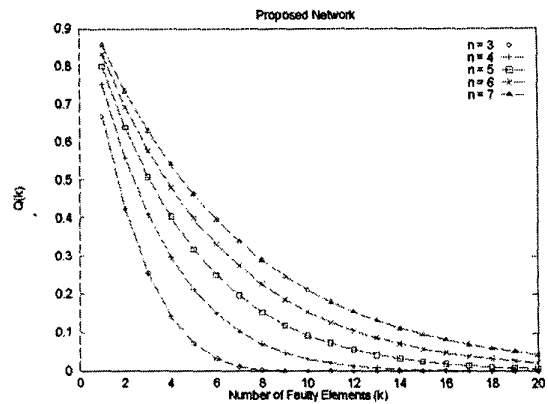

Fig. 5. Network Survivability of Proposed Network for Different Number of Faults

If the number of failures is k, where $0 \le k \le n2^{n-1}$, then the number of configurations in which k failures can occur is

$$\binom{n2^{n-1}}{k}$$

Therefore the survival probability $Q(k)$ is given by:

$$Q(k) = \binom{n2^{n-1}}{k} \bigg/ \binom{(n-1)2^{n-1}}{k}$$

The total number of switches of the proposed switch is computed and compared to Benes' and Itoh's and found that in the proposed architecture, the number of switches is reduced to about 50%. To compute the cost-effectiveness, let ε be the expected number of faulty switches that will cause the entire network to fail. Then ε can be obtained as:

$$\varepsilon = \sum_{i=2}^{L} i p(i)$$

The parameter L is the total number of switches that can fail. Using the above and assuming that all the switches of the network are prone to fail, it is seen from Table 1 that the cost-effectiveness of all the networks reduces, but still performs better when using the proposed network.

Table 1. Cost-Effectiveness

Network	n	ε	L	E/L (%)
Benes	3	5.1	40	12.75
	4	8.3	112	7.41
	5	12.1	256	4.20
Itoh	3	7.1	44	16.13
	4	15.1	132	11.43
	5	29.8	356	2.58
Proposed	3	4	12	33.33
	4	8	32	25.00
	5	16	80	20.00

V. CONCLUSION

In this paper, we have studied the survivability performance of a fault-tolerant 2X2 switch that can make use of optical and electrical components. Analysis was carried out by incorporating the proposed fault-tolerant switch into the baseline network. However, it can be employed as the basic building block for other multistage interconnection networks. The results of the survivability performance analysis show that without using any additional hardware, better network survivability is achieved. Cost-effectiveness analysis shows that the proposed switch is much more cost-effective than other fault-tolerant switches. Comparison to Itoh's and Benes' networks was performed.

REFERENCES

[1] H. S. Kim and A. Leon-Garcia, "A self-routing multistage switching network for broadband ISDN," *IEEE J. Select. Areas in Communications*, Vol. 8, No. 3, 1990, pp.459-466

[2] A. Itoh, "A fault-tolerant switching network for B-ISDN," *IEEE J. Select. Areas in Communications*, Vol. 9, No. 8, 1991, pp. 1218-1226

[3] N. Tzeng, P. Yew, and C. Zhu, "A fault-tolerant scheme for on fault-tolerant multistage interconnection network," *12th International Symposium on Computer Architecture*, June 1985, pp. 368-375

[4] M. Guizani, "Picosecond multistage interconnection networks architecture for optical computing," *Applied Optics*, Vol. 33, No. 8, 1994, pp. 1587-1599

[5] T. Y. Feng, "A survey of interconnection networks," *IEEE Computer Magazine*, Vol. 4, Dec. 1981, pp. 12-27

[6] H. Ishikawa, "High-speed packet switching systems for multimedia communications," *IEEE J. Selected Areas in Communications*, Vol. 5, Oct. 1987, pp. 1336-1345

[7] A. Hac and H. B. Mutlu, "Synchronous optical network and broadband ISDN protocols," *Computer*, Vol. 22, 11, Nov. 1989, pp. 26-34

[8] R. Handel, "Evolution of ISDN towards broadband ISDN," *IEEE Network*, Jan. 1989, pp. 7-13

[9] J. S. Turner, "New Directions in Communications," *Proc. IZS'86*, 1986, Paper A3, pp. 1-8

[10] J. J. Kulzer and W. A. Montogomery, "Statistical switching architecture for future services," *Proc. ISS'84*, 1984, paper 43A.1, pp. 1-6

[11] H. Imagawa, "A new self-routing switch driven with input-output address difference," *Proc. GLOBECOM'88*, Dec. 1988, pp. 1607-161

[12] K. Padmanabham and D. H. Lawrie, "A class of redundant path multistage interconnection networks," *IEEE Trans. On Computers*, Vol. 32, No. 12, 1983, pp. 1099-1108

[13] S. M. Reddy and V. P. Kumar, "On fault-tolerant multistage interconnection networks," *Proc. of the International Conference on Parallel Processing*. Aug. 1984, pp. 155-164

[14] G. B. Adams and H. J. Siegel, "Modifications to improve the fault-tolerance of the extra stage cube interconnection network," *Proc. of the International Conference on Parallel Processing*. Aug. 1984, pp. 169-173

[15] V. P. Kumar and A. L. Reibman, "Failure dependent performance analysis of a fault-tolerant on fault-tolerant multistage interconnection network," *IEEE Trans. on Computers*, Vol. 38}, No. 12, pp. 1703-1713