

Parallel prefix computation on extended multi-mesh network

Prasanta K. Jana^{a,*}, B. Damodara Naidu^a, Shailendra Kumar^b,
Monish Arora^b, Bhabani P. Sinha^c

^a Department of Computer Science and Engineering, Indian School of Mines, Dhanbad 826 004, India

^b CREWMaN Lab, 411 Woolf Hall, 500 First Street, Arlington, TX 76019, USA

^c Advanced Computing and Microelectronics Unit, Indian Statistical Institute, 203 B.T. Road, Calcutta 700 035, India

Received 25 April 2000; received in revised form 29 April 2002

Communicated by A. El Abbadi

Abstract

A parallel algorithm for prefix computation of $N = n^4$ elements on an $n \times n$ extended multi-mesh network is presented. The network is a modified version of an earlier multi-mesh network with a 4-regular structure. The algorithm takes $O(N^{1/4})$ time on N processors ($13N^{1/4} - 5$ communication steps and $\log N + 4$ arithmetic/logic steps).

Keywords: Prefix computation; Multi-mesh network; Parallel algorithm; Time complexity; Associative binary operation

1. Introduction

Given N data values x_1, \dots, x_N and the associative binary operation \circ , the prefix problem is to compute $P_i = x_1 \circ x_2 \circ x_3 \circ \dots \circ x_i$, $1 \leq i \leq N$.

Parallel solutions to many real-life problems such as job scheduling and knapsack depend on the efficiency at which prefix computations can be carried out. Prefix computation is also extensively used in loop optimization, evaluation of polynomials, solution of linear equations, polynomial interpolation, etc. Several parallel algorithms for prefix computation involved have been reported in [6,7,10–12,16]. Owing to the broad applications of this computation, many prefix circuits have also appeared in the literature [8,13–15]. Paper [1] describes an $O(\log N)$ time algorithm on a specialized network with $O(N \log N)$ processors. Lin and Lin [9] present an algorithm on a fully connected message passing system of N processors with no more than $\lceil 1.44 \log N \rceil + 1$ communication steps. On a CRCW-PRAM model, an algorithm requiring $O(\log N / \log \log N)$ time using $(N \log \log N) / \log N$ processors appears in [6]. Egecioglu and Srinivasan [2] give a $2\tau\sqrt{N} + O(\log N)$ time algorithm on a $\sqrt{N} \times \sqrt{N}$ mesh, where τ is the time for a single routing step. They also give an algorithm with $\sqrt{2}\tau\sqrt{N} + O(\sqrt{\tau}N^{1/4})$ time on a disc of N processors.

The multi-mesh network of [4] has several interesting topological properties, e.g., low diameter, existence of a Hamiltonian cycle, 4-regularity. Parallel algorithms for several fundamental problems including summation, matrix-multiplication, sorting, polynomial interpolation, and DFT computation, have been efficiently mapped onto this architecture and reported in [3,5]. A parallel algorithm for finding polynomial roots on this architecture with a little topological modification appears in [17]. An extension of this network is proposed in [18] for its application in designing light wave networks.

Our parallel algorithm for prefix computation on an $N^{1/4} \times N^{1/4}$ extended multi-mesh network uses a modified version of the multi-mesh network of [4]. The algorithm, which is based on the optimal prefix computation mesh given in [2], runs in time $O(N^{1/4})$ for N data values using N processors. It takes $13N^{1/4} - 5$ communication steps and $4 \log N^{1/4} + 4$ arithmetic/logic steps. This can be compared with $2\sqrt{N} + 1$ communication steps and $\log N + 1$ arithmetic/logic steps required by the algorithm of [2] on a square mesh using N processors.

2. The computational model

An $n \times n$ multi-mesh network [4] consists of n^2 meshes, each of size $n \times n$. These n^2 meshes are arranged to form an $n \times n$ lattice. Thus, it has $N = n^4$ processors. Each $n \times n$ mesh in this network is called a *block*. Fig. 1 shows a 4×4 multi-mesh network.

Let $BL(\alpha, \beta)$ denote the block in row α and column β . By $P(\alpha, \beta, x, y)$, we denote the processor in row x th row and column y th column of block $BL(\alpha, \beta)$. Within the block, processor $P(\alpha, \beta, x, y)$ is directly connected by to processors $P(\alpha, \beta, x \pm 1, y \pm 1)$, if they exist, by means of bi-directional, links, called *intra-block* links. Additional *inter-block* bi-directional links, which connect the boundary or corner processors between different blocks, are defined as follows:

- (1) $P(\alpha, \beta, 1, y)$ is connected to $P(y, \beta, n, \alpha)$ for $1 \leq y, \alpha, \beta \leq n$. As a special case, for $\alpha = y$ these links connect processors within the same block $BL(\alpha, \beta)$.
- (2) $P(\alpha, \beta, x, 1)$ is connected to $P(\alpha, x, \beta, n)$ for $1 \leq x, \alpha, \beta \leq n$. As a special case, for $\beta = x$ these links connect processors within the same block $BL(\alpha, \beta)$.

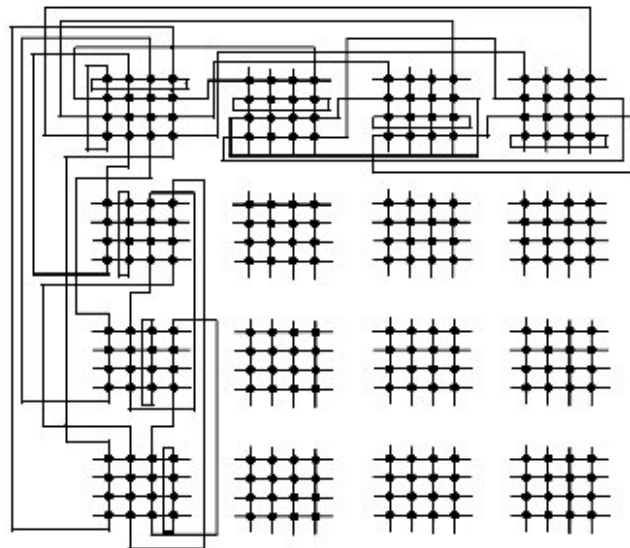


Fig. 1. 4×4 multi-mesh (all links are not shown).

For the sake of efficient mapping of prefix computation, we extend the multi-mesh with extra links as follows. We call such a multi-mesh an $n \times n$ extended multi-mesh.

- (i) $P(\alpha, \beta, n/2 + 1, n)$ is connected to $P(\alpha + 1, \beta, n/2 + 1, n)$ for $1 \leq \alpha < n, 1 \leq \beta \leq n$.
- (ii) $P(n/2, \beta, n/2 + 1, n)$ is connected to $P(n/2, \beta + 1, n/2 + 1, n)$ for $1 \leq \alpha \leq n, 1 \leq \beta < n$.
- (iii) $P(n/2 + 1, \beta, n/2 + 1, n)$ is connected to $P(n/2 + 1, \beta + 1, n/2 + 1, n)$ for $1 \leq \beta < n$.

3. The parallel prefix algorithm

The key idea in our parallel algorithm for prefix computation of $N = n^4$ data values on an $n \times n$ extended multi-mesh network is as follows. The computation is performed in two phases. In the first phase, the optimal prefix computation (Algorithm B in [2]) is applied on all the blocks in parallel, and the partial results are stored locally, in each block. In phase two, the prefix computation is computed on the whole architecture using the partial result of phase one. For this, each block is treated as a single virtual processor. To implement phase two, the links among different blocks are used to broadcast the partial results of phase one. For the sake of simplicity, we assume that n is a power of 2. However, the basic idea can easily be extended when this is not so.

The algorithm is described below stepwise. The important steps are illustrated by an example for $n = 4$ with the help of Figs. 2–6. In these figures, solid lines show the data routing using actual and virtual links. We consider, however, the actual (physical) links to calculate the time complexity required by each step of the algorithm. We assume that a communication step requires τ units of time and that a binary operation \circ requires unit time.

Parallel Algorithm:

Initialization: Initially, the data elements are distributed among the processors $P(\alpha, \beta, i, j)$ such that $P(\alpha, \beta, i, j)$ contains the following data (see Fig. 2):

$$\begin{aligned} x_{(j-1)n+i+(\alpha-1)n^2+(\beta-1)n^3} & \quad \text{for } \alpha \leq n/2 \text{ and } i \leq n/2, \\ x_{jn+n/2+1-i+(\alpha-1)n^2+(\beta-1)n^3} & \quad \text{for } \alpha \leq n/2 \text{ and } i > n/2, \\ x_{(j-1)n+i+(3n/2-\alpha)n^2+(\beta-1)n^3} & \quad \text{for } \alpha > n/2 \text{ and } i \leq n/2, \\ x_{jn+n/2+1-i+(3n/2-\alpha)n^2+(\beta-1)n^3} & \quad \text{for } \alpha > n/2 \text{ and } i > n/2. \end{aligned}$$

Step 1 (the first phase): Apply prefix computation on each block in parallel.

Apply the optimal prefix algorithm (given as Algorithm B) in [2] on each block in parallel. The contents of $P(\alpha, \beta, i, j)$ is shown below. Fig. 3 illustrates this step, where we use $i : j$ to denote $x_i \circ x_{i+1} \circ \dots \circ x_j$. This step requires parallel time $2n\tau + 2 \log n + \tau + 1$; it is step 1 of Algorithm B of [2].

$$\begin{aligned} & x[1 + (\alpha - 1)n^2 + (\beta - 1)n^3 : (j - 1)n + i + (\alpha - 1)n^2 + (\beta - 1)n^3] \\ & \quad \text{for } \alpha \leq n/2 \text{ and } i \leq n/2, \\ & x[1 + (\alpha - 1)n^2 + (\beta - 1)n^3 : jn + n/2 + 1 - i + (\alpha - 1)n^2 + (\beta - 1)n^3] \\ & \quad \text{for } \alpha \leq n/2 \text{ and } i > n/2, \\ & x[1 + (3n/2 - \alpha)n^2 + (\beta - 1)n^3 : (j - 1)n + i + (3n/2 - \alpha)n^2 + (\beta - 1)n^3] \\ & \quad \text{for } \alpha > n/2 \text{ and } i \leq n/2, \\ & x[1 + (3n/2 - \alpha)n^2 + (\beta - 1)n^3 : jn + n/2 + 1 - i + (3n/2 - \alpha)n^2 + (\beta - 1)n^3] \\ & \quad \text{for } \alpha > n/2 \text{ and } i > n/2. \end{aligned}$$

X ₁	X ₅	X ₉	X ₃	X ₆₅	X ₆₉	X ₇₃	X ₇₇	X ₁₂₉	X ₁₃₃	X ₁₃₇	X ₁₄₁	X ₁₉₃	X ₁₉₇	X ₂₀₁	X ₂₀₅
X ₂	X ₆	X ₁₀	X ₄	X ₆₆	X ₇₀	X ₇₄	X ₇₈	X ₁₃₀	X ₁₃₄	X ₁₃₈	X ₁₄₂	X ₁₉₄	X ₁₉₈	X ₂₀₂	X ₂₀₆
X ₄	X ₈	X ₁₂	X ₆	X ₆₈	X ₇₂	X ₇₆	X ₈₀	X ₁₃₂	X ₁₃₆	X ₁₄₀	X ₁₄₄	X ₁₉₆	X ₂₀₀	X ₂₀₄	X ₂₀₈
X ₃	X ₇	X ₁₁	X ₁₅	X ₆₇	X ₇₁	X ₇₅	X ₇₉	X ₁₃₁	X ₁₃₅	X ₁₃₉	X ₁₄₃	X ₁₉₅	X ₁₉₉	X ₂₀₃	X ₂₀₇
X ₁₇	X ₂₁	X ₂₅	X ₂₉	X ₈₁	X ₈₅	X ₈₉	X ₉₃	X ₁₄₅	X ₁₄₉	X ₁₅₃	X ₁₅₇	X ₂₀₉	X ₂₁₃	X ₂₁₇	X ₂₂₁
X ₁₈	X ₂₂	X ₂₆	X ₃₀	X ₈₂	X ₈₆	X ₉₀	X ₉₄	X ₁₄₆	X ₁₅₀	X ₁₅₄	X ₁₅₈	X ₂₁₀	X ₂₁₄	X ₂₁₈	X ₂₂₂
X ₂₀	X ₂₄	X ₂₈	X ₃₂	X ₈₄	X ₈₈	X ₉₂	X ₉₆	X ₁₄₈	X ₁₅₂	X ₁₅₆	X ₁₆₀	X ₂₁₂	X ₂₁₆	X ₂₂₀	X ₂₂₄
X ₁₉	X ₂₃	X ₂₇	X ₃₁	X ₈₃	X ₈₇	X ₉₁	X ₉₅	X ₁₄₇	X ₁₅₁	X ₁₅₅	X ₁₅₉	X ₂₁₁	X ₂₁₅	X ₂₁₉	X ₂₂₃
X ₄₉	X ₅₃	X ₅₇	X ₆₁	X ₁₁₃	X ₁₁₇	X ₁₂₁	X ₁₂₅	X ₁₇₇	X ₁₈₁	X ₁₈₅	X ₁₈₉	X ₂₄₁	X ₂₄₅	X ₂₄₉	X ₂₅₃
X ₅₀	X ₅₄	X ₅₈	X ₆₂	X ₁₁₄	X ₁₁₈	X ₁₂₂	X ₁₂₆	X ₁₇₈	X ₁₈₂	X ₁₈₆	X ₁₉₀	X ₂₄₂	X ₂₄₆	X ₂₅₀	X ₂₅₄
X ₅₂	X ₅₆	X ₆₀	X ₆₄	X ₁₁₆	X ₁₂₀	X ₁₂₄	X ₁₂₈	X ₁₈₀	X ₁₈₄	X ₁₈₈	X ₁₉₂	X ₂₄₄	X ₂₄₈	X ₂₅₂	X ₂₅₆
X ₅₁	X ₅₅	X ₅₉	X ₆₃	X ₁₁₅	X ₁₁₉	X ₁₂₃	X ₁₂₇	X ₁₇₉	X ₁₈₃	X ₁₈₇	X ₁₉₁	X ₂₄₃	X ₂₄₇	X ₂₅₁	X ₂₅₅
X ₃₃	X ₃₇	X ₄₁	X ₄₅	X ₉₇	X ₁₀₁	X ₁₀₅	X ₁₀₉	X ₁₆₁	X ₁₆₅	X ₁₆₉	X ₁₇₃	X ₂₂₅	X ₂₂₉	X ₂₃₃	X ₂₃₇
X ₃₄	X ₃₈	X ₄₂	X ₄₆	X ₉₈	X ₁₀₂	X ₁₀₆	X ₁₁₀	X ₁₆₂	X ₁₆₆	X ₁₇₀	X ₁₇₄	X ₂₂₆	X ₂₃₀	X ₂₃₄	X ₂₃₈
X ₃₆	X ₄₀	X ₄₄	X ₄₈	X ₁₀₀	X ₁₀₄	X ₁₀₈	X ₁₁₂	X ₁₆₄	X ₁₆₈	X ₁₇₂	X ₁₇₆	X ₂₂₈	X ₂₃₂	X ₂₃₆	X ₂₄₀
X ₃₅	X ₃₉	X ₄₃	X ₄₇	X ₉₉	X ₁₀₃	X ₁₀₇	X ₁₁₁	X ₁₆₃	X ₁₆₇	X ₁₇₁	X ₁₇₅	X ₂₂₇	X ₂₃₁	X ₂₃₅	X ₂₃₉

Fig. 2. Initialization.

1:1	1:5	1:9	1:13	65:65	65:69	65:73	65:77	129:129	129:133	129:137	129:141	193:193	193:197	193:201	193:205
1:2	1:6	1:10	1:14	65:66	65:70	65:74	65:78	129:130	129:134	129:138	129:142	193:194	193:198	193:202	193:206
1:4	1:8	1:12	1:16	65:68	65:72	65:76	65:80	129:132	129:136	129:140	129:144	193:196	193:200	193:204	193:208
1:3	1:7	1:11	1:15	65:67	65:71	65:75	65:79	129:131	129:135	129:139	129:143	193:195	193:199	193:203	193:207
17:17	17:21	17:25	17:29	81:81	81:85	81:89	81:93	145:145	145:149	145:153	145:157	209:209	209:213	209:217	209:221
17:18	17:22	17:26	17:30	81:82	81:86	81:90	81:94	145:146	145:150	145:154	145:158	209:210	209:214	209:218	209:222
17:20	17:24	17:28	17:32	81:84	81:88	81:92	81:96	145:148	145:152	145:156	145:160	209:212	209:216	209:220	209:224
17:19	17:23	17:27	17:31	81:83	81:87	81:91	81:95	145:147	145:151	145:155	145:159	209:211	209:215	209:219	209:223
.....															
49:49	49:53	49:57	49:61	113:113	113:117	113:121	113:125	177:177	177:181	177:185	177:189	241:241	241:245	241:249	241:253
49:50	49:54	49:58	49:62	113:114	113:118	113:122	113:126	177:178	177:182	177:186	177:190	241:242	241:246	241:250	241:254
49:52	49:56	49:60	49:64	113:116	113:120	113:124	113:128	177:180	177:184	177:188	177:192	241:244	241:248	241:252	241:256
49:51	49:55	49:59	49:63	113:115	113:119	113:123	113:127	177:179	177:183	177:187	177:191	241:243	241:247	241:251	241:255
33:33	33:37	33:41	33:45	97:97	97:101	97:105	97:109	161:161	161:165	161:169	161:173	225:225	225:229	225:233	225:237
33:34	33:38	33:42	33:46	97:98	97:102	97:106	97:110	161:162	161:166	161:170	161:174	225:226	225:230	225:234	225:238
33:36	33:40	33:44	33:48	97:100	97:104	97:108	97:112	161:164	161:168	161:172	161:176	225:228	225:232	225:236	225:240
33:35	33:39	33:43	33:47	97:99	97:103	97:107	97:111	161:163	161:167	161:171	161:175	225:227	225:231	225:235	225:239

Fig. 3. Situation after step 1.

Step 2 (start of the second phase): Apply the same prefix computation on the whole extended multi-mesh network.

Do steps 2.1 and 2.2 in parallel:

- 2.1** Using the communication links between processors $P(\alpha, \beta, n/2 + 1, n)$ and $P(\alpha + 1, \beta, n/2 + 1, n)$, for $1 \leq \alpha < n/2, 1 \leq \beta \leq n$, follow the steps similar to those in the algorithm of [16] to compute the prefix over all $n/2$ data values (the values computed after step 1) contained in $P(\alpha, \beta, n/2 + 1, n)$.
- 2.2** Using the communication links between processors $P(\alpha, \beta, n/2 + 1, n)$ and $P(\alpha + 1, \beta, n/2 + 1, n)$, for $n/2 + 1 < \alpha \leq n, 1 \leq \beta \leq n$, follow the steps similar to those in the algorithm of [16] to compute the prefix over all $n/2$ data values (prefix values computed after step 1) stored in $P(\alpha, \beta, n/2 + 1, n)$.

Steps 2.1 and 2.2 use the log $n/2$ arithmetic step algorithm of [16] to compute the prefix on $n/2$ values. The data communication paths are shown in Fig. 3. The number of routing steps required for each step is $n/2 - 1$. Thus, steps 2.1 and 2.2 require parallel time $(n/2 - 1)\tau + \log n/2$.

2.3 For all $\alpha, \beta, 1 < \alpha < n, 1 \leq \beta \leq n$, broadcast the currently computed prefix value at $P(\alpha, \beta, n/2 + 1, n)$ to other processors in the same block.

Step 2.3 requires $2(n - 1)\tau$ parallel time units.

2.4 For all $\beta, j, 1 \leq \beta \leq n, 1 \leq j \leq n, P(\alpha, \beta, i, j)$ computes the following:

$$\begin{aligned}
 &x[(\beta - 1)n^3 + 1 : (j - 1)n + i + (\alpha - 1)n^2 + (\beta - 1)n^3] \\
 &\quad \text{for } \alpha \leq n/2 \text{ and } i \leq n/2, \\
 &x[(\beta - 1)n^3 + 1 : jn + n/2 + 1 - i + (\alpha - 1)n^2 + (\beta - 1)n^3] \\
 &\quad \text{for } \alpha \leq n/2 \text{ and } i > n/2, \\
 &x[n^3/2 + (\beta - 1)n^3 + 1 : (j - 1)n + i + (3n/2 - \alpha)n^2 + (\beta - 1)n^3] \\
 &\quad \text{for } \alpha > n/2 \text{ and } i \leq n/2, \\
 &x[n^3/2 + (\beta - 1)n^3 + 1 : jn + n/2 + 1 - i + (3n/2 - \alpha)n^2 + (\beta - 1)n^3] \\
 &\quad \text{for } \alpha > n/2 \text{ and } i > n/2.
 \end{aligned}$$

The situation after step 2.4 is illustrated in Fig. 4. Step 2.4 requires one unit of time.

Step 3.

3.1 For all $\beta, 1 \leq \beta \leq n$, send the values computed in step 2 from processor $P(n/2, \beta, n/2 + 1, n)$ to processor $P(n/2 + 1, \beta, 1, n/2)$ and from processor $P(n/2 + 1, \beta, n/2 + 1, n)$ to $P(n/2, \beta + 1, n, n/2 + 1)$.

The data routing for this step is shown in Fig. 4. This step requires $n\tau$ units of time.

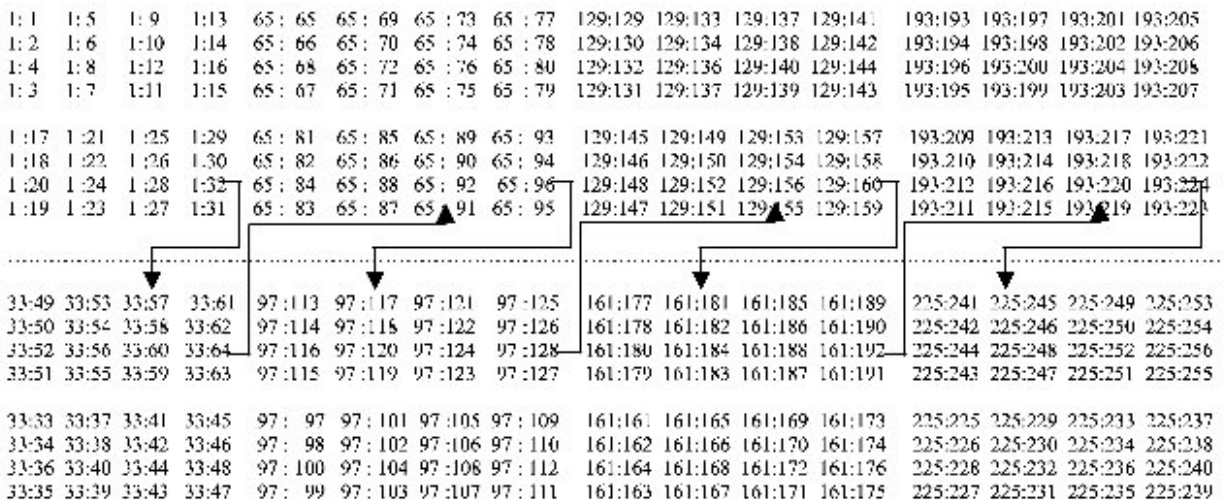


Fig. 4. After step 2.

3.2 Broadcast the data elements received in each block ($n \times n$ mesh) in step 3.1 to all other processors in that block.

Step 3.2 needs $2(n - 1)\tau$ units of time.

3.3 For all j , $1 \leq j \leq n$, $P(n/2, \beta, i, j)$ computes:

$$x[n^3/2 + (\beta - 2)n^3 + 1 : (j - 1)n + i + (\alpha - 1)n^2 + (\beta - 1)n^3]$$

for $2 \leq \beta \leq n$ and $i \leq n/2$,

$$x[n^3/2 + (\beta - 2)n^3 + 1 : jn + n/2 + 1 - i + (\alpha - 1)n^2 + (\beta - 1)n^3]$$

for $2 \leq \beta \leq n$ and $i > n/2$,

and $P(n/2 + 1, \beta, i, j)$ computes:

$$x[(\beta - 1)n^3 + 1 : (j - 1)n + i + (3n/2 - \alpha)n^2 + (\beta - 1)n^3]$$

for $1 \leq \beta \leq n$ and $i \leq n/2$,

$$x[(\beta - 1)n^3 + 1 : jn + n/2 + 1 - i + (3n/2 - \alpha)n^2 + (\beta - 1)n^3]$$

for $1 \leq \beta \leq n$ and $i > n/2$.

Step 3.3 needs one unit of time. The content of each processor after step 3 is shown in Fig. 5.

Step 4. The processors in the blocks of row $n/2 + 1$, perform their prefix computations along this row. Similarly the processors in the blocks of row $n/2$ perform their prefix computations along this row as shown in Fig. 5.

Do steps 4.1 and 4.2 in parallel.

4.1 For all β , $1 \leq \beta < n$, using the communication links between processors $P(n/2, \beta, n/2 + 1, n)$ and $P(n/2, \beta + 1, n/2 + 1, n)$, follow steps similar to those in the algorithm of [16] to compute the prefix over all n values (prefix values computed by step 3) contained in $P(n/2, \beta, n/2 + 1, n)$ and $P(n/2, n, n/2 + 1, n)$.

1:1	1:5	1:9	1:13	65:65	65:69	65:73	65:77	129:129	129:133	129:137	129:141	193:193	193:197	193:201	193:205
1:2	1:6	1:10	1:14	65:66	65:70	65:74	65:78	129:130	129:134	129:138	129:142	193:194	193:198	193:202	193:206
1:4	1:8	1:12	1:16	65:68	65:72	65:76	65:80	129:132	129:136	129:140	129:144	193:196	193:200	193:204	193:208
1:3	1:7	1:11	1:15	65:67	65:71	65:75	65:79	129:131	129:135	129:139	129:143	193:195	193:199	193:203	193:207
1:17	1:21	1:25	1:29	33:81	33:85	33:89	33:93	97:145	97:149	97:153	97:157	161:209	161:213	161:217	161:221
1:18	1:22	1:26	1:30	33:82	33:86	33:90	33:94	97:146	97:150	97:154	97:158	161:210	161:214	161:218	161:222
1:20	1:24	1:28	1:32	33:84	33:88	33:92	33:96	97:148	97:152	97:156	97:160	161:212	161:216	161:220	161:224
1:19	1:23	1:27	1:31	33:83	33:87	33:91	33:95	97:147	97:151	97:155	97:159	161:211	161:215	161:219	161:223
.....															
1:49	1:53	1:57	1:61	65:113	65:117	65:121	65:125	129:177	129:181	129:185	129:189	193:241	193:245	193:249	193:253
1:50	1:54	1:58	1:62	65:114	65:118	65:122	65:126	129:178	129:182	129:186	129:190	193:242	193:246	193:250	193:254
1:52	1:56	1:60	1:64	65:116	65:120	65:124	65:128	129:180	129:184	129:188	129:192	193:244	193:248	193:252	193:256
1:51	1:55	1:59	1:63	65:115	65:119	65:123	65:127	129:179	129:183	129:187	129:191	193:243	193:247	193:251	193:255
33:33	33:37	33:41	33:45	97:97	97:101	97:105	97:109	161:161	161:165	161:169	161:173	225:225	225:229	225:233	225:237
33:34	33:38	33:42	33:46	97:98	97:102	97:106	97:110	161:162	161:166	161:170	161:174	225:226	225:230	225:234	225:238
33:36	33:40	33:44	33:48	97:100	97:104	97:108	97:112	161:164	161:168	161:172	161:176	225:228	225:232	225:236	225:240
33:35	33:39	33:43	33:47	97:99	97:103	97:107	97:111	161:163	161:167	161:171	161:175	225:227	225:231	225:235	225:239

Fig. 5. After step 3.

4.2 For all β , $1 \leq \beta < n$, using the communication links between processors $P(n/2 + 1, \beta, n/2 + 1, n)$ and $P(n/2 + 1, \beta + 1, n/2 + 1, n)$, follow steps similar to those in the algorithm of [16] to compute the prefix over all the n values (prefix values computed by step 3) contained in $P(n/2 + 1, \beta, n/2 + 1, n)$ and $P(n/2 + 1, n, n/2 + 1, n)$.

Steps 4.1 and 4.2 calculate the prefix operation as indicated by circles in Fig. 5 in $(n - 1)$ communication steps and $\log n$ arithmetic steps [16], so these steps require parallel time $(n - 1)\tau + \log n$.

4.3 For all β , $1 < \beta \leq n$, broadcast the computed prefix values in $P(n/2, \beta - 1, n/2 + 1, n)$ and $P(n/2 + 1, \beta - 1, n/2 + 1, n)$ to all processors in the blocks $BL(n/2, \beta)$ and $BL(n/2 + 1, \beta)$, respectively.

Step 4.3 needs $(1 + (n - 1) + n/2)\tau = 1.5n\tau$ units of time.

4.4 For all j , $1 \leq j \leq n$, $P(n/2, \beta, i, j)$ computes

$$x[1 : (j - 1)n + i + (n/2 - 1)n^2 + (\beta - 1)n^3]$$

for $2 \leq \beta \leq n$ and for $i \leq n/2$,

$$x[1 : jn + n/2 + 1 - i + (n/2 - 1)n^2 + (\beta - 1)n^3]$$

for $2 \leq \beta \leq n$ and for $i > n/2$,

and processor $P(n/2 + 1, \beta, i, j)$ computes

$$x[1 : (j - 1)n + i + (n - 1)n^2 + (\beta - 1)n^3]$$

for $2 \leq \beta \leq n$ and for $i \leq n/2$,

$$x[1 : jn + n/2 + 1 - i + (n - 1)n^2 + (\beta - 1)n^3]$$

for $2 \leq \beta \leq n$ and for $i > n/2$.

Step 4.4 requires one unit of time. Fig. 6 illustrates the resulting situation after Step 4.

Step 5. Do steps 5.1 and 5.2 in parallel.

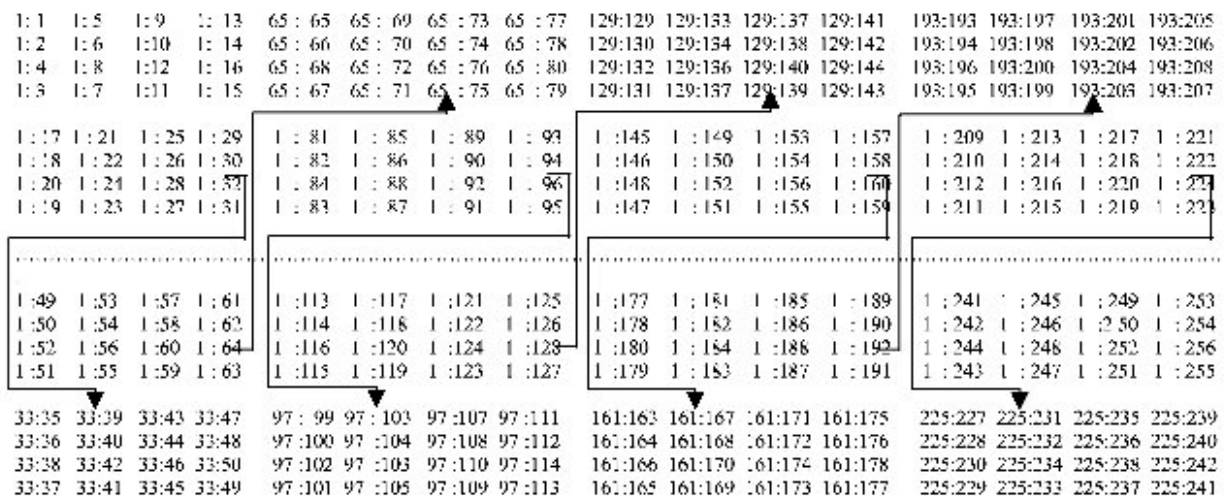


Fig. 6. Situation after step 4.

- 5.1 For all β , $1 \leq \beta < n$, send the value computed in step 4 from processor $P(n/2 + 1, \beta, n/2 + 1, n)$ to $P(n/2 - 1, \beta + 1, n, n/2 + 1)$.
- 5.2 For all β , $1 \leq \beta \leq n$, send the values computed in step 4 from the processor $P(n/2, \beta, n/2 + 1, n)$ to $P(n/2 + 2, \beta, 1, n/2)$.

The data routing for steps 5.1 and 5.2 is shown in Fig. 6. Step 5 requires parallel time $(n/2 + 1)\tau$.

Step 6. For all β , $1 < \beta \leq n$, processors $P(n/2 - 1, \beta + 1, n, n/2 + 1)$ and $P(n/2 + 2, \beta, 1, n/2)$ broadcast their computed prefix values to any one processor in each of the blocks in their corresponding columns (i.e., blocks with the same β). (This can be done by first broadcasting the data along a row (row n of block $BL(n/2 - 1, \beta + 1)$ or the first row of block $BL(n/2 + 2, \beta)$) and then sending the data to one processor of each of the other blocks with the same β by following one more link of the multi-mesh network.) Then, broadcast these values to all other processors in the corresponding blocks.

In step 6, sending data to any one processor of each of the other blocks requires $(n/2 + 1)\tau$ time units. Following that, broadcasting data within a block needs time $2(n - 1)\tau$.

Step 7. Each processor calculates its target prefix value in unit time, so that $P(\alpha, \beta, i, j)$ contains

$$\begin{aligned}
 &x[1 : (j - 1)n + i + (\alpha - 1)n^2 + (\beta - 1)n^3] \\
 &\quad \text{for } \alpha \leq n/2 \text{ and } i \leq n/2, \\
 &x[1 : jn + n/2 + 1 - i + (\alpha - 1)n^2 + (\beta - 1)n^3] \\
 &\quad \text{for } \alpha \leq n/2 \text{ and } i > n/2, \\
 &x[1 : (j - 1)n + i + (3n/2 - \alpha)n^2 + (\beta - 1)n^3] \\
 &\quad \text{for } \alpha > n/2 \text{ and } i \leq n/2, \\
 &x[1 : jn + n/2 + 1 - i + (3n/2 - \alpha)n^2 + (\beta - 1)n^3] \\
 &\quad \text{for } \alpha > n/2 \text{ and } i > n/2.
 \end{aligned}$$

The final result is shown in Fig. 7. The total computation time required by all the above steps is $(13n - 5)\tau + 4 \log n + 4$.

1: 1	1: 5	1: 9	1: 13	1: 65	1: 69	1: 73	1: 77	1: 129	1: 133	1: 137	1: 141	1: 193	1: 197	1: 201	1: 205
1: 2	1: 6	1: 10	1: 14	1: 66	1: 70	1: 74	1: 78	1: 130	1: 134	1: 138	1: 142	1: 194	1: 198	1: 202	1: 206
1: 4	1: 8	1: 12	1: 16	1: 68	1: 72	1: 76	1: 80	1: 132	1: 136	1: 140	1: 144	1: 196	1: 200	1: 204	1: 208
1: 3	1: 7	1: 11	1: 15	1: 67	1: 71	1: 75	1: 79	1: 131	1: 135	1: 139	1: 143	1: 195	1: 199	1: 203	1: 207
1: 17	1: 21	1: 25	1: 29	1: 81	1: 85	1: 89	1: 93	1: 145	1: 149	1: 153	1: 157	1: 209	1: 213	1: 217	1: 221
1: 18	1: 22	1: 26	1: 30	1: 82	1: 86	1: 90	1: 94	1: 146	1: 150	1: 154	1: 158	1: 210	1: 214	1: 218	1: 222
1: 20	1: 24	1: 28	1: 32	1: 84	1: 88	1: 92	1: 96	1: 148	1: 152	1: 156	1: 160	1: 212	1: 216	1: 220	1: 224
1: 19	1: 23	1: 27	1: 31	1: 83	1: 87	1: 91	1: 95	1: 147	1: 151	1: 155	1: 159	1: 211	1: 215	1: 219	1: 223
1: 49	1: 53	1: 57	1: 61	1: 113	1: 117	1: 121	1: 125	1: 177	1: 181	1: 185	1: 189	1: 241	1: 245	1: 249	1: 253
1: 50	1: 54	1: 58	1: 62	1: 114	1: 118	1: 122	1: 126	1: 178	1: 182	1: 186	1: 190	1: 242	1: 246	1: 250	1: 254
1: 52	1: 56	1: 60	1: 64	1: 116	1: 120	1: 124	1: 128	1: 180	1: 184	1: 188	1: 192	1: 244	1: 248	1: 252	1: 256
1: 51	1: 55	1: 59	1: 63	1: 115	1: 119	1: 123	1: 127	1: 179	1: 183	1: 187	1: 191	1: 243	1: 247	1: 251	1: 255
1: 35	1: 39	1: 43	1: 47	1: 99	1: 103	1: 107	1: 111	1: 163	1: 167	1: 171	1: 175	1: 227	1: 231	1: 235	1: 239
1: 36	1: 40	1: 44	1: 48	1: 100	1: 104	1: 108	1: 112	1: 164	1: 168	1: 172	1: 176	1: 228	1: 232	1: 236	1: 240
1: 38	1: 42	1: 46	1: 50	1: 102	1: 106	1: 110	1: 114	1: 166	1: 170	1: 174	1: 178	1: 230	1: 234	1: 238	1: 242
1: 37	1: 41	1: 45	1: 49	1: 101	1: 105	1: 109	1: 113	1: 165	1: 169	1: 173	1: 177	1: 229	1: 233	1: 237	1: 241

Fig. 7. Final result.

Acknowledgements

The first version of this paper appeared in the proceedings of the 7th Internat. Conf. on Advanced Computing and Comm. (ADCOM'99), December 20–22, 1999, Roorkee, India. The authors thank the anonymous referees for their valuable comments and suggestions, which greatly helped in improving the results.

References

- [1] S.G. Akl, *The Design and Analysis of Parallel Algorithms*, Prentice-Hall, Englewood Cliffs, NJ, 1989.
- [2] O. Egecioglu, A. Srinivasan, Optimal parallel prefix on mesh architecture, *Parallel Algorithms Appl.* 1 (1993) 191–209.
- [3] M. De, D. Das, B.P. Sinha, An new network topology with multiple meshes, Tech. Rep. No. T/Rep / E-94/01, Indian Statistical Institute, Calcutta, 1994.
- [4] D. Das, M. Dey, B.P. Sinha, A new network topology with multiple meshes, *IEEE Trans. Comput.* 68 (5) (1999) 536–551.
- [5] M. De, D. Das, M. Ghosh, B.P. Sinha, An efficient sorting algorithm on multi-mesh network, *IEEE Trans. Comput.* 46 (10) (1997) 1132–1137.
- [6] R. Cole, U. Vishkin, Faster optimal parallel prefix sum and list ranking, *Inform. and Control* 4 (1989) 334–352.
- [7] C.P. Krushkal, T. Madege, L. Rudolph, Parallel prefix on fully connected direct connection machines, in: *Proc. Internat. Conf. Parallel Process.*, 1986, pp. 278–284.
- [8] S. Lakshivarahan, S.K. Dhal, *Parallel Computing Using the Prefix Problem*, Oxford University Press, Oxford, 1994.
- [9] Y.C. Lin, C.M. Lin, Efficient parallel prefix algorithms on fully connected message passing computers, in: *Proc. 3rd Int. Conf. on High Performance Computing (HiPC)*, Trivandrum, India, December 19–22, 1996.
- [10] A. Nicolan, H. Hwang, Optimal schedule for parallel prefix computation with bounded resources, in: *Proc. 3rd ACM SIGPLAN Symp. Principles and Practice Parallel Programming*, 1991, pp. 1–10.
- [11] P.M. Kogge, Parallel solution of recurrence problems, *IBM J. Res. Development* (1974) 138–148.
- [12] R. Lander, M.J. Fisher, Parallel prefix computation, *J. ACM* 27 (1980) 831–839.
- [13] D.A. Carlson, B. Sugla, Limited width parallel prefix circuits, *J. Supercomput.* 4 (1990) 107–129.
- [14] F.E. Fich, New bounds for parallel prefix circuits, in: *Proc. 15th Symp. Theory of Computing*, 1993, pp. 100–109.
- [15] M. Snir, Depth-size trade-offs for parallel prefix computation, *J. Algorithms* 7 (1986) 185–201.
- [16] C.P. Krushkal, L. Rudolph, M. Snir, The power of parallel prefix, *IEEE Trans. Comput.* C-34 (10) (1985) 965.
- [17] P.K. Jana, Finding polynomial zeros on a multi-mesh of trees (MMT), in: *Proc. 2nd Internat. Conf. on Information Technology, Bhubaneswar, India*, December 20–22, 1999, pp. 203–206.
- [18] A. Sen, S. Bandyopadhyay, B.P. Sinha, A new architecture and a new metric for lightwave networks, *IEEE J. Lightwave Technology* 19 (2001) 913–925.