FAST ZEROX ALGORITHM FOR ROUTING IN OPTICAL MULTISTAGE INTERCONNECTION NETWORKS

T. D. SHAHIDA, *M. OTHMAN AND M. K. ABDULLAH

Dept. of Communication Technology and Network, University Putra Malaysia, 43400 UPM Serdang, Selangor Darul Ehsan, Malaysia.

{dian, mothman}@fsktm.upm.edu.my

***ABSTRACT* :** Based on the ZeroX algorithm, a fast and efficient crosstalk-free time-domain algorithm called the Fast ZeroX or shortly FastZ_X algorithm is proposed for solving optical crosstalk problem in optical Omega multistage interconnection networks. A new pre-routing technique called the inverse Conflict Matrix (iCM) is also introduced to map all possible conflicts identified between each node in the network as another representation of the standard conflict matrix commonly used in previous Zero-based algorithms. It is shown that using the new iCM, the original ZeroX algorithm is simplified, thus improved the algorithm by reducing the time to complete routing process. Through simulation modeling, the new approach yields the best performance in terms of minimal routing time in comparison to the original ZeroX algorithm as well as previous algorithms tested for comparison in this paper.

***KEYWORDS* :** *Optical Multistage Interconnection Networks (OMINs), Zero-based Routing Algorithm, Heuristics Sequential Increase Algorithm and Time Domain Approach.*

1. INTRODUCTION

Emerging of multicore systems has witnessed the trend towards powerful machines with higher computing power available at the reach of fingertips. Future telecommunication systems are expected to provide seamless support for higher transmission capacity and faster switching technology to connect these high-end systems, in line with the explosive growth of the Internet. Interconnection networks (INs) is faced with greater challenge as the direction in computing systems is fast moving towards advanced architectures such as chip multiprocessors (CMPs), systems on chip (SoC) and network on chip (NoC) [1].

Multistage interconnection networks (MINs) have been proposed as interconnecting structures in various types of communication applications ranging from parallel systems [2], switching architectures [3], to multicore systems [4]. Advances in optical technologies have drawn the interest for optical implementation in MINs to achieve high bandwidth capacity at the rate of terabits per second. Optical MINs (OMINs) are an attractive solution that offers a combination of high bandwidth, low error probability, and large transmission capacity [5].

However, OMINs introduce optical crosstalk, which results from coupling two signals within a switching element (SE). Optical crosstalk degrades the performance of OMINs in terms of reduced signal-to-noise ratio and limits the size of the network [6]. Limited by the properties of optical signals, it is not possible to route more than one message simultaneously, without optical crosstalk, over a switching element in an OMIN. Reducing the effect of optical crosstalk has been a challenging issue considering trade-offs between aspects i.e. performance, hardware and software complexity. The three common approaches by which the effect of crosstalk can be reduced are through network dilation in either the space, time or wavelength domain [6]. In this paper, our interest is on the time domain approach for solving optical crosstalk in the optical Omega network, a class of self-routable networks, topologically equivalent to many other topologies i.e. the baseline, butterfly, and cube networks [7].

1.1 Optical Crosstalk in Optical Omega Networks

An Omega network connects N input to N output nodes using n stages, where $n = \log_2 N$ with each stage containing 2^{n-1} switching elements (SEs). Figure 1(a) shows an example of a permutation connection when $N = 16$. Permutation refers to the one-to-one mapping of an input to an arbitrary output port in the Omega network.
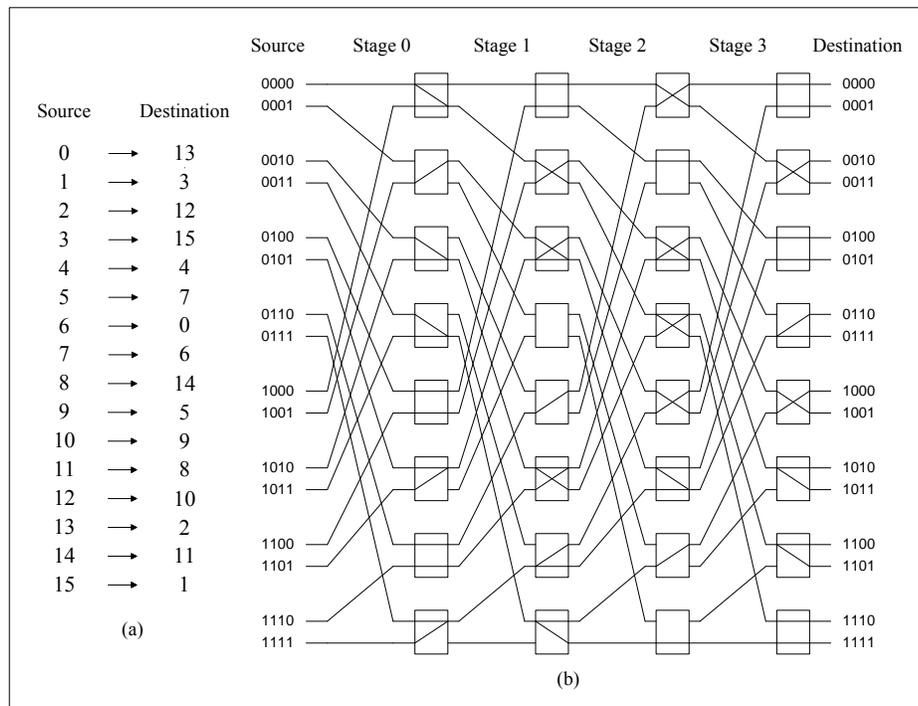

Fig. 1: (a) A 16 x 16 Permutation, and (b) Routing the Permutation in an Optical Omega Network.

To route an optical packet from an input to an output port, the routing path is set according to its destination address. In the Omega network, each optical packet must go through a number of switching stages before reaching to their destination. Each of the SEs

in the network can be of two logical states; either straight or cross connection. To utilize the available network bandwidth, it is a practical approach to route permutation simultaneously. Figure 1(b) illustrates routing the permutation in a 16 x 16 Omega network. It is clear in Fig. 1(b) that there is optical crosstalk contribution in most of the SEs when more than one link is active in the SE. Depending on the amount of voltage at the junction of the two waveguides that carry two input signals into the SE, a portion of either of the two optical signals can be detected at either of the two outputs signals. Figure 2(a) shows the two types of logic states of an SE while Fig. 2(b) shows the an illustration of the optical crosstalk occurrence within the SE connected with straight logic state. Consequently, the input signal will be distorted at the output due to loss and crosstalk accumulated along the connection path.
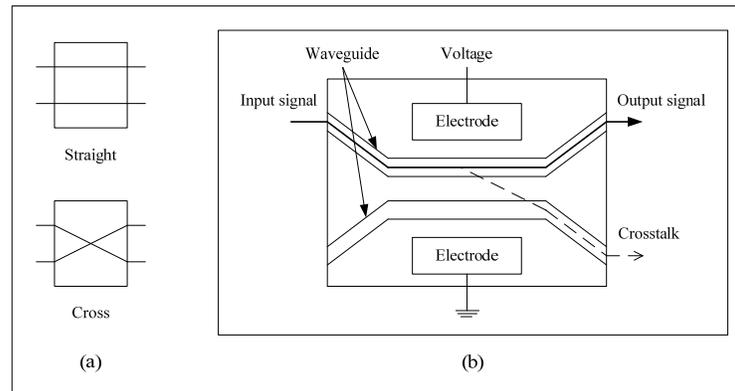

Fig. 2: (a) Straight or Cross Logic State of a 2x2 SE, and (b) Optical Crosstalk Effect in an Electro-Optic SE.

1.2 Time Domain Approach

Among the three approaches to avoid optical crosstalk, the space domain approach duplicates and combines an MIN to avoid crosstalk within individual SE. Using this approach, an *N* x *N* network is dilated into a network that is essentially equivalent to a *2N* x *2N* network, but only half of the input and output ports used for routing. Based on this idea, a dilated Benes network has been proposed [8], where up to *N* connections can be established without sharing any SE. However, it uses more than double of the number of switches required for the same connectivity, in the original Benes network. The wavelength domain [9] implements an approach where, the crosstalk between two signals passing through the same SE is suppressed by ensuring two wavelengths to be far apart by routing [10] or by using wavelength converters [11], which limits the efficiency of bandwidth utilization and/or increases cost and complexity [5].

Time-domain approach tackles the crosstalk problem by allowing only one input and output link to be active at a time within each SE in the architecture. A set of permutation connection is partitioned into several groups or passes, called semi-permutations in such a way that the elements within each group are crosstalk-free [11]. Then, each independent group is routed to its corresponding destination in a different time slot. In this paper, the focus is on how to efficiently realize permutations using the time domain approach in

order to avoid crosstalk in optical Omega network. Figure 3 shows the general framework of the time-domain approach.
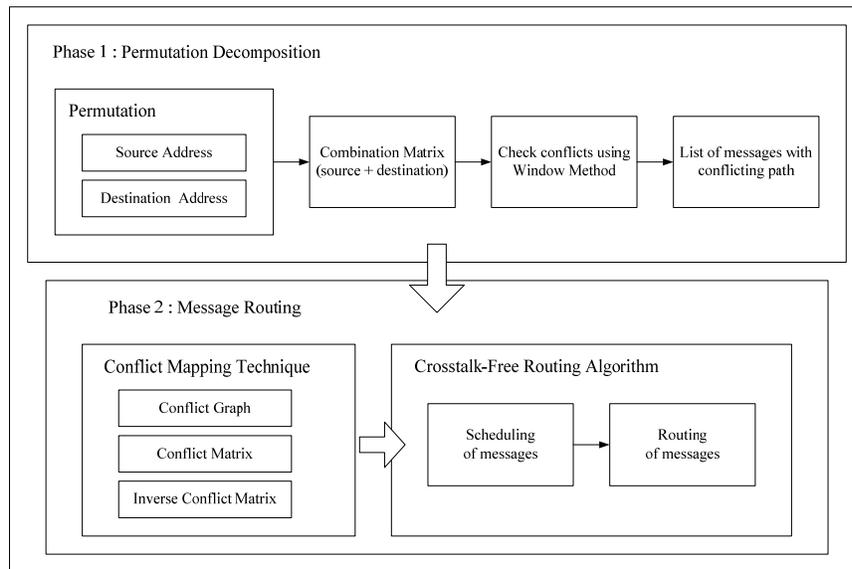

Fig. 3: Time-Domain Approach Framework.

In permutation decomposition phase, the primary goal is to identify conflicts among all messages in the network. The arbitrary source and destination address of the permutation is first combined to build a combination matrix. Then, based on the combination matrix, conflicts are checked using a pattern-checking method i.e. the Window Method (WM). Finally, messages with conflicting path between them are listed and mapped into an array representing the conflicts. In the following message routing phase, once the conflicts are discovered, it will be mapped into some array using one of the conflict mapping techniques. Based on the array of conflict generated, the messages are then sorted and selected for scheduling into crosstalk-free groups before routed to the intended destination one group per time slot.

2. PREVIOUS TIME-DOMAIN ALGORITHMS

Unlike electronic MINs where the main concern is merely avoiding path conflicts in the network, a different set of crosstalk avoiding algorithms have been proposed specifically for use in OMINs. From the literature, among the previous routing and scheduling algorithms that is based on the time domain approach are the four Heuristic algorithms i.e. Sequential Increasing, Sequential Decreasing, Degree Ascending and Degree Descending [13], Simulated Annealing (SA) [8], Genetic Algorithm (GA) [15], Ant Colony Optimization (ACO) [16], Remove Last Pass (RLP) [17], Zero algorithms [4] and Bitwise algorithms [18]. Each algorithm has the same pattern that is to select an input and an output pair from the permutation for the current pass according to some order, and if no crosstalk occurs, the pair is scheduled in the current pass; otherwise, schedule the

pair in the next pass. It is only in the order of which message to be scheduled that these algorithms primarily differ from each other.

In terms of performance wise, there exist trade-offs between the total number of passes generated and the time taken to completely route a permutation. The execution time is calculated beginning from generating conflict-free groups to routing all messages in the permutation. Previous researches have shown that the SA, GA, ACO and RLP algorithms give best results in the average number of passes produced for an arbitrary permutation connection. Alternatively, the four Heuristic algorithms, Zero algorithms and Bitwise algorithms performs best in the average execution time achieved by each algorithm. In this paper, the main objective is to design and develop conflict-free routing algorithms with better performance in terms of the execution time. From the literature, Zero algorithms is one of the algorithms that have been shown to perform best in execution time. The Bitwise algorithms have also been developed based on the Zero algorithms called the Bitwise Improved Zero (BIZ) algorithms [19].

3. THE FAST ZEROX ALGORITHM

Recent research has demonstrated that crosstalk may still occur when routing using the original ZeroX algorithm. In [16] the Improved ZeroX (IZ_X) algorithm is proposed to solve the problem, yet crosstalk is found to occur for some permutation. For example, consider the permutation for $N = 8$ as shown in Fig. 4.

```
000   ⟶   000
001   ⟶   010
010   ⟶   100
011   ⟶   110
100   ⟶   001
101   ⟶   011
110   ⟶   101
111   ⟶   111

Input      Output
```

Fig. 4: An 8 x 8 Source to Destination Permutation.

Based on the IZ_X algorithm, the messages will be scheduled in three passes including messages 000, 011, 101 and 110 for the first pass, messages 001, 010 and 100 for the second pass, and message 111 for the last pass. In the first pass, routing input messages 101 and 110 simultaneously will result in optical crosstalk since destination 101 and 100 for each respective input message is connected to the same switching element in the network. Throughout this section, the FastZ_X algorithm is shown to eliminate the abovementioned crosstalk completely. We describe the algorithm in two subsections with the first subsection explaining permutation decomposition methods followed by the algorithm in the second.

3.1 Permutation Decomposition

Because routing messages simultaneously causes optical crosstalk, it is important to make sure a permutation is decomposed in a crosstalk-free order. Permutation decomposition involves finding messages that may cause conflicts in the network and mapping these conflicts into some array before it can be selected for scheduling. The Bitwise Window Method (BWM) [18] is used to check conflicts between messages in the network. In BWM, each (n-1)-bit binary optical window where ($n = log_2N$) and N is network size, is simplified and converted into equivalent decimal representation using bitwise operations. As a result, the number of columns in each optical window is reduced to n, instead of 2*n for an N x N OMIN. Once the conflicts are identified, it can be mapped into the conflict graph [20] or the conflict matrix [4] to alleviate the scheduling procedure.

In this paper, we introduce a new pre-routing technique called the inverse Conflict Matrix (iCM) to map conflicts discovered using BWM. The iCM for the permutation in Fig. 4 can be constructed as in Fig. 5. The idea in iCM is extended from the conflict matrix used in previous Zero-based algorithms [4, 18], only that the iCM provides a more complete mapping of all conflicts in the network. Using iCM, scheduling is made simple and straightforward without prior column summation of the conflict matrix, which is a mandatory step in original ZeroX algorithm.

Destination Address	000	001	010	011	100	101	110	111
000	0	1	1	0	1	0	0	0
001	1	0	0	1	0	1	0	0
010	1	0	0	1	0	0	1	0
011	0	1	1	0	0	0	0	1
100	1	0	0	0	0	0	1	1
101	0	1	0	0	0	0	1	1
110	0	0	1	0	1	1	0	0
111	0	0	0	1	1	1	0	0

Fig. 5: The Inverse Conflict Matrix.

3.2 The Algorithm

Figure 6 shows the framework of the Fast Z_X algorithm. As indicated by its name, the algorithm is developed to optimally minimize the execution time of ZeroX algorithms. Based on our analysis, it is concluded that the original and IZ_X algorithm involves a lot of iterative procedures in the algorithm's functions for the summation of rows or columns, finding the intersections to refine the conflict matrix, multiple conflict-checking to ensure no conflicts when adding successful intersections to a group, reducing the matrix and calculating new summation for the reduced matrix. The advantage of the FastZ_X algorithm over the original algorithm is that based on the iCM a message can be determined whether it has conflicts and should not be scheduled in the same group directly

without prior column summation compulsory in ZeroX algorithm. Without the summation of the conflict matrix, the Unique Case and Refine functions in ZeroX algorithm were removed. Furthermore, the iCM need not be reduced to smaller matrix size since the iCM provides all the information needed for scheduling.

In terms of crosstalk avoidance, the FastZ_X algorithm successfully avoids crosstalk for any given permutation. Let the permutation as in Fig. 4. After the iCM in Fig. 5 is obtained, we select the first message 000 and scheduled it in the first pass for initialization. Then, to schedule the next message 001, check from the iCM, the intersection between messages 000 and 001. If the intersection has the value equals 0, then add message 001 to the current pass consisting message 000. Otherwise, if the intersection value equals 1, then go to the next pass and check the intersection value again but with the current entries in the pass. The process is repeated for the rest of the messages until all messages are scheduled.
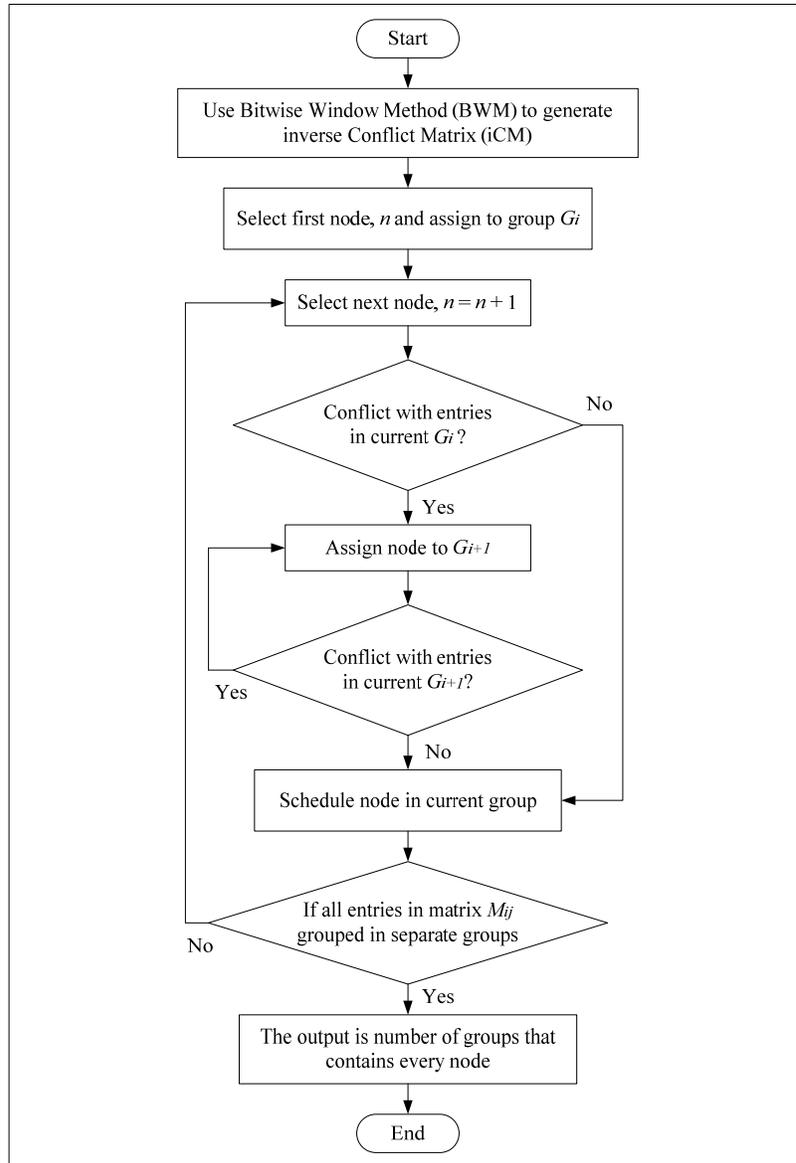
Fig. 6: Fast ZeroX Algorithm Framework.

In contrast to the results from the IZ_X algorithm discussed previously, applying the FastZ_X algorithm to the same permutation results in three conflict-free passes. The first pass contains messages 000, 011 and 101, the second pass contains messages 001, 010 and 100, and the third pass contains messages 110 and 111. The number of passes maybe the same but the messages that will be routed differs to avoid optical crosstalk in the new FastZ_X algorithm.

4. PERFORMANCE ANALYSIS

We present the test results for FastZ_X algorithm in this section. FastZ_X algorithm is then compared with previous Zero-based algorithms as well as the Heuristic's Sequential Increase algorithm since equal routing results can be obtained between all algorithms under test. Each algorithm is run 10,000 times and the results are presented in terms of the average number of passes as well as the average execution time. The results of the ZeroX algorithm presented in this paper refers to the result of a modified version of the algorithm, called the Modified ZeroX (MZ_X) algorithm developed for comparative analysis purposes. This is due to the fact that crosstalk may still occur in both the original and IZ_X algorithms. Therefore, in this paper the MZ_X will be regarded as the original ZeroX algorithm as we want the original algorithm to be crosstalk-free.

The MZ_X algorithms further extends the work in IZ_X algorithms and solved the optical crosstalk that occurs as a result from the algorithm's Unique Case function. To avoid the optical crosstalk, the MZ algorithms implemented two modifications to the function. First, the next entry that equals to zero in the same row or column of the group's unique entry, be checked for conflicts before it can be added to the group when the Unique Case function is executed. Second, the number of entries that can be added to the group is no longer limited to three entries only. In fact, a maximum of $N/2$ entries can be added to the group, where N is the size of the network. The rest of the MZ_X algorithms follows the framework of IZ_X algorithms i.e. the Refine function.
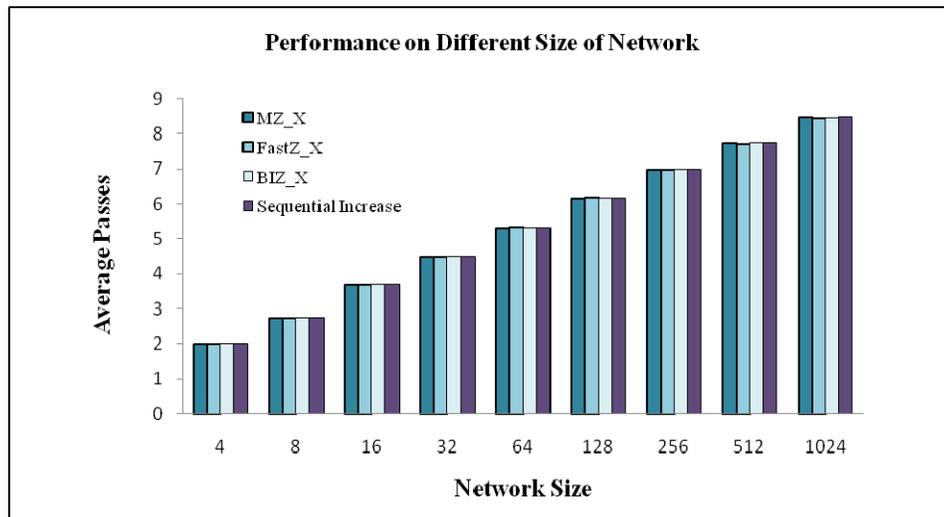

Fig. 7: Average Passes for FastZ_X Algorithm.

Figure 7 presents the average number of passes or conflict-free groups required to route a permutation between all algorithms. It is clear from the figure that there is almost no difference in the average number of passes realized by all the algorithms being tested. The slight difference shown is a result most probably due only to the different random source and destination of the permutation generated in each run. Otherwise, the number of passes realized is the same because the messages are selected based on the same order for scheduling starting from message 0 to message N where N is the size of the network.

In terms of the execution time, FastZ_X algorithm has shown consistent improvement in the average execution time compared to that of the original ZeroX algorithm (MZ_X algorithm) with 42% improvement as shown in Fig. 8. Furthermore, as can be seen, the larger the network size, the more significant the improvement in speed. Compared to the other previous algorithms such as the BIZ_X and Sequential Increase algorithm, FastZ_X outperforms all the algorithms with the best execution time achieved and an average of 16% improvement for both BIZ_X and Sequential Increase algorithms.
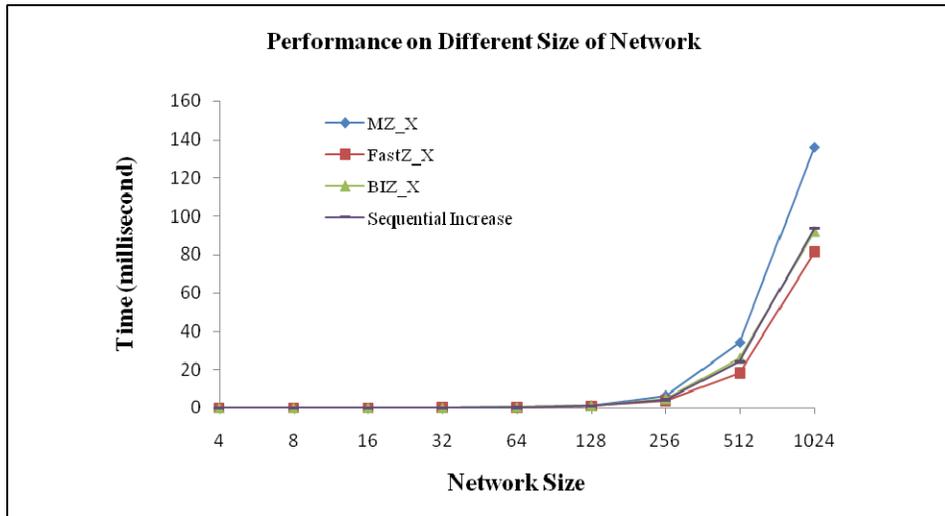

Fig. 8: Average Execution Time for FastZ_X Algorithm.

The advantage of the FastZ_X algorithm is that scheduling decisions is made directly based on the iCM simply by checking the intersections value between the message to be scheduled and the existing entries of the current pass. Through the use of iCM, the Refine and Unique Case function in the ZeroX algorithm is also removed in FastZ_X algorithm. Furthermore, as stated before, since the iCM provides the complete mapping of all conflicts in the network, prior summation of all columns in the conflict matrix is no longer necessary in the FastZ_X algorithm.

5. CONCLUSIONS

Throughout this paper, we have presented the Fast ZeroX algorithm, a new algorithm to improve the performance of scheduling and routing in OMINs without crosstalk. We also proposed the inverse Conflict Matrix or iCM to manage conflict mapping effectively. Through simulation modeling, FastZ_X algorithm was proven to significantly reduce the average execution time with the same number of crosstalk-free groups obtained using the original ZeroX algorithm. The new algorithm was also shown to achieve the best in execution time compared to other previous algorithms tested. As for future considerations, the iCM approach can be applied to other Zero-based algorithms as well as other time domain algorithms to reduce the execution time. The RLP algorithm can also be integrated with FastZ_X algorithm to improve the number of passes needed for routing based on the time-domain approach.

REFERENCES

[1]   J. D. Owens, W. J. Dally, R. Ho, D. N. Jayasimha, S. W. Keckler and L. S. Peh, "Research Challenges for On-Chip Interconnection Networks". IEEE Micro, Vol. 27, No. 5, pp. 96-108, 2007.

[2]   Y. Yang, J. Wang and Y. Pan, "Routing Permutations with Link-Disjoint and Node-Disjoint Paths in a Class of Self-Routable Interconnects". IEEE Transactions on Parallel and Distributed Systems, Vol. 14, No. 4, pp. 383-393, 2003.

[3]   D. Tustch and G. Hommel, "MLMIN: A Multicore Processor and Parallel Computer Network Topology for Multicast". Computers and Operations Research Journal, Elsevier, Vol. 35, No. 12, pp. 3807-3821, 2008.

[4]   M. A. Al-Shabi and M. Othman, "A New Algorithm for Routing and Scheduling in Optical Omega Network". International Journal of the Computer, The Internet and Management, Vol. 16, No. 1, pp. 26-31, 2008.

[5]   E. Lu and S. Q. Zheng, "High-Speed Crosstalk-Free Routing for Optical Multistage Interconnection Networks". Proceedings of the 12th International Conference on Computer Communications and Networks, pp. 249-254, 2003.

[6]   S. C. Chau, T. Xiao and A. W. C. Fu, "Routing and Scheduling for a Novel Optical Multistage Interconnection Networks". Euro-Par 2005 Parallel Processing, Lecture Notes in Computer Science, Vol. 3648, pp. 984-993, 2005.

[7]   C. L. Wu, and T. Y. Feng, "On a Class of Multistage Interconnection Networks". IEEE Transactions on Computers, Vol. 29, No. 8, pp. 694-702, 1980.

[8]   K. Padmanabhan and A. N. Netravali, "Dilated Networks for Photonic Switching". IEEE Transactions on Communications, Vol. 35, No. 12, pp. 1357-1365, 1987.

[9]   M. M. Vaez and C. T. Lea, "Space-Wavelength Tradeoff in the Design of Nonblocking Directional-Coupler-Based Networks Under Crosstalk Constraint". Journal of Lightwave Technology, Vol. 8, No. 16, pp. 1373-1379, 1998.

[10] J. Sharony, K. W. Cheung, and T. E. Stern, "The Wavelength Dilation Concept in Lightwave Networks – Implementation and System Considerations". Journal of Lightwave Technology, Vol. 11, No. 5/6, pp. 900-907, 1993.

[11] X. Qin and Y. Yang, "Permutation Capacity of WDM Switching Networks with Limited-Range Wavelength Conversion". IEEE Proceedings of 2001 International Conference on Parallel Processing Workshops on Optical Networks, pp. 271-276, 2001.

[12] Y. Yang, J. Wang and Y. Pan, "Permutation Capability of Optical Multistage Interconnection Networks". Journal of Parallel and Distributed Computing, Vol. 60, No. 1, pp. 72-91, 2000.

[13] . D. Shahida, M. Othman and M. Khazani, "Routing Algorithms in Optical Multistage Interconnection Networks: Revisited". Proceedings of the 3rd World Engineering Congress, pp. 63-70, 2007.

[14] A. K. Katangur, Y. Pan and M. D. Fraser, "Message Routing and Scheduling in Optical Multistage Networks Using Simulated Annealing". Proceedings of the International Parallel and Distributed Processing Symposium, pp. 201-208, 2002.

[15] T. W. Manikas and J. T. Cain, "Genetic Algorithms vs. Simulated Annealing: A Comparison of Approaches for Solving the Circuit Partitioning Problem". Technical Report 96-101, Department of Electrical Engineering, University of Pittsburgh, 1996.

[16] A. K. Katangur, S. Akkaladevi, Y. Pan and M. D. Fraser, "Applying Ant Colony Optimization to Routing Optical Multistage Interconnection Networks with Limited Crosstalk". Proceedings of the 18th International Parallel and Distributed Processing Symposium, pp. 163-170, 2004.

[17] S. C. Chau, and T. Xiao, "A New Algorithm for Routing and Scheduling in Optical Multistage Interconnection Networks". Proceedings of the 4th IASTED International Multi-Conference on Wireless and Optical Communications, Banff, Canada, pp. 749-755, 2004.

[18] F. Abed and M. Othman, "Fast Method to Find Conflicts in Optical Multistage Interconnection Networks". International Journal of The Computer, The Internet and Management, Vol. 16, No. 1, pp. 18-25, 2008.

[19] X. Qin and Y. Yang, "Permutation Capacity of WDM Switching Networks with Limited-Range Wavelength Conversion". IEEE Proceedings of 2001 International Conference on Parallel Processing Workshops on Optical Networks, pp. 271-276, 2001.

[20] M. Abdullah, M. Othman and R. Johari, "An Efficient Approach for Message Routing in Optical Omega Network". International Journal of the Computer, The Internet and Management, Vol. 14, No. 1, pp. 50-60, 2006.