EFFICIENT ROUTING ALGORITHM ON THE OTIS-NETWORKS

AHMAD M. AWWAD
Computer Science Dept.
Zarka Private University
ahmad_awwad@zpu.edu.jo

A. AL-AYYOUB
Dept. of Math. & Computer Sci.
Arab Open University
Amman, Jordan.

M. OULD-KHAOUA
Computer Science Dept.
Glasgow University
Glasgow, U.K.

Abstract

A number of recent studies have revealed that Optical Transpose Interconnection Systems (OTIS) are promising candidates as future high-performance parallel computers. In this paper we have proposed an efficient routing algorithm for Optical transpose Interconnection System. Some topological properties for the OTIS-Networks are derived including size, degree, diameter, and number of links. The proposed routing algorithm and the derived properties are general for all OTIS-Network, which will save the researchers effort to work on each on the OTIS-Networks individually. Two well-studied OTIS-Networks; OTIS-mesh and OTIS-hypercube are presented furthermore a broadcasting algorithm for OTIS-networks is discussed. This study provides new means for further testing the viability of the OTIS as alternative parallel computer architecture.

Index Terms: Optoelectronic systems, OTIS, interconnection networks, parallel algorithms.

1. INTRODUCTION

The choice of the network topology for a high-speed parallel computer is an important design decision that involves inherent trade-offs in terms of efficient algorithm support and network implementation cost. For instance, networks with large bisection width allow fast and reliable communication. However, such networks are difficult to implement using today's electronic technologies that are two dimensional in nature [10]. In principle, free-space optical technologies offer several fronts to improve this trade-off. The improved transmission rate, power consumption, and signal interference are few examples on these fronts [2, 10, 11, 12, 14].Optoelectronic and optical networking will become the key enabling technologies of the future communications infrastructure through the elimination of the difficult limitation of bandwidth and bit-error rate inherent in traditional electromagnetic signal-based communications [5, 6]. Electromagnetic signals carried over copper (or coaxial) wires suffer from loss of strength and are subject to errors due to noise and hence such systems have limited data rates [13]. When copper or coax is replaced by fiber technology the achievable bandwidth is in excess of 50 terabits/second with an almost zero bit-error rate [13]. The full implications of essentially huge bandwidth and extremely low loss rates are only beginning to be recognized and will radically reshape the future network technologies. While in the past the communication link was the bottleneck, this link now holds the potential to become the enabler of new modes of computing far beyond those existing today [13].

2. OPTICAL TRANSPOSE INTERCONNECTION SYSTEMS (OTIS)

This paper focuses on a specific optical interconnect, namely the Optical Transpose Interconnection Systems (OTIS). Marsden et al were the first to propose the OTIS [5]. A number of computer architectures have subsequently been proposed in which the OTIS were used to connect different processors [5]. Krishnamoorthy et al [12] have shown that the power consumption is minimized and the bandwidth rate is maximized when the OTIS computer is partitioned into N groups of N processors each. Zane et al [7] have limited their study to this type of the OTIS. In this paper, we focused on OTIS-networks where the number of processors in each group is equal to the number of groups; the terms OTIS-computer and OTIS-network refer to parallel architectures based on the OTIS and will be used interchangeably.

OTIS-networks are implemented using free-space optoelectronic technology [5]. In this model, processors are partitioned into groups, where each group is realized on a separate chip with electronic inter-processor connects. Processors on separate chips are interconnected through free space interconnects. The philosophy behind this separation is to utilize the benefits of both the optical and electronic technologies [18].

The advantage of using the OTIS as optoelectronic architecture lies in its ability to maneuver the fact that free space optical communication is superior in terms of speed and power consumption when the connection distance is more than few millimeters [12]. In the OTIS, shorter (intra-chip) communication is realized by electronic interconnects while longer (inter-chip) communication is realized by free space interconnects.

Extensive modeling results for the OTIS have been reported in [2]. The achievable terabit throughput at a reasonable cost makes the OTIS a strong competitor to the electronic alternatives [5, 12]. These encouraging findings prompt the need for further testing of the suitability of the OTIS for real-world parallel applications. A number of recent studies have been conducted in this direction [1, 3, 8, 9, 12, 15]. Sahni and Wang [8, 9] have presented and evaluated various algorithms on the OTIS-mesh and OTIS-hypercube including basic data rearrangements, routing, selection and sorting. They have also developed algorithms for various matrix multiplication operations [15, 16] and image processing [3]. Zane et al [7] have shown that the OTIS-mesh efficiently embeds four-dimensional meshes.

By using different electronic topologies, we arrive at different classes of OTIS-networks. For Instance, the OTIS-mesh and OTIS-hypercube are two different classes of the OTIS-networks where the mesh and hypercube topologies are used to realize the electronic interconnect. A brief description of the OTIS-mesh and OTIS-hypercube along with some of their topological properties is discussed in the following section.

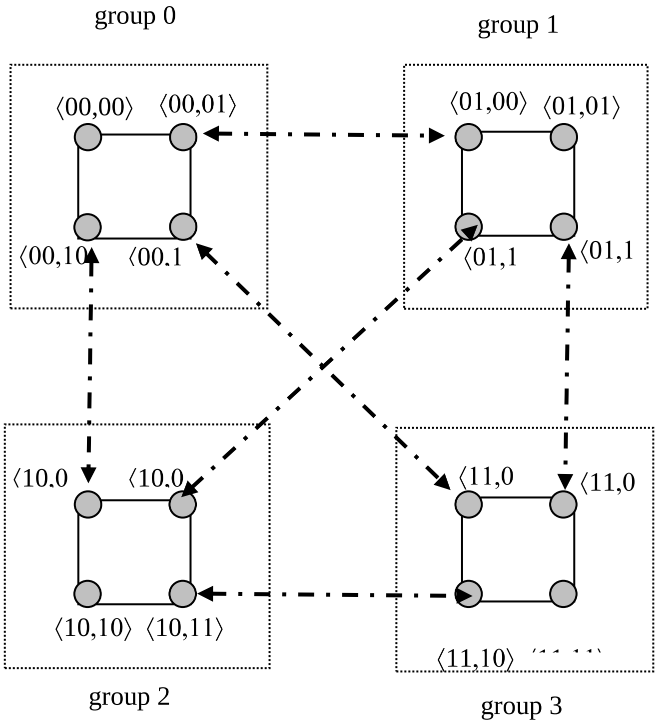

Figure 1: The OTIS-mesh with 16 processors.

3. THE OTIS-MESH AND OTIS-HYPERCUBE

The OTIS-mesh consists of an N2 processor; there are N groups and each group forming $\sqrt{N} \times \sqrt{N}$ mesh. In the OTIS-mesh, processors in the same group are connected as two-dimensional mesh [1, 5, 7]. Figure 1 shows a 16 processor OTIS-mesh, the notation ⟨g, p⟩ is used to refer to the group and processor addresses, respectively. Two nodes ⟨g1, p1⟩ and ⟨g2, p2⟩ are connected if, and only if, g1 = g2 and (p1, p2)∈E0 (such that E0 is the set of edges in the factor network) or g1 = p2 and p1 = g2 and in this case the tow nodes are connected by transpose edges.

1- The distance in the OTIS-mesh is defined as the shortest path between any two processors, ⟨g1, p1⟩ and ⟨g2, p2⟩, and involves one of the following forms [1]:

2- When g1 = g2 then the path involves only one electronic moves. ...(a)

3- When g1 ≠ g2 and if the number of optical moves is an even number of moves and more than two, then the paths can be compressed into shorter path of the form:

$\langle g1,\ p1\rangle \xrightarrow{E} \langle g1,\ p2\rangle \xrightarrow{O} \langle p2,\ g1\rangle \xrightarrow{E} \langle p2,\ g2\rangle \xrightarrow{O} \langle g2, p2\rangle$ . ...(b)

Where the symbols O and E stand for optical and electronic moves respectively.

4- When $g_1 \neq g_2$, the path involves an odd number of OTIS moves. In this case the paths can be compressed into a shorter path of the form:

$\langle g_1,\ p_1\rangle \xrightarrow{E} \langle g_1,\ g_2\rangle \xrightarrow{O} \langle g_2,\ g_1\rangle \xrightarrow{E} \langle g_2,\ p_2\rangle$. ...(*c*)

Theorem 1:

The length of the shortest path between any two processors $\langle g_1, p_1\rangle$ and $\langle g_2, p_2\rangle$ in OTIS-mesh is $d(p_1, p_2)$ when $g_1 = g_2$ and $\min\{d(p_1, p_2) + d(g_1, g_2) +2,\ d(p_1, g_2) + d(g_1, p_2) + 1\}$ when $g_1 \neq g_2$, where $d(p, g)$ stands for the shortest distance between the two processors *p* and *g* using any path *a*, *b* and *c* [1].

It is obvious from the above theorem that when $g_1 = g_2$, then the length of the path between the two processors $\langle g_1, p_1\rangle$ and $\langle g_2, p_2\rangle$ is $d(p_1, p_2)$. From the path construction methods in (*b*) and (*c*) above, it can be easily verified that the length of the path equal min $\{d\ (p_1, p_2) + d(g_1, g_2) +2,\ d(p_1, g_2) + d(g_1, p_2) + 1$ when $g_1 \neq g_2\}$.

Since each group in the OTIS-mesh is an $\sqrt{N} \times \sqrt{N}$ mesh, all the distances $d\langle p_1, p_2\rangle$, $d\langle p_2, g_1\rangle$, $d\langle p_1, g_2\rangle$ and $d\langle g_1, g_2\rangle \leq 2(\sqrt{N}$ -1). It follows that any two processors in the OTIS-mesh are at a distance at most $4(\sqrt{N}$ -1) +1 (= $4\sqrt{N}$ -3), which is the diameter of the OTIS-mesh [17].

The OTIS-hypercube is another example of OTIS-networks. In this class of networks the electronic interconnect follows the hypercube topology. The OTIS-hypercube consists of N^2 processors and *N* groups, each group is a hypercube of dimension $\log_2 N$ and of size equal $N=2^h$ (where h is the dimension of the hypercube). Figure 2 shows the topology of the OTIS-hypercube with 4 groups where each group consists of 4 processors. Two nodes $\langle g_1, p_1\rangle$ and $\langle g_2, p_2\rangle$ are connected if, and only if, $g_1 = g_2$ and $(p_1, p_2) \in E_0$ (such that E_0 is the set of edges in the factor network) or $g_1 = p_2$ and $p_1 = g_2$ and in this case the tow nodes are connected by transpose edges [9].

The length of the shortest path from processor i to the processor j in the hypercube is defined as d(i, j). If ⟨g1, p1⟩ and ⟨g2, p2⟩ are any two processors of the OTIS-hypercube, then the shortest path between these two processors follow the same path construction method as described above, i.e. paths of type a, b or c.

It can be easily shown that The diameter of the OTIS-hypercube is 2h+1 where h is the dimension of the hypercube [9].

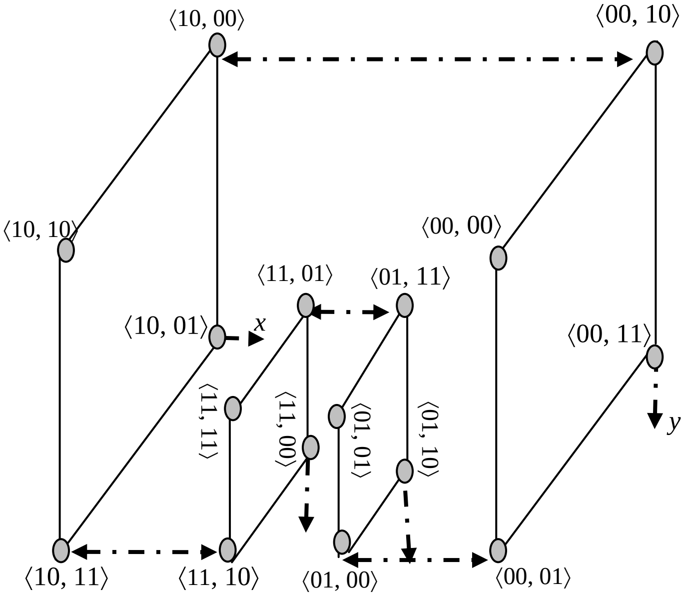

Figure 2: The OTIS-hypercube with 16 processors.

4. THE TOPOLOGICAL PROPERTIES OF THE OTIS-NETWORKS

This section discusses and derives some of the basic topological properties of the OTIS-networks including size, degree, diameter, number of links, broadcasting, and routing. Table 1 summarises the topological properties of the OTIS-star, OTIS-hypercube and OTIS-mesh along with their electronic counterparts star, hypercube and mesh. The symbols *g* and *p* in the table stands for the group and processor addresses respectively. Let us discuss and derive the properties that are itemised in Table 1.

Size: If G_0 is the factor network of size N, then the size of the OTIS-G_0 is N^2.

Degree*:* Let $\langle g, p\rangle$ be any node in OTIS-G0. Then the degree (or deg) of the OTIS-G0 is as follows:

$$deg_{\text{OTIS-}G_0}(g, p) = \begin{cases} deg_{G_0}(p) & \text{if } g = p \\ deg_{G_0}(p) + 1 & \text{if } g \neq p \end{cases}$$

Number of Links: Let N_0 be the number of links in the factor network and let M be the number of nodes in the G_0. The number of links in the OTIS-G_0 = $(M^2 - M)/2 + N_0 \cdot M$. For instance, the number of links in the OTIS-mesh consisting of 16 processors is $(4^2 - 4)/2 + 4 \cdot 4 = 12/2 + 16 = 22$.

Diameter: Let $\langle g_1, p_1\rangle$ and $\langle g_2, p_2\rangle$ be two different processors in the OTIS-G_0. To transmit data originated in the source node $\langle g_1, p_1\rangle$ to the destination node $\langle g_2, p_2\rangle$ there are 3 possible paths to follow: *a*, *b* and *c*. These paths are in the following forms:

1- When $g_1 = g_2$, then the path employs only electronic moves.

Path *a*: $\langle g, p_1\rangle \xrightarrow{E} \langle g, p_2\rangle$

2- When $g_1 \neq g_2$ then the path employs 2 electronic moves and 2 optical moves.

Path *b*: $\langle g_1, p_1\rangle \xrightarrow{E} \langle g_1, p_2\rangle \xrightarrow{O} \langle p_2, g_1\rangle \xrightarrow{E} \langle p_2, g_2\rangle \xrightarrow{O} \langle g_2, p_2\rangle$.

3- Or when $g_1 \neq g_2$ then the path employs 2 electronic moves and one optical move.

Path *c*: $\langle g_1, p_1\rangle \xrightarrow{E} \langle g_1, g_2\rangle \xrightarrow{O} \langle g_2, g_1\rangle \xrightarrow{E} \langle g_2, p_2\rangle$.

To send a message M from the source node $\langle g_1, p_1\rangle$ to the destination node $\langle g_2, p_2\rangle$ it must follow a route along one of the three possible paths *a*, *b*, and *c*. The length of the shortest path between the nodes $\langle g_1, p_1\rangle$ and $\langle g_2, p_2\rangle$ is one of the forms:

$$\text{Length} = \begin{cases} d(p_1, p_2) & \text{if } g_1 = g_2 \\ min(\, d(p_1, g_2) + d(g_1, p_2) + 1,\ d(p_1, p_2) + d(g_1, g_2) + 2) & o.w. \end{cases} \quad \text{...(I)}$$

Where $d(p_1, p_2)$ is the length of the shortest path between any two processors $\langle g_1, p_1\rangle$ and $\langle g_1, p_2\rangle$. If δ_0 is the diameter of the factor network, G_0 then from (1) it follows that the diameter of the OTIS-G_0 is $2\delta_0 + 1$.

5. ROUTING

Routing is an important issue when considering a new interconnection network [4]. We assume that a routing algorithm in the factor network, G_0, is known and given with a function $Next_{G_0}(p_1, p_2)$; this function returns the forwarding node after p_1 in routing towards p_2. Also a function $Dist._{G_0}(p_1, p_2)$ returns the minimum distance between the two processors p_1 and p_2 in G_0. The message (M) is originated at the source node $\langle g_1, p_1\rangle$ and to be transmitted to the destination node $\langle g_3, p_3\rangle$. Figure 3 summarises the routing algorithm in the OTIS-G_0.

In Figure 3 the parameter M, stand for the message to be transmitted from the source node, while the parameters g_1, g_2, and g_3 are the source group, current group, and the destination group, respectively. Furthermore the parameters p_1, p_2, and p_3 stand for the source processor, current processor and the destination processor. The node $\langle g_1, p_1\rangle$ is defined to be the source node, $\langle g_2, p_2\rangle$, is the current node where the message has arrived and $\langle g_3, p_3\rangle$ is the destination node.

When routing between the source node $\langle g_1, p_1\rangle$ and destination node $\langle g_3, p_3\rangle$, the distance between the source and destination nodes is in one of the forms: -

- If $g_1 = g_3$ then minimum distance between the source node and the destination node is $d = Dist_{G_0}(p_1, p_3)$. and the source node will call the function $Route_{\text{OTIS-}G_0}(M, \langle g_1, p_1\rangle, \langle g_2, p_2\rangle, \langle g_3, p_3\rangle, a)$ to route the message.
- If $g_1 \neq g_3$ then $d_1 = Dist_{G_0}(p_1, p_3) + Dist_{G_0}(g_1, g_3) + 2$ or $d_2 = Dist_{G_0}(p_1, g_3) + dist_{G_0}(g_1, p_3) + 1$.

If $d_1 < d_2$ then the source node will call the function

$Route_{\text{OTIS-}G_0}(M, \langle g_1, p_1\rangle, \langle g_2, p_2\rangle, \langle g_3, p_3\rangle, b)$

else

$Route_{\text{OTIS-}G_0}(M, \langle g_1, p_1\rangle, \langle g_2, p_2\rangle, \langle g_3, p_3\rangle, c)$.

```
Function Route_{OTIS-G_0} $(M, \langle g_1, p_1\rangle, \langle g_2, p_2\rangle, \langle g_3, p_3\rangle$, path type)
Begin {function}
  if $\langle g_2, p_2\rangle \neq \langle g_3, p_3\rangle$ then
        {where $\langle g_2, p_2\rangle$ is the current node and $\langle g_3, p_3\rangle$ is the destination node}
    if path = a then send $(M, \langle g_1, p_1\rangle, Next_{G_0}(p_2, p_3), \langle g_3, p_3\rangle, a)$
    else if path = b then
        {path b: $\langle g_1, p_1\rangle \xrightarrow{E} \langle g_1, p_2\rangle \xrightarrow{O} \langle p_2, g_1\rangle \xrightarrow{E} \langle p_2, g_2\rangle \xrightarrow{O} \langle g_2, p_2\rangle$}
          if $g_2 = g_1$ then
            if $p_2 \neq p_3$ send $(M, \langle g_1, p_1\rangle, Next_{G_0}(p_2, p_3), \langle g_3, p_3\rangle, b)$
            else send $(M, \langle g_1, p_1\rangle, \langle p_3, g_2\rangle, \langle g_3, p_3\rangle, b)$
          else
            if $g_2 \neq g_3$ then send $(M, \langle g_1, p_1\rangle, Next_{G_0}(g_2, g_3), \langle g_3, p_3\rangle, b)$
            else send $(M, \langle g_1, p_1\rangle, \langle g_3, p_3\rangle, \langle g_3, p_3\rangle, b)$
    else if path = c then {path c: $\langle g_1, p_1\rangle \xrightarrow{E} \langle g_1, g_2\rangle \xrightarrow{O} \langle g_2, g_1\rangle \xrightarrow{E} \langle g_2, p_2\rangle$}
          if $g_2 = g_1$ then
            if $p_2 \neq g_3$ then send $(M, \langle g_1, p_1\rangle, Next_{G_0}(p_2, g_3), \langle p_3, g_3\rangle, c)$
            else send $(M, \langle g_1, p_1\rangle, \langle g_3, g_2\rangle, \langle g_3, p_3\rangle, c)$
          else
            if $g_2 \neq p_3$ then send $(M, \langle g_1, p_1\rangle, Next_{G_0}(g_2, p_3), \langle g_3, p_3\rangle, c)$
  else
    stop, destination reached
End {function}. be observed in Figure 3.
```

Figure 3: The routing algorithm in the OTIS-G_0 using one of the paths *a, b* or *c*

***Broadcasting*: ** Data broadcasting is one of the most important operations required by many parallel applications [4]. Assume that the data is initially in the source node $\langle g_1, p_1\rangle$ where g_1 is the group index and p_1 is the processor index, and is to be broadcasted to all of the nodes N^2 in the OTIS-G_0. The data can be transmitted to all processors in the different groups using the following algorithm steps:

Step 1: The source processor $\langle g_1, p_1\rangle$ broadcasts the data to all the processors in its group using electronic moves.

Step 2: Each processor in the group then broadcasts the data to one processor in each group in the network using one OTIS move.

Step 3: Each processor in the group broadcasts the data again to the remaining processors using electronic moves.

In the above algorithm, following Step 1 each processor of the group containing the source processor will receive a copy of the data from the source processor using electronic move. After Step 2, one processor from each group in the network will receive a copy using OTIS move. In Step 3, each processor of the OTIS-G_0 will obtain a copy of the data using electronic moves. Step 1 and 3 will take one electronic move and Step 2 will take one OTIS move. We can also conclude that the cost of broadcasting in the OTIS-G_0 is $2 \times \beta_{G_0} + 1$ (*i.e.* two electronic moves plus one optical move) where β_{G_0} is the cost of broadcasting in G_0.

6. CONCLUSIONS

The major contribution of this paper is proposing an efficient routing algorithm for OTIS-Networks, furthermore a broadcasting algorithms has been discussed. The proposed routing algorithm is general in the sense that no factor network is assumed. We have presented two well-studied OTIS-networks OTIS-mesh and OTIS-hypercube. Some of the basic topological properties such as the degree and diameter of the OTIS-mesh and OTIS-hypercube have been discussed. Moreover, some of the general properties of OTIS-networks have been derived including the size, degree, diameter, and number of links. Our Future work is to extend our investigation into more other important issues in the OTIS-Networks including fault tolerance, and embedding some of the well known algorithms like Gaussian elimination, matrix computation problems, Fourier transforms.

REFERENCES

[1] S. Sahni and C. Wang, "BPC Permutations on the OTIS-mesh Optoelectronic Computer," *Proceedings of the IEEE Conference on Massively Parallel Programming using Optical Interconnect*, 1997, pp. 130-135.

[2] W. Hendrick , O. Kibar, P. Marchand, C. Fan, D. Blerkom, F. McCormick, I. Cokgor, M. Hansen, and S. Esener, "Modeling and Optimization of the Optical Transpose Interconnection System," *In Optoelectronic Technology Centre*, Program Review, Cornell University, Sept. 1995.

[3] C. Wang and S. Sahni, "Image Processing on the OTIS-mesh Optoelectronic Computer," *IEEE Trans. Parallel and Distributed Systems*, vol. 11, no. 2, 2000, pp. 97-109.

[4] K. Day and A. Tripathi, "Arrangement Graphs: A Class of Generalized Star Graphs," *Information Processing Letters*, vol. 42, 1992, pp. 235-241.

[5] G. Marsden, P. Marchand, P. Harvey, and S. Esener, "Optical Transpose Interconnection System Architecture," *Optics Letters*, 18(13), 1993, pp. 1083-1085.

[6] W. Hendrick , O. Kibar, P. Marchand, C. Fan, D. Blerkom, F. McCormick, I. Cokgor, M. Hansen, and S. Esener, "Modeling and Optimization of the Optical Transpose Interconnection System," *In Optoelectronic Technology Centre*, Program Review, Cornell University, Sept. 1995.

[7] F. Zane, P. Marchand, R. Paturi, and S. Esener, "Scalable Network Architecture Using the Optical Transpose Interconnection System (OTIS)," *Journal of Parallel and Distributed Computing* 60(2000), pp. 521-538.

[8] C. Wang and S. Sahni, "Basic Operations on the OTIS-mesh Optoelectronic Computer," *IEEE Trans. Parallel and Distributed Systems*, vol. 9, no. 12, 1998, pp.1226-1236.

[9] S. Sahni and C. Wang, "BPC Permutations on the OTIS-hypercube Optoelectronic Computer," *Informatica*, vol. 22, 1998, pp. 263-269.

[10] W.J. Dally, "Performance Analysis of *k*-ary *n*-cubes Interconnection Networks," *IEEE Trans. Computers*, vol. 39, no. 6, 1990, pp. 775-785.

[11] G. Yayla, P. Marchand, and S. Esener, "Speed and Energy Analysis of Digital

Interconnections: Comparison of on-chip, off-chip, and Free-space Technologies," *Applied Optics*, vol. 37, no. 2, 1998, pp. 205-227.

[12] A. Krishnamoorthy, P. Marchand, F. Kiamilev, and S. Esener, "Grain-size Considerations for Optoelectronic Multistage Interconnection Networks," *Applied Optics*, vol. 31, no. 26, 1992, pp. 5480- 5507.

[13] S. Chatterjee and S. Pawlowski, "Enlightening the Effects and Implications of Nearly Infinite Bandwidth," *Communications of the ACM*, vol. 42, no.6, 1999, pp. 75-83.

[14] S. Sahni, "Models and Algorithms for Optical and Optoelectronic Parallel Computers," *Proceedings of the International Symposium on Parallel Algorithms and Networks*, IEEE Computer Society Press, 1999, pp. 2-7.

[15] C. Wang and S. Sahni, "Matrix Multiplication on the OTIS-mesh Optoelectronic Computer," *Proceedings of the 6th IEEE International Conference on Parallel Interconnects*, 1999, pp. 131-138.

[16] C. Wang and S. Sahni, " Matrix Multiplications on the OTIS-Mesh Optoelectronic Computer," *IEEE Trans. Computers*, vol. 40, no. 7, July 2001, pp. 635-646.

[17] S. Sahni and C. Wang, "BPC Permutations on the OTIS-mesh Optoelectronic Computer," Technical Report 97-008, CISE department, University of Florida, 1997.

[18] A. M. Awwad, A. Al-Ayyoub, M. Ould-Khaoua, and K. Day, "Solving Linear Systems Equations Using the Grid Structural Outlook," Proceedings of the 13th IASTED Parallel and Distributed Computing and Systems (PDCS' 2001), Anaheim, USA, August 21-24, 2001, pp. 365-369.

Ahmad Awwad finished his B.Sc. in Computer Science 1987 and his M.Sc. degree from Tennessee State University, USA, 1989. He gained his Ph.D. in Computer Science from university of Glasgow, United Kingdom, 2001. His research interests are in parallel processing, interconnection networks, product networks, opto-electronic networks, Algorithms Design for HSPC, and vertex product networks. Dr. Awwad is currently the head of computer science department at Zarka private university. Also he is a member of general secretariat for Arab Conference on Information Technology (ACIT).