# A New Routing Algorithm for a Three-Stage Clos Interconnection Networks

**Zahra sadat Ghandriz[1] and Esmaeil Zeinali Kh.[2]**

**[1] Department of Computer Engineering, Islamic Azad University, Qazvin Branch, QIAU, Iran**

**[2] Department of Computer Engineering, Islamic Azad University, Qazvin Branch, QIAU, Iran**

**Abstract**

Clos Interconnection network is one of the known connection networks in processing systems and distributed systems, which is used extensively in many fields such as Telecommunication networks, ATM switches and Data transmission. In order to eliminate the blocking in such networks, various routing algorithm have been proposed, each imposing extra costs due to hardware use and re-routing algorithm. This study offers a routing algorithm which takes a blocking-avoidance approach hence avoiding related costs. There is no blocking while the primary routing is performed from the input to output. This method has the complexity of $O(N\sqrt{N})$. The results show that this algorithm is simpler than the algorithms previously proposed.

***Keywords:*** *Clos Interconnection Networks, Routing Algorithm, Blocking Avoidance, Three Stage Interconnection Networks.*

## 1. Introduction

As one of the most important parts of parallel processing systems, interconnection networks make connection between switches [1, 2, 15]. The topologies typically found in processing systems are often regular. They are classified into two groups: statistic and dynamic. The networks using dynamic topology fall in the following categories of structures: bus networks, crossbar switching networks, multi-stage connection networks. Multistage interconnection networks consist of blocking networks, rearrangeable non-blocking networks, wide-sense non-blocking networks and strictly non-blocking networks [3, 4, 5, 6].

Routing algorithm can be divided into two groups: 1- blocking avoidance routing, 2- routing irrespective of blocking. In the first method, routing in the algorithm is performed in such a way that there is no blocking but in the second method, initially the routing is performed from input set to output set. In case of any blocking, attempts are made to eliminate it through changing the arrangements in network switching, using routing algorithm.

Routing algorithm in multi-stage connection can be performed by 2 methods: 1- Graph coloring algorithm, 2- decomposition algorithm using matrixes. Generally, graph coloring algorithm enjoys a better time complexity than decomposition algorithm, using matrix. However, as the network size is too large, it will be inefficient. Decomposition algorithm using matrixes has the following advantages: Directly locating the problems and enjoying simplicity in switching settings .Unfortunately, many proposed decomposition algorithms using matrix are incomplete. Neiman algorithm has been implemented, using decomposition matrix and time complexity $O(N\sqrt{N})$ [13]. GS algorithm with complexity $O(N\sqrt{N})$ uses 2 matrixes, namely: Specification and Count. But Siu, Chiu, Lee and Carpinelli believes GS algorithm to be incomplete [7, 13]. Later, a new method called the modified methods of GS was proposed [7]. This algorithm was proposed at time $O(N\sqrt{N}Log\sqrt{N})$ , by adding 3 steps to the main algorithm and deleting 2 steps for purpose of eliminating indefinite loops. Later, an algorithm on the basis of Heuristic Routing Algorithm using minimum distribution priority scheme was introduced for routing Clos networks. This method had the capability of accessing all non-blocking routings, reducing its time complexity to $O(N\sqrt{N})$ in the worst situation [14].

This study presents the routing mechanism in Clos interconnection networks, taking blocking-avoidance approach. This method can be applied extensively in distributed and parallel networks for purpose of

processing. Section 2 introduces the Clos networks and blocking-avoidance routing algorithm. Section 3 compares this new method with the previously proposed methods and section 4 concludes this study.

## 2. Blocking Avoidance Routing Algorithm in Clos Networks

### 2.1 Clos Network

Using small crossbar switches, Charles Clos introduced a type of interconnection network which is extensively studied and applied as a framework for ATM switches because it is economical, regular, scalable, fault-tolerant and highly efficient. Special attention has been paid to Clos three-stage networks as they are rearrangeable for developing multi-stage networks. These three stage-networks are intended to be used for data communication and parallel computing system.

A switching network is composed of one or more switch stages that can create various paths through creating various connections between their inputs and outputs. Clos three-stage network is an example of multi-stage switching networks (Fig.1).

Clos three-stage networks $N \times M$ are represented as $c(n_1, n_2, m, r_1, r_2)$ where N represents the overall inputs of network and M as overall outputs of network. $n_1, n_2, m, r_1, r_2$, represent the number of inputs of each switch of input stage, the number of outputs of each switch of output stage, number of switches of middle stage ,number of switches of input stage and number of switches of output stage, respectively. If $N = M$ then $r_1 = r_2$ , $n_1 = n_2$. In this case, Clos symmetrical networks are shown as $c(n, r, m)$ [15, 16]. The first stage of a three-stage network is called input stage which includes $r$ $(n \times m)$ switches. The second stage is called middle stage which includes $m$ $(r \times r)$ switches. The third stage is called output stage which includes $r$ $(m \times n)$ switches. The network is capable of connecting one to one and one to many with $N = nr$ and $m \geq n$ . There is a link between 2 switches in 2 continuous stages [7, 8]. As $c(n, r, m)$ knows all possible permutations between inputs and outputs. A link can be accessed between stages, provided it is usable and not engaged.

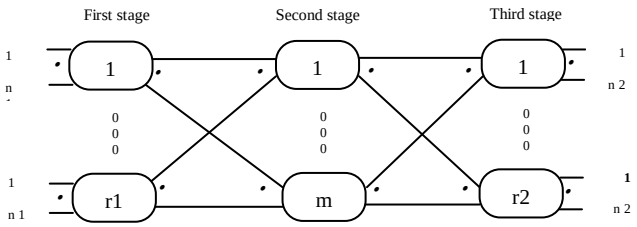


Fig. 1. Clos Three-stage network [15].

The Clos networks have the following advantages over other interconnection networks [8, 9, 10, 11, 12]:

1. This network is more efficient in terms of switching data.
2. The lower hardware costs of this network have drawn the attention of many researchers.
3. Clos networks have the capability to add switching packages in their configuration, increasing fault-tolerance compared to one-stage interconnection networks.

### 2.2 Vector Modeling of Clos Networks

The connections between each input and output can be shown as follows:

$$P = \begin{bmatrix} 0 & 1 & \cdots & i & \cdots & N-1 \\ \pi(0) & \pi(1) & \cdots & \pi(i) & \cdots & \pi(N-1) \end{bmatrix} \quad (1)$$

Where input i is connected to output $\pi(i)$ and $0 \leq i \leq N - 1$, $N = nr$ . The switch between input and output is assumed to be of non-blocking type. P may be changed between input switches and output switches [13].

In this algorithm, the matrix is used in such a way that network initially has its connections. Then, the new inputs are added into the network in such a way that there will be not any blocking. This is made possible through 4 main matrixes (A, B, C, D) and an mono dimensional array (e). The matrixes consist of elements 0, 1. 0 represents the free link and 1 represents the engaged link in the network.

Now the matrixes and array (e) are explained briefly:

1. Matrix A represents a connection between input switches and inputs of each input switch. The number of rows (i) represents input switches and the number of columns (j) represents the inputs of each input switch.
2. Matrix B represents the connection between middle switches and inputs of each middle switch. The number of rows (i1) represents the middle switches and the number of columns (j1) represents inputs of each middle switch.

3. Matrix C shows the connections between output switches and inputs of each output switch. The number of rows (i3) encompasses output switches and the number of columns (j3) represents the input of each output switch.
4. Matrix D indicates the connection between the outputs of each output switch and output switches. Here, the row (i4) represents output switches and columns (j4) represent the outputs of each output switch.
5. Array (e): each member of this array represents the number of the engaged input of each middle switch. The length of the array equals the number of middle stages.

**Definition:** All steps are complete when the elements of matrix A are transferred to matrix D. That is, all contents of input switches are added in the network and the input stage links are free for receiving new inputs.

### 2.3 Blocking-Avoidance Routing Algorithm

1. Initially, 0 is assigned to matrixes B, C, D and array (e).
2. Matrix A is derived from the inputs.
3. In matrixes B and A, a movement is made in parallel in such a way that in matrix A, the movement is made from row to row and in matrix B form column to column. When 1 is seen in matrix A, a free link is needed in middle switches to which the inputs are transferred. Then we turn to matrix B, searching for 0. On finding the first 0 in matrix B, we should replace it by 1 in matrix A.
4. Having transferred all engaged links of matrix A (input stage) to matrix B (middle stage) in such a way that there is no blocking in the input switches, now the output of middle stages should be arranged. To this purpose, the number of engaged input links into each middle switch is calculated, then added to the array (e). Having calculated all engaged links, we should assign the passive output links in middle switches (matrix B) to the engaged input links in middle switches (in array (e)).
5. We should arrange matrix B (middle stage) to matrix C (output stage) in parallel so that matrix B is examined in terms of columns and matrix C in terms of rows. On seeing 1 in matrix B, we should replace it by 0 in matrix C.
6. Finally, in order to transfer the engaged links to the output, we should make use of matrix D. That is, the movement in matrix C should be made from row to row and in metric D from the row to row in parallel. On seeing 1 in matrix C, it is replaced by 0 in matrix D.

Now, pseudo code of blocking-avoidance routing algorithm is presented as follows:

**Step1)** Initialize by Setting Middle Matrix $(B)$, $Output_1$ Matrix$(C)$, $Output_2$ Matrix$(D)$ and Middle array (e) = False.

**Step 2)** Read Input Matrix (A).

**Step 3)** For each of row elements of Matrix A, that is "True", if $(A[i,j] == 1)$ then {Find The First column elements of Matrix B, that is "False", if $(B[i_1,j_1] == 0)$ {Swap $(A[i,j], B[i_1,j_1]);$}}

**Step 4)** For each of row elements of Matrix B, that is "True", if $(B[i_1,j_1] == 1)$ then {S++;}e [k++]=S;
if $(e[k > 0])$ then { $B[k][j_{11}] = 1$ , k-- ;}

**Step 5)** For each of column elements of Matrix$(B)$, that is "True" , if $(B[i_1,j_1] == 1)$ then {Find The First row elements of Matrix C, that is "False", if $(C[i_3,j_3] == 0)$ {Swap $(B[i_1,j_1], C[i_3,j_3])$ ; }}

**Step 6)** For each of row elements of Matrix C , that is "True", if $C[i_3,j_3] == 1$ then { Find The First row elements of Matrix D, that is "False", if $(D[i_4,j_4] == 0)$ { Swap$(C[i_3,j_3], D[i_4,j_4]);$ }}

As you can see, all elements in matrixes A, B and C can be transferred to the next stage matrix in a parallel manner. This speaks for the parallel routing in the proposed manner.

## 3. Comparing blocking- avoidance routing algorithm with the previous methods

To compare the algorithm proposed by this study with the previous ones, we can examine them from two perspectives:

- The complexity of routing algorithm
- Time

### 3.1 The complexity of routing algorithm

The following should be taken into account to compute the complexity of routing algorithm:

- m = n, this indicates that the number of middle stage switches (m) equals the number of input ports of the switches of input stage (n).
- N= r×n, this indicates that overall network inputs are obtained by multiplying the number of switches of input stage (r) by the input ports of each switch of input stage (n).
- $n = \sqrt{N}$.

Each matrix is read by time $O(N)$ and the main body of algorithm is formed at $O(Nn)$ . thus, the overall complexity of proposed algorithm is :

$$O(Nn) + 3O(N) = O(Nn) \xrightarrow{n=\sqrt{N}} O\left(N\sqrt{N}\right)$$

Table1 shows the complexity of blocking-avoidance algorithm, drawing on the previously proposed algorithms.

Table 1: shows the complexity of blocking-avoidance algorithm in comparison with the previously proposed algorithms.

| Algorithm name | Complexity of Routing Algorithm |
|---|---|
| GS Algorithm | $N\sqrt{N}$ |
| GS Modification Algorithm | $N\sqrt{N}Log\sqrt{N}$ |
| Heuristic Algorithm | $N\sqrt{N}$ |
| Blocking-Avoidance Algorithm | $N\sqrt{N}$ |

As table 1 shows, the proposed routing algorithm outperforms the GS modified routing algorithm, with the former maintaining the network hardware. The proposed routing algorithm also equals Heuristic routing algorithm and GS routing algorithm in terms of algorithm complexity.

In GS routing Algorithm, there are indefinite loops which don't appear in the proposed routing algorithm in this study, resulting in higher efficiency of the proposed method.

In heuristic routing method, rearrangement is required. This leads to additional costs. While, in blocking avoidance approach, there is no such cost. This is because the blocking is eliminated while routing is being done.

Hence the better performance of this method over other methods.

### 3.2 Time

Regarding time, the algorithm proposed by this study outperforms the previously proposed methods. The previous methods start routing without taking account of blocking. That is, when routing is over, these methods make an attempt to eliminate the blocking through using some new algorithm, taking some time. This method proposed by this study starts routing so as to prevent the occurrence of any blocking, saving a lot of time.

## 4. Discussion and Conclusion

This study presents a routing mechanism in Clos interconnection networks, taking a blocking-avoidance approach so that the routing is performed properly from input to output without blocking the links in network. While in the previously proposed methods, the strictly non-blocking network was defined by simple routing which results in high hardware costs. Or alternatively, the routing was performed irrespective of blocking. Then, a new algorithm was used to eliminate these blocking, leading to some costs. The method proposed by this study has solved this problem. On the other hand, this algorithm completes its cycle at time $O(N\sqrt{N})$. This type of routing can be used in communications switching as well as in data transmission networks for purpose of reducing the delay in transference time and for controlling the network traffic. The future studies can examine the ways in which new algorithms can be used to reduce routing time and the use of memory and to reduce the complexity of algorithm.

## References

[1] H. Changa, J.Y. Guoa, F.K. Hwanga and C.K. Linb, "Wide-sense nonblocking for symmetric or asymmetric 3-stage Clos networks under various routing strategies", Theoretical Computer Science, No. 314, 2004, pp. 375–386.
[2] L.M. Ni and P.K. Mckinley, "A survey of wormhole routing techniques in direct network", IEEE Computer Society, VOL. 26, NO. 2, February 1993.
[3] P. Haynos and Y. Yang, "An Analytical Model on the Blocking Probability of a Fault Tolerant Network", IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEMS, VOL. 10, NO. 10, October 1999.
[4] Q. Ngo, "A new routing algorithm for multirate rearrangeable Clos networks", Theoretical Computer Science, No. 290, 2003, pp. 2157–2167.
[5] Holmberg, "Optimization Models for Routing in Switching Networks of Clos Type with Many Stages", An Electronic International Journal - AMO Advanced Modeling and Optimization, Volume 10, Number 1, 2008, pp. 1841-4311.

[6] S. Sharma, P.K. Bansal and K.S. Kahlon, “On a class of multistage interconnection network in parallel processing “, IJCSNS International Journal of Computer Science and Network Security, VOL.8, No.5, May 2008.
[7] F.K. Hwang, "A Modification to a Decomposition Algorithm of Gordon and Srikanthan", IEEE TRANSACTIONS ON COMPUTERS, VOL. 46, August 1997.
[8] X. Duan, D. Zhang and X. Sun, "Topology and Routing Schemes for Fault-tolerant Clos Network", International Conference on Networks Security, Wireless Communications and Trusted Computing, 2009.
[9] Y.Yang and J.Wang, “A New Design for Wide-Sense Nonblocking Multicast Switching Networks“, IEEE TRANSACTIONS ON COMMUNICATIONS, VOL. 53, NO. 3, March 2005.
[10] Y.Yang and J.Wang, "A Fault-Tolerant Rearrangeable Permutation Network", IEEE TRANSACTIONS ON COMPUTERS, VOL. 53, NO. 4, April 2004.
[11] Yuanyuan Yang, "A New Conference Network for Group Communication", IEEE TRANSACTIONS ON COMPUTERS, VOL. 51, NO. 9, September 2002.
[12] uanyuan Yang, "A Class of Multistage Conference Switching Networks for Group Communication", IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEMS, VOL. 15, NO. 3, March 2004.
[13] Y. Lee, F.k. Hwang and J.D. Carpinelli, "A New Decomposition Algorithm for Rearrangeable Clos Interconnection Networks", IEEE TRANSACTIONS ON COMMUNICATIONS,VOL. 44, NO. 11, November 1996.
[14] Duan and S. Liu, "A Heuristic Routing Algorithm for Clos Network", IEEE World Congress on Proceedings of the 7th, Intelligent Control and Automation, Chongqing, China, June 2008, pp. 25-27.
[15] Esmaeil Zeinali Kh. and Mehdi N. Fesharaki, "NMR: A New Approach for Optimal Desgin of Strictly Non-Blocking Multistage Interconnection Networks", Journal of Applied Sciences, Vol. 8, No. 19, 2008, pp. 3453-3459.
[16] Esmaeil Zeinali Kh., Mehdi N. Fesharaki and Mohammad Teshnehlab, "Optimal Design of Switch Sizes in Strictly Non-Blocking Clos Three-Stage Interconnection Networks", Journal of Applied Sciences, Vol. 8, No. 22, 2008, pp. 4234-4237.

**Zahra Sadat Ghandriz** was born in Yazd, Iran. She is currently a Master student in Hardware Engineering at Azad University of Qazvin, Iran. She received her B.Sc. degree in Software Engineering from Azad University, Maybod, Iran. Her research interests include Parallel Processing in Interconnection Networks and Parallel Algorithms.

**Esmaeil Zeinali Khasraghi** earned his B.Sc. degree in Computer Engineering (Software) from Qazvin Islamic Azad University (QIAU), and his M.Sc. and PhD degrees in Computer Engineering (Computer Architecture and Hardware) from Science and Research University, Tehran, Iran. He is an Assistant Professor of computer engineering and the QIAU Vice-President for Education since 2004. Working as a full-time faculty member of QIAU, he also directs the High Performance Computing and Networking Research Labs at the university. His research interests include Parallel and Distributed Systems, Optical and Wireless Networks and High Performance Computer Architecture.