SCALABLE COMPUTING: PRACTICE AND EXPERIENCE
Volume 9, Number 1, pp. 69–75. http://www.scpe.org
ISSN 1895-1767

AN MIMD ALGORITHM FOR FINDING MAXIMUM AND MINIMUM ON OMTSE ARCHITECTURE

S. C. PANIGRAHI* AND G. SAHOO*

Abstract. Optical Multi-Trees with Shuffle Exchange (OMTSE), recently proposed, is an efficient model of optoelectronic parallel computer. The OMTSE interconnection system consists of n^2 factor networks called TSE networks, which are organized in the form of an $n \times n$ matrix. Each factor network has n number of leaf nodes. The network has a total of $3n^3/2$ nodes. The diameter and bisection width of the network is $6 \log n - 1$ and $n^3/4$ respectively. In this paper we present a synchronous MIMD algorithm to find the maximum and minimum of $n(n-1)^2$ data elements on OMTSE with $(4 \log n + 4)$ electronic moves and 3 optical moves.

Key words. parallel processing, time complexity, optical multi-trees with shuffle exchange, OTIS mesh, optoelectronic computers.

1. Introduction. The interconnection network is the heart of a parallel processing system, and many systems have failed to meet their design goals for the design of their essential components. The performance of most digital systems today is limited by their interconnection network bandwidth. The bandwidth limitation of the electronic interconnects prompted the need for exploring alternatives that overcome this limitation. Optics is considered as an alternative that is capable of providing inherent communication, parallelism, high connectivity and large bandwidth. It is well known that when the communication distances exceed few millimeters, optical interconnects provide advantage over the electronic interconnects in term of power, speed and crosstalk property [1, 10]. Therefore, in the construction of very powerful and large multiprocessor systems, it is advantageous to interconnect close processors (with in same physical package, e.g. chip) physically using electronic links and far processors (kept in other package) using optical links. Motivated by these observations new hybrid computer architecture utilizing both optical and electronic technologies have been proposed and investigated in [9, 8].

Marsden et al. [9], Hendrick et al. [14] and Zane et al. [8] have proposed an architecture in which the processors are partitioned into groups where processors within each group are connected by electronic links, while those in different groups are connected using optical interconnections. The Optical Transpose Interconnect System (OTIS) proposed by Marsden et al. [9] is an example of such hybrid architecture in which processors are partitioned into groups of same size, and processor i of group j is connected to the processor j of group i via an optical link. Krishnamoorthy et al. [1] have shown that when number of processors in each group is equal to the total number of groups, then bandwidth and power efficiency are maximized, and system area and volume are minimized.

The OTIS mesh optoelectronic computer is a class of OTIS computers where processor in each group are interconnected by electronic links followings the two-dimensional mesh paradigm. An N-processor OTIS-Mesh [4] has a diameter of $4N^{1/4} - 3$. Mapping algorithms of various fundamental problems occurring in real-life applications on the OTIS-Mesh has been studied by several authors, e.g., Wang and Sahni [4, 5, 6, 7], Rajasekarna and Sahni [13], Osterloh [1].

Optical Multi-Trees with Shuffle Exchange (OMTSE) is a new interconnection network proposed by Panigrahi et al. [11] for optoelectronic parallel computers. The network consists of a total of $3n^3/2$ processors are built around n^2 factor networks called TSE networks. Each factor network consists of n leaf nodes. The diameter and bisection width of the OMTSE network is shown to be $6 \log n - 1$ and $n^3/4$. Basic broadcast operations and several parallel algorithms including summation, prefix computation, matrix transpose, matrix multiplication, sorting have shown to map on OMTSE [11, 12] in lesser time than the OMULT [3].

In this paper we have introduced a synchronous MIMD algorithm for finding maximum and minimum of m elements on OMTSE where $m = n(n-1)^2$. All processors of OMTSE interconnection system operate in a lock step fashion. The algorithm has shown to run in $(4 \log n + 4)$ electronic moves and 3 optical moves. For the purpose of the complexity analysis of our algorithm, we count the data moves along the electronic links (known as electronic moves) and optical links (known as optical moves) separately.

The rest of the paper is organized as follows. The topology of the network is described in section 2. The proposed algorithm is described in the section 3 followed by the feasibility analysis and conclusion presented in section 4 and section 5 respectively.

2. Topology of OMTSE. The factor network used in OMTSE topology is a two layer TSE (Trees with Shuffle Exchange) network, which is nothing but an interconnection network containing a group of 2^r, $r \geq 1$, complete binary trees of height one and the roots of these binary trees are connected with Shuffle-Exchange fashion. The proposed OMTSE

*Deaprtment of Computer Science and Engineering, Birla Institute of Technology, Mesra, Ranchi 835215, India (satish.panigrahi@gmail.com, gsahoo@bitmesra.ac.in).

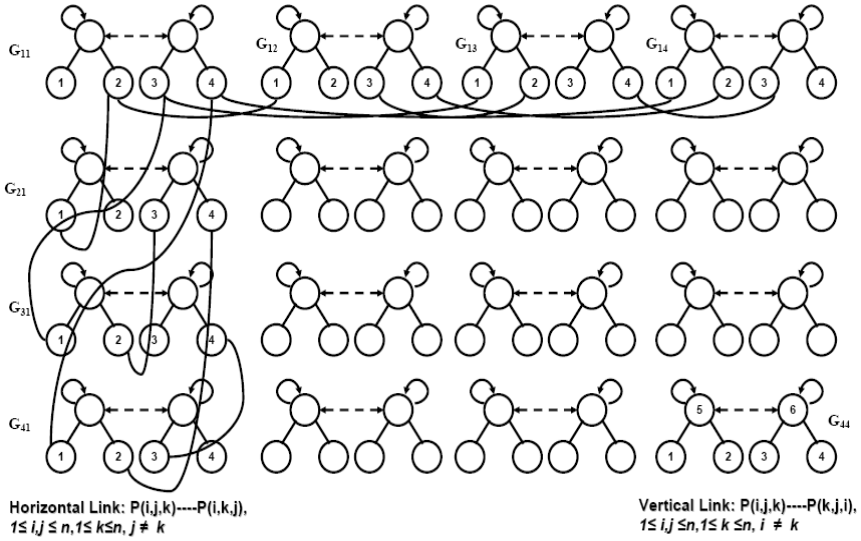

FIG. 2.1. *An example of OMTSE topology with* $n = 4$.

interconnection system consists of n^2 TSE networks, which are organized in the form of an $n \times n$ matrix. We denote the TSE network placed at i^{th} row and j^{th} column of this matrix by G_{ij}, $1 \leq i, j \leq n$. Each TSE network having n nodes at layer 2 and $n/2$ nodes at layer 1. There are, therefore, $N = 3n^3/2$ processors in total. The nodes within each TSE network are interconnected by usual electronic links, while the nodes at layer 2 (i. e. leaf processors) of different TSE networks are interconnected by optical links according to the rules given below. Let us label the nodes in each TSE network G_{ij}, $1 \leq i, j \leq n$, by distinct integers from 1 to $3n/2$ in reverse order, i. e., the nodes at layer 2 of TSE network are numbered from 1 to n in order from left to right, and nodes at layer 1 also numbered from left to right. The node k in a TSE network G_{ij} will be referred to by the processor $P(i, j, k)$, $1 \leq i, j \leq n$, $1 \leq k \leq 3n/2$. The optical links interconnecting only the leaf nodes in different TSE networks are used in the following way

1. Processor $P(i, j, k)$, $1 \leq i, j \leq n$, $1 \leq k \leq n$, $j \neq k$, is connected to the processor $P(i, k, j)$ by bidirectional optical link called horizontal inter-TSE link.

2. Processor $P(i, j, k)$, $1 \leq i, j \leq n$, $1 \leq k \leq n$, $i \neq k$, is connected to the processor $P(k, j, i)$ by bidirectional optical link called vertical inter-TSE link.

The diameter of a network is the maximum distance between any two processing nodes in the network. Hence starting from a node $P(i, j, k)$, $1 \leq i, j \leq n$, $1 \leq k \leq n$ we can reach another node $P(i', j', k')$, $1 \leq i', j' \leq n$, $1 \leq k' \leq n$ of the OMTSE interconnection system by traversing the following path

$$P(i, j, k) \rightarrow P(i, j, j') \rightarrow P(i, j', j) \rightarrow P(i, j', i') \rightarrow P(i', j', i) \rightarrow P(i', j', k')$$

It can easily be seen that the diameter of OMTSE topology is $6 \log n - 1 = O(\log n)$, with $6 \log n - 3$ electronic links and 2 optical links. Similarly, we can find out the bisection width of OMTSE topology which is equal to $n^3/4$. An example of OMTSE topology for $n = 4$ with partial links is shown in Fig. 2.1.

3. Proposed Algorithm. We assume that each processor $P(i, j, 1)$, $1 \leq i, j \leq n$ has two registers $A(i, j, 1)$ and $B(i, j, 1)$, where as the rest of the processors has only one register $A(i, j, k)$, $1 \leq i, j \leq n$, $2 \leq k \leq 3n/2$. Suppose we have $m = n(n-1)^2$ data elements stored in A-register of n leaf nodes of all TSE networks excluding the TSE networks in the first row and first column of OMTSE. Hence each factor networks excluding the factor networks in the first row and in the first column of OMTSE contain n data elements in its leaf nodes. We can find the maximum and minimum of these elements in $(4 \log n + 4)$ electronic moves and 3 optical moves. The steps of the proposed algorithm have been described below, where we refer the left child and right child of the processor $P(i, j, k)$ as $\text{LCHILD}(P(i, j, k))$ and $\text{RCHILD}(P(i, j, k))$ respectively.

Algorithm: (ALGO_MAX_MIN). Initial Condition: The m elements stored in A-registers of $P(i, j, k)$, $2 \leq i, j \leq n$, $1 \leq k \leq n$ and initialize $A(i, 1, 1)$ as zero for $1 \leq i \leq n$, and $A(1, j, 1)$ as ∞ for $2 \leq j \leq n$.

1. For all i, j, k, $2 \leq i, j \leq n$ and $n+1 \leq k \leq 3n/2$, do in parallel $A(i, j, k) \leftarrow \max\big(A(LCHILD(P(i, j, k))) + A(RCHILD(P(i, j, k)))\big)$

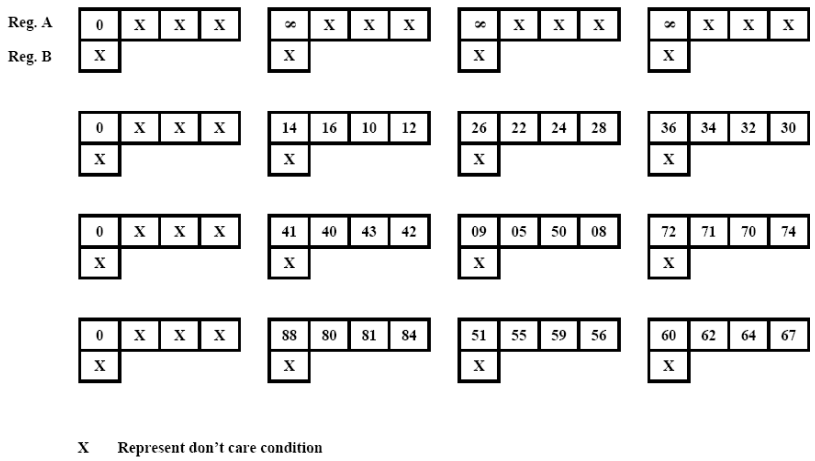

FIG. 3.1. *Initial Distribution of Data Elements.*

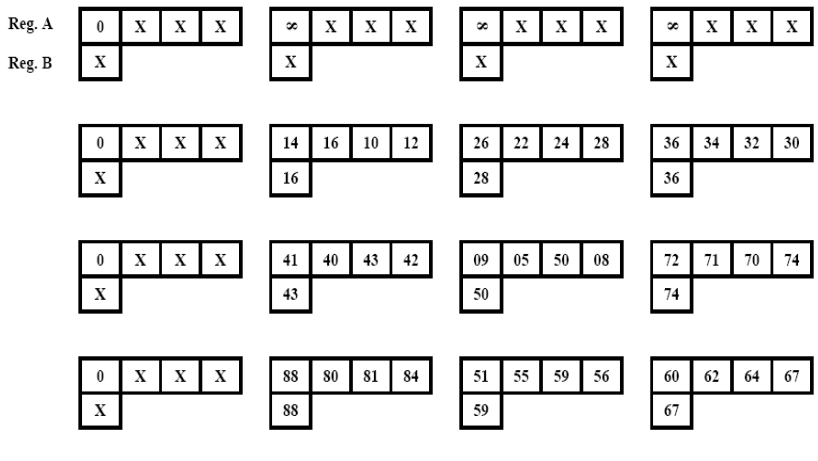

FIG. 3.2. *Data Distribution after step 3.*

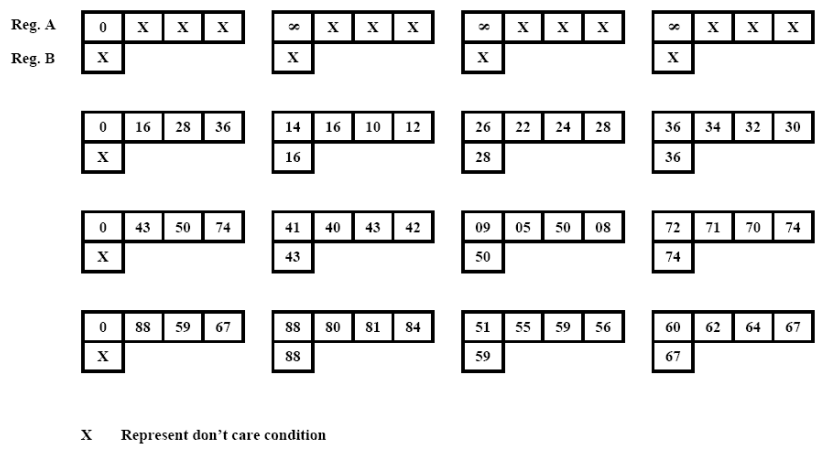

FIG. 3.3. *Data Distribution after step 4.*

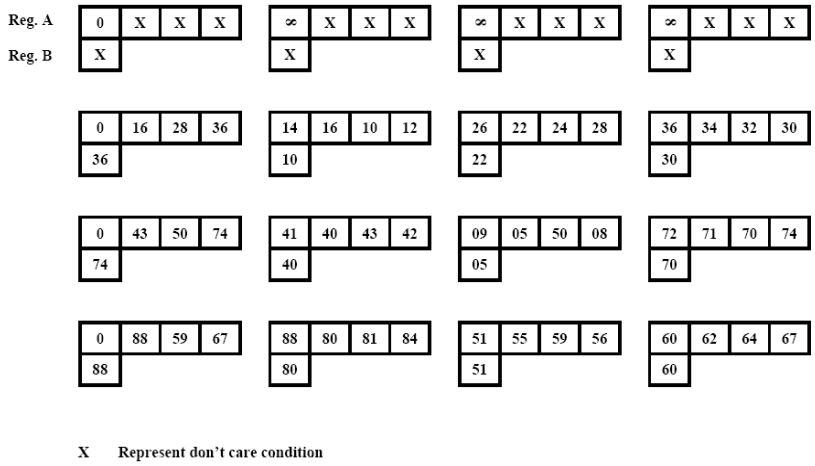

FIG. 3.4. *Data Distribution after step 7.*

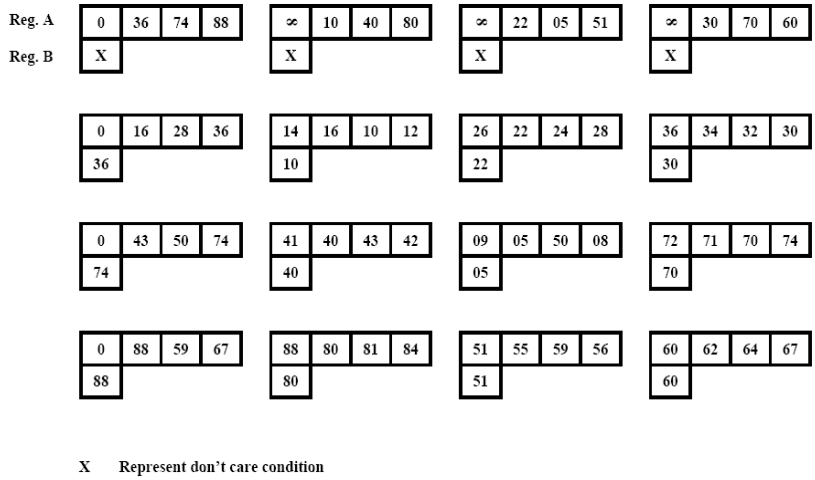

FIG. 3.5. *Data Distribution after step 8.*

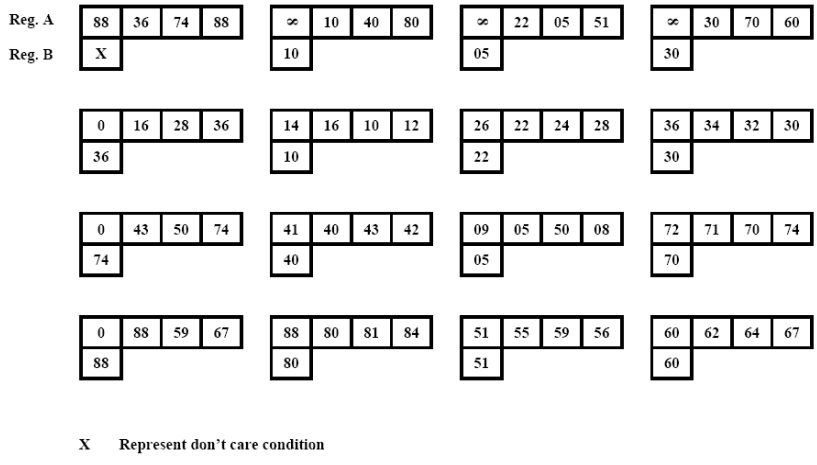

FIG. 3.6. *Data Distribution after step 11.*

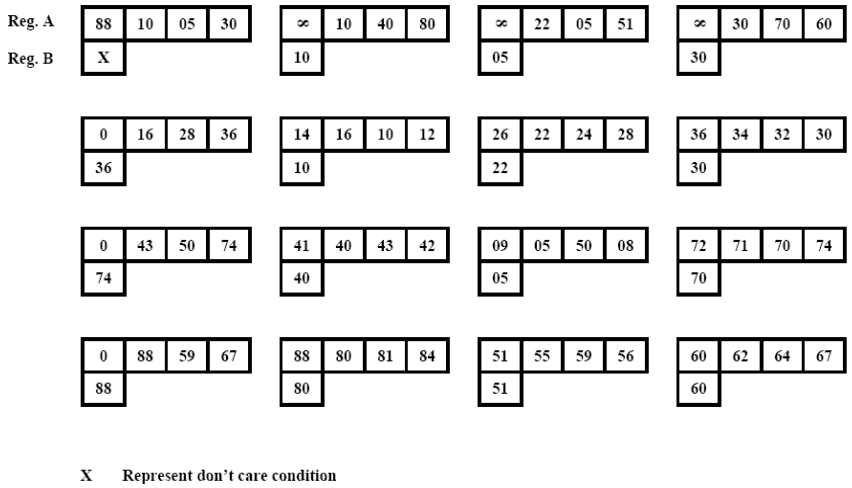

FIG. 3.7. *Data Distribution after step 12.*

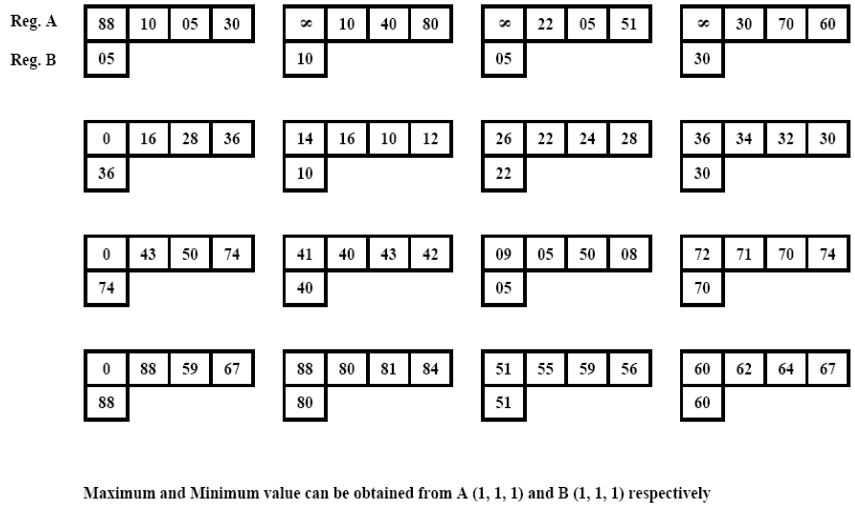

FIG. 3.8. *Maximum and Minimum value can be obtained from* $A(1,1,1)$ *and* $B(1,1,1)$.

2. Find the maximum from the data stored in layer 1 of all TSE networks, G_{ij}, $2 \leq i,j \leq n$, of OMTSE interconnection system.
3. For all i,j, $2 \leq i,j \leq n$;
   $B(i,j,1) \leftarrow A(i,j,n+1)$
4. For all i,j, $2 \leq i,j \leq n$;
   $A(i,1,j) \leftarrow B(i,j,1)$ /* horizontal inter-TSE link */
5. do in parallel
   (i) For all i,k; $2 \leq i \leq n$ and $n+1 \leq k \leq 3n/2$, do in parallel
   $A(i,1,k) \leftarrow \max\big(A(LCHILD(P(i,1,k))) + A(RCHILD(P(i,1,k)))\big)$
   (ii) For all i,j,k, $2 \leq i,j \leq n$ and $n+1 \leq k \leq 3n/2$, do in parallel
   $A(i,j,k) \leftarrow \min\big(A(LCHILD(P(i,j,k))) + A(RCHILD(P(i,j,k)))\big)$
6. do in parallel
   (i) Find the maximum from the data stored in layer 1 of all TSE networks, G_{i1}, $2 \leq i \leq n$, of OMTSE interconnection system.
   (ii) Find the minimum from the data stored in layer 1 of all TSE networks, G_{ij}, $2 \leq i,j \leq n$, of OMTSE interconnection system.

7. For all $i, j, 2 \leq i \leq n, 1 \leq j \leq n$, do in parallel
   $B(i, j, 1) \leftarrow A(i, j, n + 1)$
8. For all $i, j, 2 \leq i \leq n, 1 \leq j \leq n$, do in parallel
   $A(1, j, i) \leftarrow B(i, j, 1)$ /* vertical inter-TSE link */
9. do in parallel
   (i) For all k; $n + 1 \leq k \leq 3n/2$, do in parallel
   $A(1, 1, k) \leftarrow \max\big(A(LCHILD(P(1, 1, k))) + (RCHILD(P(1, 1, k)))\big)$
   (ii) For all j, k; $2 \leq j \leq n$ and $n + 1 \leq k \leq 3n/2$, do in parallel
   $A(1, j, k) \leftarrow \min\big(A(LCHILD(P(1, j, k))) + A(RCHILD(P(1, j, k)))\big)$
10. do in parallel
   (i) Find the maximum from the data stored in layer 1 of all TSE networks, G_{11} of OMTSE interconnection system.
   (ii) Find the minimum from the data stored in layer 1 of all TSE networks, G_{1j}, $2 \leq j \leq n$, of OMTSE interconnection system.
11. do in parallel
   (i) $A(1, 1, 1) \leftarrow A(1, 1, n + 1)$
   (ii) For all j, $2 \leq j \leq n$,
   $B(1, j, 1) \leftarrow A(1, j, n + 1)$
12. For all j, $2 \leq j \leq n$,
   $A(1, 1, j) \leftarrow B(1, j, 1)$ /* horizontal inter-TSE link */
13. For all k; $n + 1 \leq k \leq 3n/2$, do in parallel
   $A(1, 1, k) \leftarrow \min\big(A(LCHILD(P(1, 1, k))) + (RCHILD(P(1, 1, k)))\big)$
14. Find the minimum from the data stored in layer 1 of all TSE networks, G_{11} of OMTSE interconnection system.
15. $B(1, 1, 1) \leftarrow A(1, 1, n + 1)$.

The above algorithm has been illustrated through Fig. 3.1 to 3.8, for m distinct positive numbers, all of which are less than a given number (say, ∞). The output can be taken respectively form $A(1, 1, 1)$ and $B(1, 1, 1)$ for maximum and minimum of m elements. Steps 4, 8 and 12 require only one move through optical link each, step 2, 6, 10, and 14 require $(\log n - 1)$ data moves through electronic link where as rest of the instructions require one move through each electronic link. The overall time of the computation is $(4 \log n + 4)$ electronic links and 3 optical links.

The comparative analysis of the above algorithm with the corresponding SIMD algorithm [12] in respect to the size of data set is given in the following section.

4. Feasibility Analysis. The proposed MIMD algorithm for finding the maximum and minimum of m data elements on OMTSE requires an overall $4 \log n + 7$ number of moves where as the SIMD approach require $6 \log n + 10$ for n^3 data elements [12]. In both cases the algorithm uses a total of $3n^3/2$ processors, built around n^2 TSE networks of OMTSE interconnection system. Taking the number of data elements into consideration, the study of feasibility of the proposed algorithm has been done with a total number of $n^4(n - 1)^2$ data elements. For these number of data elements it can be seen that SIMD and MIMD algorithms can be repeated to a maximum of $n^3 - 2n^2 + n + 2$ and $n^3 + 4$ times respectively. Consequently the total number of data movement (both optical and electronic) can respectively be obtained as $\lambda(n) = (6 \log n + 10)(n(n - 1)^2 + 2)$ and $\mu(n) = (4 \log n + 7)(n^3 + 4)$. For various values of n, $\lambda(n)$ and $\mu(n)$ are depicted in Fig. 4.1.

5. Conclusion. This paper introduces a synchronous MIMD algorithm for finding both maximum and minimum of a set of data elements on OMTSE interconnection system with $O(\log n)$ time. It turns out from above discussion that the overall performance of the proposed algorithm is much more satisfactory than the existing SIMD algorithm discussed in [12].

REFERENCES

[1] A. KRISHNAMOORTHY, P. MARCHAND, F. KIAMILEV AND S. ESNER, *Grain-Size Considerations for Optoelectronic Multistage Interconnection Networks*, Applied Optics, Vol. 31, no. 26, 1992, pp. 5480–5507.

[2] A. OSTERLOH, *Sorting on the OTIS-Mesh*, Proc. 14th International Parallel and Distributed Processing Symposium (IPDPS 2000), 2000, pp. 269–274.

[3] B. P. SINHA, S. BANDYOPADHYAY, *OMULT: An Optical Interconnection System for Parallel Computing*, Proc. of International Conference on Computers and Devices for Communications, 2004.

[4] C. -F. WANG, S. SAHNI, *Basic operations on the OTIS-Mesh optoelectronic computer*, IEEE Transaction on Parallel and Distributed Systems, Vol. 9, no.12, 1998, pp. 1226–1236.

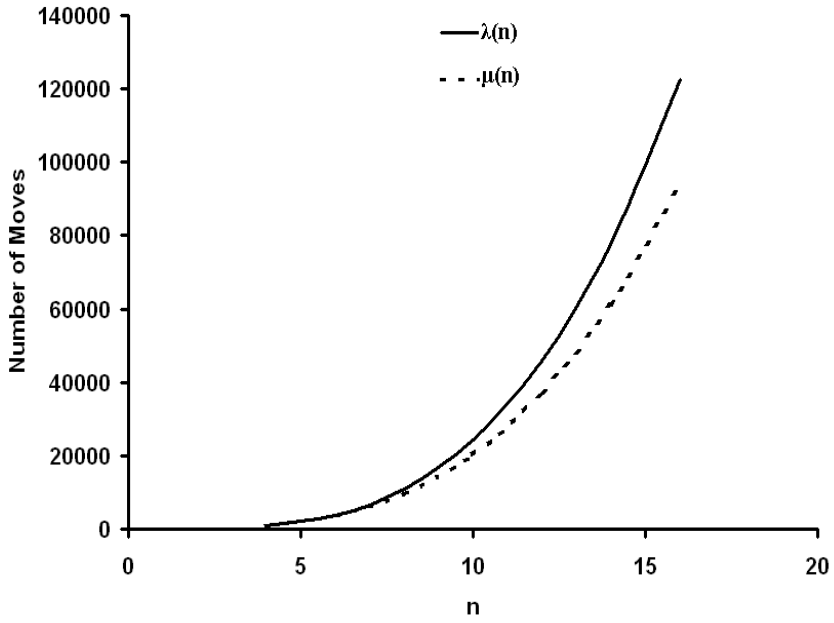

FIG. 4.1. *Comparison between $\lambda(n)$ and $\mu(n)$.*

[5] C. -F. WANG, S. SAHNI, *BPC permutations on the OTIS-Mesh optoelectronic computer*, Proc. of the International Conference on Massively Parallel processing using Optical Interconnections (MPPOI), 1997, pp. 130–135.
[6] C. -F. WANG, S. SAHNI, *Image Processing on the OTIS-Mesh optoelectronic computer*, IEEE Transaction on Parallel and Distributed Systems, Vol. 11, no.2, 2000, pp. 97–109.
[7] C. -F. WANG, S. SAHNI, *Matrix Multiplication on the OTIS-Mesh optoelectronic computer*, IEEE Transaction on Computers, Vol. 50, No.7, 2001, pp. 635–646.
[8] F. ZANE, P. MARCHAND, R. PATURI, S. ESENER, *Scalable Network Architectures Using Optical Transpose Interconnection System (OTIS)*, Journal of Parallel and Distributed Computing, Vol. 60, no. 5, 2000, pp. 521–538.
[9] G. C. MARSDEN, P. J. MARCAAND, P. HARVEY, S. C. ESNER, *Optical Transpose Interconnection System Architectures*, Optics Letters, Vol. 18, no.13, 1993, pp. 1083–1085.
[10] M. FELDMAN, S. ESENER, C. GUEST, S. LEE, *Comparison between Electrical and Free-Space Optical Interconnects Based on Power and Speed Consideration*, Applied Optics, Vol. 27, no.9, 1988, pp. 1742–1751.
[11] S. C. PANIGRAHI, S. PAUL, G. SAHOO, *OMTSE - An Optical Interconnection System for Parallel Computing*, Proc. 14th International Conference on Advanced Computing and Communication (ADCOM), 2006, pp. 626–627.
[12] S. C. PANIGRAHI, S. PAUL, G. SAHOO, *Parallel Prefix Computation, Sorting and Reduction Operation on OMTSE architecture*,Proc. of the International Conference on Advanced Computing and Communications (ICACC), 2007, pp. 616–620.
[13] S. RAJASEKARAN, S. SAHNI, *Randomized routing, selection and sorting on the OTIS-Mesh optoelectronic computer*,IEEE transaction on Parallel and Distributed Systems, Vol. 9, No.9, 1998, pp. 833–840.
[14] W. HENDRICK, O. KIBAR, P. MARCHAND, C. FAN, D. V. BLERKOM, F. MCCORMIC, I. COKGOR, M. HANSEN, S. ESENER, *Modeling and Optimization of the Optical Transpose Interconnection System*,Optoelectronic Technology Center, Program Review, Cornell Univ, 1995.

Edited by: Marcin Paprzycki
Received: Nov 16, 2007
Accepted: Nov 21, 2007