Simulation of Meshes with Separable Buses by Meshes with Multiple Partitioned Buses

Susumu MATSUMAE
Department of Information System
Tottori University of Environmental Studies
Tottori-City, Tottori 689-1111, Japan
s.matsumae@computer.org

Abstract

This paper studies the simulation problem of meshes with separable buses (MSB) by meshes with multiple partitioned buses (MMPB). The MSB and the MMPB are the mesh connected computers enhanced by the addition of broadcasting buses along every row and column. The broadcasting buses of the MSB, called separable buses, can be dynamically sectioned into smaller bus segments by program control, while those of the MMPB, called partitioned buses, are statically partitioned in advance and cannot be dynamically re-configurable. In the MSB model, each row/column has only one separable bus, while in the MMPB model, each row/column has L partitioned buses (L \geq 2). We consider the simulation and the scaling-simulation of the MSB by the MMPB, and show that the MMPB of size \( n \times n \) can simulate the MSB of size \( n \times n \) in \( O(n^{1/(2L)}) \) steps, and that the MMPB of size \( m \times m \) can simulate the MSB of size \( n \times n \) in \( O(n^{1/(2L)} \log^2 n) \) steps (\( m < n \)). The latter result implies that the MMPB of size \( m \times m \) can simulate the MSB of size \( n \times n \) time-optimally when \( m \leq n^\alpha \) holds for \( \alpha = \frac{1}{1+1/(2L)} \).

1. Introduction

The mesh architecture is one of promising models for parallel computing. Its natural structure is suitable for VLSI implementation and allows a high degree of integration. However, the mesh has a crucial drawback that its communication diameter is quite large due to lack of broadcasting mechanism. To overcome this problem, many researchers have considered adding broadcasting buses to the mesh [1, 2, 3, 8, 9, 10, 11]. Among them, in this paper we deal with the meshes with separable buses (MSB) [3, 9] and a variant of the meshes with partitioned buses [1, 2] called the meshes with multiple partitioned buses (MMPB).

The MSB and the MMPB are the mesh connected computers enhanced by the addition of broadcasting buses along every row and column. The broadcasting buses of the MSB, called separable buses, can be dynamically sectioned into smaller bus segments by program control, while those of the MMPB, called partitioned buses, are statically partitioned in advance and cannot be dynamically re-configurable. In the MSB model, each row/column has only one separable bus, while in the MMPB model, each row/column has L partitioned buses (L \geq 2).

In this paper, we show that the MMPB of size \( n \times n \) can simulate the MSB of size \( n \times n \) in \( O(n^{1/(2L)}) \) steps. This is the extension of our previous result that the MMPB of size \( n \times n \) with \( L = 1 \) can simulate the MSB of size \( n \times n \) in \( O(n^{1/2}) \) steps [4, 5]. From a theoretical point of view, since we have shown that the MSB of size \( n \times n \) can simulate the reconfigurable mesh [7, 11] (or PARBS, the processor array with reconfigurable bus systems) of size \( n \times n \) in \( O(\log^2 n) \) steps [6], we can show that any problem that is solved in \( T \) steps by the reconfigurable mesh of size \( n \times n \) can be solved in \( O(Tn^{1/(2L)} \log^2 n) \) steps by the MMPB of size \( n \times n \). It has been argued that the reconfigurable mesh can be used as a universal chip capable of simulating any equivalent-area architecture without loss in time [7], our result gives the upper bounds in time for the MMPB for simulating other equivalent-area architectures. Furthermore, we also consider the scaling-simulation problem of the MSB, and show that the MMPB of size \( m \times m \) can simulate the MSB of size \( n \times n \) in \( O(n^{1/(2L)} \log^2 n) \) steps (\( m < n \)). This result implies that the MMPB of size \( m \times m \) can simulate the MSB of size \( n \times n \) time-optimally\(^1\) when \( m \leq n^\alpha \) holds for \( \alpha = \frac{1}{1+1/(2L)} \).

This paper is organized as follows. Section 2 describes the MSB and the MMPB models. Section 3 presents an algorithm that simulates the \( n \times n \) MSB on the \( n \times n \) MMPB, and Section 4 gives an scaling-simulation algo-

\(^1\)It takes \( \Omega(\frac{n^2}{m^2}) \) steps to simulate \( n \times n \) processors using \( m \times m \) processors.
2. Models

An \( n \times n \) mesh consists of \( n^2 \) identical SIMD processors or processing elements (PE’s) arranged in a two-dimensional grid with \( n \) rows and \( n \) columns. The PE located at the grid point \((i, j)\), denoted as PE\([i, j]\), is connected via bi-directional unit-time communication links to those PE’s at \((i \pm 1, j)\) and \((i, j \pm 1)\), provided they exist \((0 \leq i, j < n)\). PE\([0, 0]\) is located in the top-left corner of the mesh. Each PE\([i, j]\) is assumed to know its coordinates \((i, j)\).

An \( n \times n \) mesh with separable buses (MSB) and an \( n \times n \) mesh with multiple partitioned buses (MMPB) are the \( n \times n \) meshes enhanced with the addition of broadcasting buses along every row and column. The broadcasting buses of the MSB, called separable buses, can be dynamically sectioned through the PE-controlled switches during execution of programs, while those of the MMPB are statically partitioned in advance by a fixed length. In the MSB model, each row/column has only one separable bus (Figure 1), while in the MMPB model each row/column has \( L \) partitioned buses for \( L \geq 2 \) (Figure 2). Those \( L \) partitioned buses of the MMPB are indexed as level-1, level-2, \ldots, level-\( L \), respectively. For each level-\( l \), the value \( \ell_l \) denotes the length of a bus segment of the partitioned bus in level-\( L \).

A single time step of the MSB and the MMPB is composed of the following three substeps:

**Local communication substep:** Every PE communicates with its adjacent PE’s via local links.

**Broadcast substep:** Every PE changes its switch configurations by local decision (this operation is only for the MSB). Then, along each broadcasting bus segment, several of the PE’s connected to the bus send data to the bus, and several of the PE’s on the bus receive the data transmitted on the bus.

**Compute substep:** Every PE executes some local computation.

\(^2\)In this paper, we assume that the partitioned buses of the MMPB are equally partitioned by the same length if they belong to the same level.
The bus accessing capability is similar to that of Common-CRCW PRAM model. If there is a write-conflict on a bus, the PE’s on the bus receive a special value ⊥ (i.e., PE’s can detect whether there is a write-conflict on a bus or not). If there is no data transmitted on a bus, the PE’s on the bus receive a special value φ (i.e., PE’s can know whether there is data transmitted on a bus or not).

3. Simulation of the $n \times n$ MSB by the $n \times n$ MMPB

In this section, we consider how to simulate a single step of the $n \times n$ MSB using the $n \times n$ MMPB. Given a single step of the simulated MSB in such a way that each PE[i, j] of the simulating MMPB knows only how corresponding PE[i, j] of the MSB behaves at this single step, we consider how to achieve the same computational task of the step on the MMPB. We assume that the computing power of PE’s, the bandwidth of local links, and that of broadcasting buses are equivalent in both the MSB and the MMPB.

In what follows, we focus on how to simulate the broadcast substep of the MSB using the MMPB, because the local communication and the compute substeps of the simulated MSB can be easily simulated in a constant time by the MMPB.

To begin with, we consider the case where $L = 2$.

Lemma 1 For any single step of the $n \times n$ MSB, the broadcasts taken on the separable bus in row $i$ (resp. column $i$) of the MSB can be simulated in row $i$ (resp. column $i$) of the $n \times n$ MMPB in $O(n^{1/4})$ time ($L = 2$).

Proof: Assume $\ell_1 = n$ and $\ell_2 = n^{1/2}$, and take any single step $S$ of the MSB and any row index $i \in \{0, 1, \ldots, n - 1\}$. Let us consider simulating the broadcasts taken on the separable bus along row $i$ of the MSB only, those on the bus along column $i$ of the MSB can be simulated similarly. To simplify the exposition, let $P_j$ and $P_j'$ respectively denote PE[i, j] of the $n \times n$ MSB and PE[i, j] of the $n \times n$ MMPB ($0 \leq j < n$).

First, we define some notations to describe the broadcasts to be simulated. To distinguish the two ports through which a PE has access to the separable bus, we refer to the port on the left side of the sectioning switch as port L and the other as port R, as shown in Figure 1. Then, the broadcasts are carried out in the following way: (1) several of $P_0, P_1, \ldots, P_{n-1}$ section the bus, (2) several of $P_0, P_1, \ldots, P_{n-1}$ send data to the bus through port L and/or R, and (3) several of $P_0, P_1, \ldots, P_{n-1}$ receive data from the bus through port L and/or R. W.r.t. these broadcasts performed in the row-separable bus of the MSB, we define $C^r_j, s^j_0, s^r_j$ (0 ≤ $j < n$, $i \in \{L, R\}$) as follows:

- $C^r_j = \{(k, y) |$ port $x$ of $P_j$ and port $y$ of $P_k$ belong to the same bus segment after the broadcasting bus being sectioned\}.
- $s^j_0 = a$ if $P_j$ sends data $a$ to port $x$, otherwise $s^j_0 = \phi$.
- $r_j^a = ($the data received by $P_j$ from port $x$).

To describe each $r_j^a$ using $C^a_j$ and $s^a_j$, we define a binary commutative operator $\oplus$ in such a way that it satisfies the following equations for any $x$ and $y$:

$$ x \oplus \phi = \phi \oplus x = x, $$
$$ x \oplus \perp = \perp \oplus x = \perp, $$
$$ x \oplus y = y \oplus x = x \quad \text{if } x = y, $$
$$ x \oplus y = y \oplus x = \perp \quad \text{if } x \neq y, x \neq \phi, \text{ and } y \neq \phi. $$

It is not difficult to confirm that $\oplus$ is well-defined and enjoys the associative law. Then, each $r_j^a$ is expressed as

$$ r_j^a = \oplus_{(k, y) \in C^r_j} s_k^y. $$

Next, we consider how to inform every $P_j'$ of the values $r_j^l$ and $r_j^R$ when every $P_j$ is given $s_j^l, s_j^R$, and the switch configuration taken by $P_j$. We divide $P_0, P_1, \ldots, P_{n-1}$ into $n^{1/2}$ disjoint blocks $B_0, B_1, \ldots, B_{n^{1/2}-1}$ in such a way that each $B_p$ consists of $P_j$ ($p n^{1/2} \leq j < (p + 1) n^{1/2}$). Let $L_{BP}$ (resp. $R_{BP}$) denote the leftmost PE (resp. the rightmost PE) of $B_p$. For each $j \in \{0, \ldots, n-1\}$ and $x \in \{L, R\}$, we let

$$ r_j^x = \oplus_{(k, y) \in C^x_j} s_k^y, $$

where $C^x_j = C^x_j \cap \{(k, y) | P_j \text{ and } P_k \text{ are in the same block and } y \in \{L, R\}\}.

Then, each $P_j'$ can compute the values $r_j^l$ and $r_j^R$ by the following 3 phases:

(Phase 1) {local simulation}

Every $P_j'$ computes the values $r_j^l$ and $r_j^R$. For each $B_p$, every $P_l'$ and $P_r'$, where $l$ is the index of $L_{BP}$ and $r$ is of $R_{BP}$, check whether the port $L$ of $L_{BP}$ and the port $R$ of $R_{BP}$ are connected in the bus-configuration of $S$.

(Phase 2) {global simulation of boundary ports}

For each $B_p$, every $P_l'$ and $P_r'$ compute the value $r_l^R$ and $r_l^k$ where $l$ is the index of $L_{BP}$ and $r$ is of $R_{BP}$.

(Phase 3) {adjustment}

Every $P_j'$ computes the values $r_j^l$ and $r_j^R$. We divide $P_0', P_1', \ldots, P_{n-1}'$ into $n^{1/2}$ disjoint blocks $B_0', B_1', \ldots, B_{n^{1/2}-1}'$ in a way that each $B_p'$ consists of $P_j'$ ($p n^{1/2} \leq j < (p + 1) n^{1/2}$). Then, for each block $B_p'$, the PE’s in $B'_p$ and a bus segment of the level-2 partitioned bus form the linear processor array of $n^{1/2}$ PE’s.
with a single broadcast bus. With these observations, we can now consider the time cost for each phase. Since the linear array of $n$ processors with a single broadcast bus can simulate the broadcast operation of the $1 \times n$ MSB (i.e., the MSB of one row and $n$ columns) in $O(\sqrt{n})$ time [6]. Phase 1 can be executed in $O(\sqrt{n^{1/2}}) = O(n^{1/4})$ time. Phase 2 is essentially the same problem as simulating the broadcast operation of the $1 \times n^{1/2}$ MSB using a linear array of $n^{1/2}$ processors with a single broadcasting bus, it can be completed in $O(\sqrt{n^{1/2}}) = O(n^{1/4})$ as well. Phase 3 can be done in $O(n^{1/4})$ time similarly to Phase 1. All the 3 phases can be executed in $O(n^{1/4})$ time. Hence, the conclusion follows.

Now, we can prove the following lemma.

**Lemma 2** For any single step of the $n \times n$ MSB, the broadcasts taken on the separable bus in row $i$ (resp. column $i$) of the $n \times n$ MSB can be simulated in row $i$ (resp. column $i$) of the $n \times n$ MMPB in $O(n^{1/(2L)})$ time ($L \geq 2$).

**Proof:** Let us consider simulating the broadcasts taken on the separable bus along row $i$ of the MSB only, those on the bus along column $i$ of the MSB can be simulated similarly.

Let $T_k(n)$ denote the time cost for simulating the broadcasts taken along the separable bus in row $i$ of the $n \times n$ MSB using row $i$ of the $n \times n$ MMPB with $L = k$ ($k \geq 2$).

We prove the lemma by Mathematical Induction.

- Base case: For $k = 2$, from Lemma 1, we have $T_2(n) = O(n^{1/4}) = O(n^{1/(2k)})$.

- Inductive case: For $k > 2$, we prove $T_k(n) = O(n^{1/(2k)})$ assuming $T_{k-1}(n) = O(n^{1/(2(k-1))})$ holds. We let $\ell_1 = n$. We modify the 3-phase simulation algorithm proving Lemma 1 in such a way that we divide $P_0, P_1, \ldots, P_{n-1}$ into $n^{1/k}$ disjoint blocks $B_0, B_1, \ldots, B_{n^{1/k}-1}$, each $B_p$ consists of $P_j (p\leq j < (p+1)n^{1/(k-1)})$. Then, we can tune $k - 1$ partitioned buses other than the level-1 bus, Phase 1 and 3 can be executed in $T_{k-1}(n^{1/(k-1)}) = O(n^{(k-1)/(2(k-1))})$ time, and Phase 2 can be completed in $O(n^{1/(2k)})$ time. Thus, we have $T_k(n) = O(n^{1/(2k)})$.

Hence, the conclusion follows.

Since the local communication and compute substeps of the MSB can be simulated obviously in a constant time by the MMPB, Lemma 2 immediately implies the following theorem:

**Theorem 1** Any single step of the $n \times n$ MSB can be simulated in $O(n^{1/(2L)})$ time by the $m \times m$ MMPB ($L \geq 2$).

---

4In Phase 2, the number of active PE’s is $2n^{1/2}$, and each of those PE’s is connected to next such PE’s via either a local communication link or a bus segment of the level-2 partitioned bus.

### 4. Scaling-Simulation of the $n \times n$ MSB by the $m \times m$ MMPB

In this section, we consider simulating a single step of the $n \times n$ MSB by the $m \times m$ MMPB ($m < n$). Let $P_{ij}$ and $P'_{ij}$ respectively denote PE($i,j$) of the MSB and PE($i,j$) of the MMPB. To simplify the exposition, we assume that $n \mod m = 0$. We define the processor mapping as follows: each $P'_{i,j}$ simulates $P_{x,y}$ ($\frac{i}{m} < x < (i + 1)\frac{1}{m}$, $\frac{j}{m} < y < (j + 1)\frac{1}{m}$).

Given a single step of the simulated MSB in such a way that each $P'_{i,j}$ knows only how each assigned $P_{x,y}$ behaves at this single step ($\frac{i}{m} < x < (i + 1)\frac{1}{m}$, $\frac{j}{m} < y < (j + 1)\frac{1}{m}$), we consider how to achieve the same computational task of the step using the MMPB. We assume that the computing power of PE’s, the bandwidth of local links, and that of broadcasting buses are equivalent in both the MSB and the MMPB.

By the results in the previous section, we can prove the following lemma.

**Lemma 3** For any single step of the $n \times n$ MSB, the broadcasts taken on the separable bus in row $i$ (resp. column $i$) of the MSB can be simulated in row $\frac{i}{m}$ (resp. column $\frac{i}{m}$) of the $m \times m$ MMPB in $O(n^{1/(L+2L)})$ time ($L \geq 2$).

**Proof:** Here, let us consider simulating the broadcasts taken on the separable bus along row $i$ of the MSB using row $\frac{i}{m}$ of the MMPB only, those on the bus along column $i$ of the MSB can be simulated similarly.

$P_{ij}$, $C_{ij}$, $s_{ij}$, and $r_{ij}$ are defined in the same way as in the proof of Lemma 1. $P'_{ij}$ is defined as the PE($\frac{i}{m}, j$) of the MMPB. We divide $P_0, P_1, \ldots, P_{m-1}$ into $m$ disjoint blocks $B_0, B_1, \ldots, B_{m-1}$, each $B_p$ consists of $P_j (p \leq j < (p + 1)\frac{1}{m})$. W.r.t. such defined blocks, $r_{ij}'$ is defined similarly to the definition in the proof of Lemma 1. Then, we modify the 3-phase simulation algorithm proving Lemma 1 as follows:

**(Phase 1) \{local simulation\}**

Each $P'_p$ computes the values $r_{ij}'$ and $r_{ij}'R$ for $j \in \{ h \mid P_h \text{ is in } B_p \}$, and checks whether the port $L$ of LP$_{B_p}$ and the port $R$ of RP$_{B_p}$ are connected in the bus-configuration.

**(Phase 2) \{global simulation of boundary ports\}**

Every $P'_p$ computes the value $r_{ij}'$ and $r_{ij}'R$ where $l$ is the index of LP$_{B_p}$ and $r$ is the index of RP$_{B_p}$.

**(Phase 3) \{adjustment\}**

Every $P'_p$ computes the values $r_{ij}'$ and $r_{ij}'R$ for $j \in \{ h \mid P_h \text{ is in } B_p \}$.
Phase 1 and 3 can be completed in $O\left(\frac{n}{m}\right)$ time because each block $B_p$ is simulated by $P_p^L$ alone and we have only to sequentially scan the data of size $O\left(\frac{n}{m}\right)$. Phase 2 is essentially the same problem as simulating the broadcast operation of the $1 \times m$ MSB using the $1 \times m$ MMPB, hence it can be done in $O(m^{1/(2L)})$ time from Theorem 1. Thus, the conclusion follows.

Now, we obtain the following theorem.

**Theorem 2** Any single step of the $n \times n$ MSB can be simulated by the $n \times n$ MMPB in $O\left(\frac{n}{m} \left(\frac{m}{n} + m^{1/(2L)}\right)\right)$ time ($L \geq 2$).

**Proof**: The MMPB can simulate the broadcast substep of a single step of the MSB, by first simulating the broadcasts taken along rows, and then simulating those along columns similarly. Since each row (resp. column) of the MMPB has to simulate $\frac{n}{m}$ rows (resp. columns) of the MSB, this takes $O\left(\frac{n}{m} \left(\frac{m}{n} + m^{1/(2L)}\right)\right)$ time from Lemma 3. As for the local communication and compute substeps, the MMPB can simulate them in $O\left(\frac{n}{m^2}\right)$ steps in each PE. Thus, the conclusion follows.

## 5. Concluding Remarks

We consider the simulation and the scaling-simulation problem of the MSB by the MMPB, and obtained the following results:

1. The MMPB of size $n \times n$ can simulate the MSB of size $n \times n$ in $O(n^{1/(2L)})$ steps ($L \geq 2$).

2. The MMPB of size $m \times m$ can simulate the MSB of size $n \times n$ in $O\left(\frac{n}{m} \left(\frac{m}{n} + m^{1/(2L)}\right)\right)$ steps ($m < n$, $L \geq 2$).

The latter implies that the MMPB of size $m \times m$ can simulate the MSB of size $n \times n$ time-optimally when $m \leq n$ holds for $\alpha = \frac{n}{m^{1+1/(2L)}}$.

From a practical view point, compared to the MSB, the MMPB model has the advantage that the propagation delay introduced by the length of the bus (signal propagation delay) and those switch elements inserted to the bus (device propagation delay) can be small, and hence our simulation algorithms are useful when the mesh size becomes so large that we cannot neglect the delay.

### Acknowledgment

This work is supported by the MEXT Grant-in-Aid for Young Scientists (B) (13780231).

### References