

MULTISTAGE INTERCONNECTION NETWORKS: A TRANSITION TO OPTICAL

Sandeep Kumar¹, Arpit Kumar²

¹Sekhawati Engg. College, Dundlod, Dist. - Jhunjhunu (Raj.), 1987san@gmail.com, ²KIIT, Gurgaon (HR.),

Abstract

Several types of interconnection networks have been proposed for parallel processing. One of them is MIN's (Multistage interconnection networks). With the better way of communication between processor and memory, MIN's are less complex, good fault tolerance, high reliability, fast in communication and low cost. To achieving reliable, fast and flexible communication optical communication are necessary. Through optical networks we can achieve high performance and low latency also. With their great advantages over electronic networks, OMIN's also having their own demands and problems. This paper focuses on the qualities of Optical MINs. Here we compare Optical MINs with electronic MIN's with their design issues and gives some solution approaches for optical MINs to solve crosstalk.

Index Terms— Multistage Interconnection Networks (MIN), Optical networks, Crosstalk, Window methods.

1. INTRODUCTION

In earlier days of parallel computing systems, the metal-based communication technology becomes a chokepoint to meet the increasing demands of high performance computing applications for high channel bandwidth and low communication latency. Now the requirement arise to improve the existing interconnection technology or to introduce some new interconnection technology in parallel computing systems. Electro-optic technologies have made optical communication a promising network choice to meet the increasing demands because Fiber optic communications offer a combination of high bandwidth, low error probability and very high speed transmission capacity.

For parallel computing systems, multistage interconnection networks have been extensively accepted as an interconnecting scheme. As optical technology advances, there is considerable interest in using optical technology to implement interconnection network and switches. A multistage interconnection network is composed of several stages of switch elements by which any input port can be connected to any output port in the network. This network consists of N inputs, N outputs and n stages ($n = \log_2 N$). Each stage has $N/2$ switching elements each SE has two inputs and two outputs connected in a certain pattern. The most widely used MINs are the electronic MINs. In electronic MINs,

electricity is used and in optical MINs, light is used to transmit the messages. Although electronic MINs and optical MINs have many similarities but there are some fundamental differences between them. Optical MINs were built mainly on banyan or baseline or omega networks because they are fast in switch setting (self-routing) and also have a small number of

switches between an input-output pair [1]. Path dependent loss is a challenge in optical MIN's; it means optical signals become weak after passing through optical path. Crosstalk in optical networks is one of the major shortcomings in optical switching networks, and avoiding crosstalk is an important for making optical communication properly. To avoid a crosstalk, many approaches have been used such as time domain and space domain approaches.

2. MULTISTAGE INTERCONNECTION NETWORKS:

Multistage interconnection networks (MINs) consist of more than one stages of small interconnection elements called switching elements and links interconnecting them. A MIN normally connects N inputs to N outputs and is referred as an $N \times N$ MIN. The parameter N is called the size of the network. Figure 1 illustrates a structure of multistage interconnection network. This figure shows the connection between p inputs and b outputs, and connection between these is via number of stages.

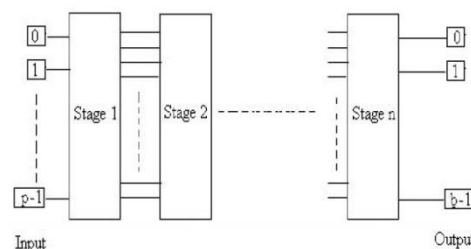


Figure 1: A Multistage Network

Multistage interconnection network is actually a compromise between crossbar and shared bus networks of various types of multiprocessor interconnections networks. Multistage interconnection networks (MINs) are used in multiprocessing systems to provide cost-effective, high bandwidth communication between processors and/or memory modules. Multistage interconnection networks attempt to reduce cost and decrease the path length [14].

TABLE 1 Properties Of Different Inter Connection Techniques

<i>Property</i>	<i>Bus</i>	<i>Crossbar</i>	<i>Multistage</i>
Speed	Low	High	High
Cost	Low	High	Moderate
Reliability	Low	High	High
Configurability	High	Low	Moderate
Complexity	Low	High	Moderate

3. OPTICAL MULTISTAGE INTERCONNECTION NETWORKS

Implementation of OMIN can be either free-space optics or guided wave technology. It uses (TDM) Time Division Multiplexing. Wavelength Division multiplexing (WDM) technique is used to exploit the huge optical bandwidth of fiber. With WDM, the optical spectrum is divided into many different logical channels, to a unique wavelength. Optical switching, involves the switching of optical signals, rather than electronic signals as in conventional electronic systems. Two types of guided wave optical switching systems can be used. The first is a hybrid approach in which optical signals are switched, but the switches are electronically controlled.

With this approach, the use of electronic control signals means that the routing will be carried out electronically. As such, the speed of the electronic switch control signals can be much less than the bit rate of the optical signals being switched. So, with this approach there is a big speed mismatch occur due to the high speed of optical signals. The second approach is all-optical switching. This has removed the problem that occurred with the hybrid approach but, such systems will not become practical in the future and hence only hybrid optical MINs are considered. In hybrid optical MINs, the electronically controlled optical switches, such as lithium niobate directional couplers, can have switching speeds from hundreds of picoseconds to tens of nanoseconds.

A. Optical networks switching

In optical networks Packet switching is not possible with Optical Multistage Interconnection Networks. If packet switching is used, the address information in each packet must be decoded in order to determine the switch state. In a hybrid MIN, it means it require conversions from optical signals to electronic ones, which could be very costly. For this reason, circuit switching is usually preferred in optical MINs.

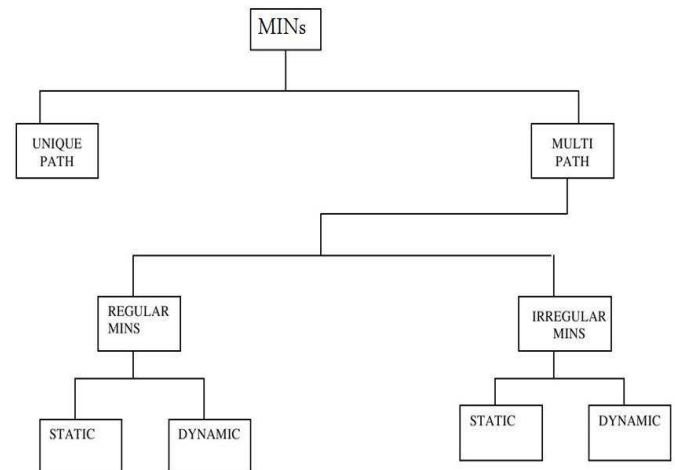


Figure 2: Types of Multistage Networks

B. Electronic or Optical Networks: A Comparison

There are lots of advantages of optical networks over the electronic ones. The main advantage of the optical networks over the electronic network is the high speed of the Optical signals. In the optical networks light is transmitted which has a very good speed but in the electronic Multistage interconnection networks electricity is used which has very slow speed. The second advantage is the Bandwidth. Now a day's applications in communication require high bandwidth, the optical networks gives combination of very high bandwidth and low latency. Therefore, they have been used in the parallel processing applications. Optical MINs are also used in wide area networks, which require less error probability and very high bandwidth. Fiber optic transmission distance is significantly greater than the electronic ones, signal need not to be regenerated in optical networks. Optical fiber has very less weight in comparison to electronic MINs. Thus Optical networks give the combination of high bandwidth and low latency.

4. PROBLEMS IN OPTICAL NETWORKS

Due to the difference in speeds of the electronic and optical switching elements and the nature of optical signals, optical MINs also hold their own challenges.

A. Path Dependent Loss

Path dependent loss means that optical signals become weaker after passing through an optical path. In a large MIN, a big part of the path-dependent loss is directly proportional to the number of couplers that the optical path passes through. Hence, it depends on the architecture used and its network size. So, if the optical signal has to pass through more no of stages or switches the path dependent loss will be more.

<i>Characteristics</i>	<i>Electronic MIN</i>	<i>Optical MIN</i>
Latency	High	Less
Energy Transmitted	Electricity	Light
Bandwidth	Used for less bandwidth	Used for high bandwidth
Cost	Less	More
Speed	Less	High
Weight	More	Less
Switching	Packet switching	Circuit Switching
Complexity	More	Less
Structure	2Dimensional	3 Dimensional
Error Probability	High	Less
Path	Multi path from S to D	Single path from S to D

B. Crosstalk

Optical crosstalk occurs when two signal channels interact with each other. There are two ways in which optical paths can interact in a switching network. The channels carrying the signals could cross each other. Alternatively; two paths sharing a switch could experience some undesired coupling from one path to another within a switch. Crosstalk problem is more dangerous than the path-dependent loss problem with current optical technology. Thus, switch crosstalk is the most significant factor that reduces the signal-to-noise ratio and limits the size of a network.

5. APPROACHES TO SOLVE CROSSTALK

A. Space Domain Approach

One way to solve crosstalk problem is a space domain approach, where a MIN is duplicated and combined to avoid crosstalk [8]. The number of switches required for the same connectivity in networks with space domain approach is slightly larger than twice that for the regular network. This approach uses more than double the original network hardware to achieve the same. Thus for the same permutation the hardware or we can say the no of switches will be double. Thus cost will be more with the networks using space domain approach.

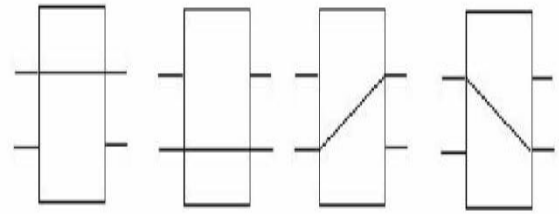


Figure 3: Crosstalk avoidance using space domain approach

B. Time Domain Approach

Another way to solve the problem of crosstalk is the time domain approach. With the time domain approach, the same objective is achieved by treating crosstalk as a conflict; that is, two connections will be established at different times if they use the same SE. Whereas we want to distribute the messages to be sent to the network into several groups, a method is used to find out which messages should not be in the same group because they will cause crosstalk in the network. A set of connections is partitioned into several subsets such that the connections in each subset can be established simultaneously in a network. There is no crosstalk in these subsections. This approach makes importance in optical MINs for various reasons. First, most of the multiprocessors use electronic processors and optical MINs. There is a big mismatch between the slow processing speed in processors and the high communication speed in networks carrying optical signals. Second, there is a mismatch between the routing control and the fast signal transmission speed.

6. METHODS FOR MESSAGE PARTITIONING IN TDM APPROACH

A. Window method

Window method is the method that is used to find the messages that are not in the same group because it causes crosstalk in the network. If we consider the network of size $N \times N$, there are N source and N destination address. Combination matrix is formed by combining source and its destination address. From this, optical window size is $M-1$, where $M = \log_2 N$ and N is size of network. In window method, number of windows is equal to number of stages.

After finding conflicts using window method, conflict graph is generated shown in figure 5.2. The number of nodes is the size of the network. The nodes that are having conflict are connected through edge. Degree of each message is the number of conflicts to the other message. The conflict matrix is a square matrix with $N \times N$ entry, it consists of the output of the window method shown in table 5.1. The definition of Conflict Matrix is the matrix M_{ij} with size $N \times N$. N is the size of the network.

B. Bitwise combination matrix

For Bitwise combination matrix, all binary bits of single rows is each windows are converted to decimal, no. of window is reduced to n. By this method, time is reduced approximately by ten times. It is very effective method even when the network is very large.

For example the bitwise combination matrix for 8 8 network size is demonstrated in Fig. 5. The number of columns in WM is 6 ($C_i, i = 2n$) and for bitwise WM is 2 ($C_i, i = n$) [11,12].

0 0 0 1 0 1	message 000 and 100 have conflict
0 0 1 0 0 1	message 001 and 101 have conflict
0 1 0 0 1 1	message 010 and 110 have conflict
0 1 1 1 1 0	message 011 and 111 have conflict
1 0 0 0 0 0	
1 0 1 0 1 0	
1 1 0 1 0 0	
1 1 1 1 1 1	
Step 1(w0)	
0 0 0 1 0 1	message 000 and 110 have conflict
0 0 1 0 0 1	message 001 and 101 have conflict
0 1 0 0 1 1	message 010 and 100 have conflict
0 1 1 1 1 0	message 011 and 111 have conflict
1 0 0 0 0 0	
1 0 1 0 1 0	
1 1 0 1 0 0	
1 1 1 1 1 1	
Step 2(w1)	
0 0 0 1 0 1	message 000 and 110 have conflict
0 0 1 0 0 1	message 001 and 100 have conflict
0 1 0 0 1 1	message 010 and 101 have conflict
0 1 1 1 1 0	message 011 and 111 have conflict
1 0 0 0 0 0	
1 0 1 0 1 0	
1 1 0 1 0 0	
1 1 1 1 1 1	

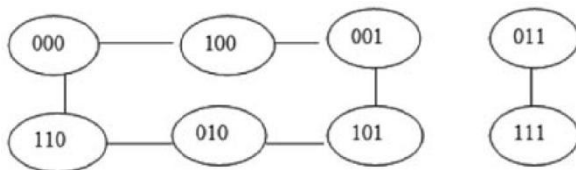


Figure 4. Conflict graph

msg	000	1	0	1	0	1	0	1
0	0	0	0	0	1	0	1	0
1	0	0	0	0	1	1	0	0
0	0	0	0	0	1	1	1	0
1	0	0	0	0	0	0	0	1
0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0

Figure 5: Conflict Matrix

C. Improved window method

In this method the first window is eliminated for this we make the conflict matrix initialized to 0, here number of windows is M-1. It takes less time to find conflicts than the windows method. Therefore, it is called improved window method.

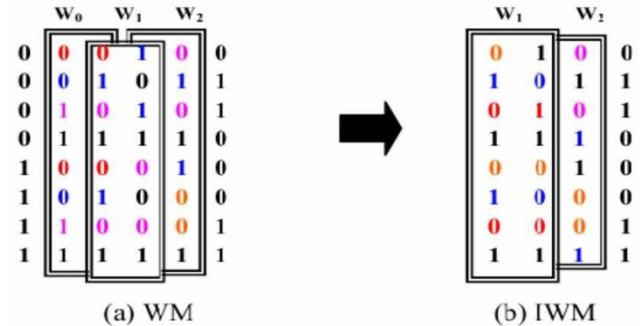


Figure 6: Improved window method

C_0	C_1	C_2	C_3	C_4	C_5	$C_{1,2}$	$C_{2,3}$	$C_{3,4}$
0	0	0	1	0	0	0	1	2
0	0	1	0	1	1	1	2	1
0	1	0	1	0	1	2	1	2
0	1	1	1	1	0	3	3	3
1	0	0	0	1	0	0	0	1
1	0	1	0	0	1	1	2	0
1	1	0	0	0	0	2	0	0
1	1	1	1	1	1	3	3	3

Combination Matrix

Bitwise Combination Matrix

D. Bitwise window method

In this method, source and destination address is in decimal format. Number of windows is $\log_2 N$. Thus, from combination matrix, the optical window size is only one for a different network size and the number of window is $\log_2 N$. In other words, there are only one decimal number in each row and each window for comparison and finding a conflict.

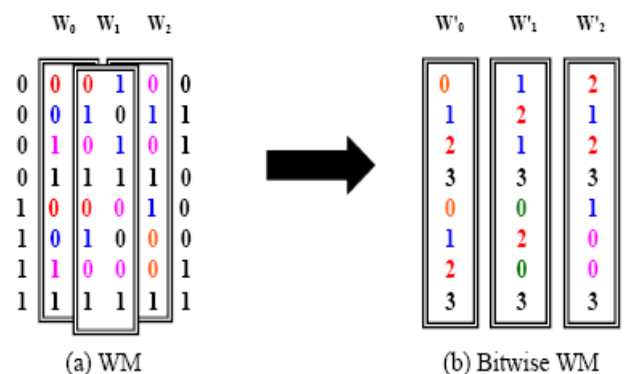


Figure 7: Bitwise window method

E. Sequential Window Method:

SWM is used to find the conflicts among the messages to be sent. It can be described briefly as follows [16]. Given a permutation, we combine each source address and its corresponding destination address to produce a matrix. The optical window size is the $m-1$ where $m = \log_2 N$ and N is the size of the network. We use this window on the produced matrix from left to right except the first column and last column. If two messages have the same bit pattern in any of the optical window, they will cause conflict in the network. That means they cannot be in the same group, hence, they have to be routed in different passes. In Figure 5.6, take the second and third columns as a matrix, messages 000 and 100 in this window have the same bit pattern of 00 inside the window and hence have a conflict. The bit patterns can be any of the four combinations of 00, 01, 10, and 11, and hence are shaded using different colors.

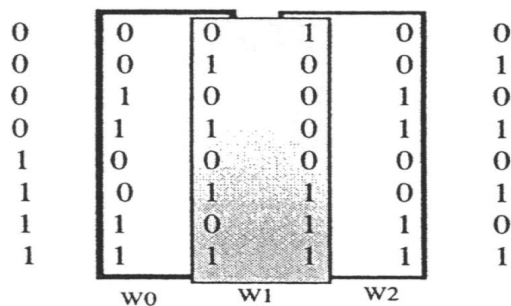


Figure 8: Sequential Window method

From this window, the messages 000, 001, 010 and 011 have conflict with messages 100, 101, 110 and 111 respectively.

F. Unbalanced Parallel Window Method:

In this algorithm, the master processor will send the whole window to the slave processor which in turn will do the comparisons and then put the result in a buffer. Thereafter, the slave will send the buffer to the master. Finally the master processor will generate the conflict matrix. It starts with description of operations of the master process P_0 and finishes by the same description of the slave processes P_p . In this algorithm, some processors are not working while the others are working. In another meaning, there is neither throughput nor load balancing. As an example of an 8×8 network size, when the first window is eliminated by IWM algorithm, the number of window will be only two. The master process will send the first window to process 0 while the second window will be sent to process P_1 as shown in figure 8.

In this algorithm, the master process works as a control process just in the beginning, when the master is waiting the results from slaves, it does the comparison of one window. After that, it totally works as a control process.

G. Balanced Parallel Window Method:

Load balancing is the task of equally dividing work among the available processors. This can easily be done when the same operations are being performed by all the processes. In this algorithm, an efficient approach is proposed to solve the unbalancing problem. An independent problem is found in window itself. The problem is divided into small sub problems. In other words, each window is divided into many Sub Windows. In each sub window, there is a conflict. Each window in network size 8×8 can be divided into seven independent sub windows. In general, each window in network size $N \times N$ can be derived into $N-1$ sub windows. The division of window is made according to the contents of figure 9.

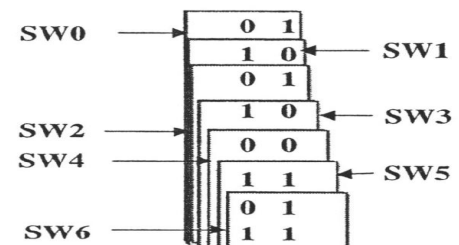


Figure 9: Decompose Window in the BPWM

CONCLUSIONS

In this paper properties of electronic and optical MINs have been explained and compared. It is concluded that for today's applications such as in wide area networks (WANs) optical networks is the promising choice to meet the high demand of Speed and Bandwidth. The paper also describes the problems and solutions of optical MINs. Various methods available in literature, which are used for crosstalk avoidance in TDM approach, are described in detail. It has been observed that the improved window method takes lesser time to find conflicts as compared to window method. The bitwise window method reduces the execution time ten times than the other algorithms even when the network size is large. Because of the dependence of each window, we can execute this method in parallel to reduce the execution time. Through load balancing, Balanced Parallel Window Method increases the performance of a network.

REFERENCES

- [1] Sanjeev Claire and Ashwani Kumar Singla “Comparison of SA and GA Algorithm for Minimizing Crosstalk in Optical MIN”, International Journal of P2P Network Trends and Technology- Volume2Issue1- 2011
- [2] A.K. Katangur, Y. Pan and M.D Fraser, “Message Routing and Scheduling in Optical Multistage Networks Using Simulated Annealing”, International Proceedings of the Parallel and Distributed Processing Symposium (IPDPS), 2002.
- [3] C. Qiao, and R. Melhem, “A Time Domain Approach For Avoiding Crosstalk In Optical Blocking Multi-stage Interconnection Networks”, Journal of Lightwave Technology, vol. 12 no. 10, 1994, pp. 1854-1862.
- [4] C. Siu and X. Tiehong, “New Algorithm for Message Routing and Scheduling in Optical Multistage Interconnection Network”, Proceedings of International Conference on Optical Communications Systems and Networks, 2004.
- [5] D. K. Hunter and I. Andonovic, “Guided wave optical switch architectures,” International Journal of Optoelectronics, vol. 9, no. 6, 1994, pp. 477-487.
- [6] Hasan, “Rearrangeability of $(2n - 1)$ -Stage Shuffle-Exchange Networks”, Society for Industrial and Applied Mathematics, vol. 32, no. 3, 2003, pp. 557-585.
- [7] J. T. Blaket and K.S. Trivedi, “Reliabilities of Two Fault-Tolerant Interconnection Networks”, Proceeding of IEEE, 1988, pp. 300-305.
- [8] K. Padmanabhan and A. Netravali, “Dilated Networks for Photonic Switching”, IEEE Transactions on Communication, vol. 35, no. 12, 1987, pp. 1357-1365.
- [9] L. N. Bhuyan and D.P. Aggarwal, “Design and performance of generalized interconnection networks”,
- [10] L. N. Bhuyan, Q. Yang Qing and D.P. Aggarwal, “Performance of Multiprocessor Interconnection Networks”, IEEE Computers, vol. 22, 1989, pp. 25-37.
- [11] M. A. M. Othman and R. Johari, “An efficient approach to avoid crosstalk in optical Omega Network”, International Journal of Computer, Internet and Management, vol. 14, no. 1, 2005, pp. 50-60.
- [12] M. Ali, M. Othman, R. Johari and S. Subramaniam, “New Algorithm to Avoid Crosstalk in Optical Multistage Interconnection Networks”, Proceedings of IEEE International Conference on Network (MICC-ICON), 2005, pp. 501-504.
- [13] A. Verma and C.S. Raghvendra, “Interconnection Networks for Multiprocessors and Multicomputers: Theory and Practice”, IEEE Computer Society Press, Los Alamitos, California, 1994.
- [14] Bin Zhou and M. Atiquzzaman, “Efficient Analysis of Multistage Interconnection Networks Using Finite Output-Buffered Switching Elements”