Bus based parallel computers: A viable way for massive parallelism

A. Ferreira* A. Goldman vel Lejbman** S. W. Song***

Abstract. In most distributed memory MIMD multiprocessors, processors are connected by a point-to-point interconnection network. Since interprocessor communication frequently constitutes serious bottlenecks, several architectures were proposed that enhance point-to-point topologies with the help of multiple bus systems so as to improve the communication efficiency. In this paper we study global communication on parallel architectures where the communication means are constituted *solely* by busses. We focus on the *hyperpath* and the *hypergrid* architectures, which are the bus-based versions of the well used point-to-point linear and grid interconnection networks. Using (hyper) graph theoretic concepts in order to model inter-processor communication in such networks, we developed a new tool called *simplification* in order to give extremely efficient algorithms for gossiping (all-to-all or total exchange communications).

1 Introduction

In most distributed memory MIMD multiprocessors, processors are connected by a point-to-point interconnection network. Since interprocessor communication frequently constitutes serious bottlenecks, several architectures were proposed to *enhance* point-to-point topologies with the help of multiple bus systems so as to improve the communication efficiency.

In this paper we study massively parallel systems where processors are connected by using *solely* by busses. These architectures can use the power of bus technologies, providing a way to interconnect much more processors in a simple and efficient manner. This represents a major issue in the construction of parallel computers, because a multiprocessor system with a large number (on the order of thousands) of processors cannot be provided with full connectivity and need to operate by communication between processors. Furthermore, the multiple bus

* CNRS - Laboratoire de l'Informatique du Parallélisme, École Normale Supérieure de Lyon, 46, Allée d'Italie, 69364 Lyon Cédex 07 - France. Partially supported by the PRC C^3 and ANM of the French CNRS.

** Department of Computer Science, Institute of Mathematics and Statistics, University of São Paulo, C.P. 20570 - São Paulo, SP 01498-970 - Brazil. Supported by FAPESP - Proc. No. 92/3991-0

*** Department of Computer Science, Institute of Mathematics and Statistics, University of São Paulo, C.P. 20570 - São Paulo, SP 01498-970 - Brazil. Supported by FAPESP - Proc. No. 93/0603-1 and CNPq - Proc. No. 306063/88-3 and PROTEM-CC/SP.

architecture has the desirable ability to expand. While a crossbar architecture needs a quadratic number of switches, a multiple bus system of P processors and B busses requires only BP switches, though with reduced connectivity. Without the underlying point-to-point network, such architectures are feasible and suitable for VLSI implementation, being able to interconnect a lot more processors, while keeping the diameter of the network small (see, for instance, [14] for a cell design incorporating a processor, its local memory and bus switches). For instance, 8100 processors can compose a 2-dimensional hypergrid – as defined in the next section –, where each processor is connected to at most four busses, each bus has 30 processors, and the distance between the farthest appart processors is as low as 10. Notice that this is better than a usual hypercube of the same size, and not as difficult to build, since it requires only 36 busses.

Common networks are usually modeled by a graph where processors are nodes and communication links are edges. On the other hand, bus interconnection networks (BINs, for short) are best modeled by a *hypergraph* [2], in which an edge is a subset of nodes. Each bus in the network is thus represented by an edge of the hypergraph [3, 13]. Several papers used derived models to propose and study new classes of interconnection networks for massively parallel computers. In [12], Scherson proposes the concept of "orthogonal graphs", that can describe, for instance, binary m-cubes or meshes with busses in rows and columns. He also studies the routing problem in such architectures. Fiduccia [8] tackles the problem of pin optimality of interconnection networks, concluding that bused n-node hypercubes, where n busses run along the $\log n$ dimensions of the hypercube, are pin-optimal. Hypergraphs are used by Szymanski [13] to model some photonic networks, where processors communicate through optical fibers carrying laser beams that have different wavelengths. Although little work has been done concerning algorithms in these new models, Bermond and Ergincan propose an excellent survey on structural problems related to BINs [4].

Another approach is to enhance existing point-to-point networks with multiple bus systems [1, 5, 10, 14]. However this is out of the scope of this paper.

For deriving efficient communication patterns in BINs we developed a new tool called *simplification*. The idea is to construct a graph, to be called *representative graph*, from the original hyper-topology, in such a way that it will become easy to describe and perform communication schemes to the former that will fit to the latter, because the simplification concept also allows us to partially use some already known communication algorithms for usual networks.

By generalizing the notion of a linear network, we can design a BIN with the following features. Let $\mathcal{P}_{n,m}$ be a parallel computer having nm processors, $n-1$ busses, with $2m$ processors per bus, each processor being connected to at most two busses. In such an architecture, we show how to implement a gossiping (all-to-all) operation in at most $m+n-2$ steps. If we start from a 2-d grid, our BIN will be denoted $\mathcal{G}_2 = \mathcal{P}^1_{n_1,m_1} \times \mathcal{P}^2_{n_2,m_2}$. This is a parallel computer with $n_1m_1n_2m_2$ processors, n_1n_2 busses, each bus with $2m_1$ or $2m_2$ processors, where each processor is connected to at most four busses. We show that, in $\mathcal{G}_2$, we can complete gossiping within $n_1+n_2+m_1+m_2$ rounds.

The algorithms presented here are optimal for hyperpaths, since they asymptotically match the lower bound given by the diameter of the network. On the other hand, we conjecture that our gossiping algorithms for hypergrids are also asymptotically optimal. Because of the methods developed in this paper, proving its optimality would imply proofs of optimality for gossiping in point-to-point networks, that are open problems at the time of this writing. Moreover, proving tight lower bounds for BINs would need new tools because of the degree of freedom introduced by multiple bus communications.

Our paper is organized as follows. In the next Section we discuss some preliminary material on hypergraphs. In Section 3 we state the models of communication best suited for BINs. Section 4 introduces our main tool, called simplification, to deal with global communication in such massively parallel systems. Then, in Section 5, we thoroughly study gossiping schemes in hyperpath and d-hypergrid BINs. Due to space limitations, this text has been shortened. We refer the interested reader to its full version in [7].

2 Preliminaries

A graph is denoted $G(V, E)$, where V is a set of nodes or vertices $\{v_1, \ldots, v_n\}$ and E is a set of edges $\{e_1, \ldots, e_m\}$, where each edge $e_i \subset V, |e_i| = 2$. A hypergraph shall be denoted $\mathcal{G}(V, E)$, where V is a set of nodes $\{v_1, \ldots, v_n\}$ and E is a set of hyperedges $\{e_1, \ldots, e_m\}$ where each hyperedge $e_i \subseteq V \mid \bigcup_{e_i \subset E} = V$. Notice that a hypergraph $\mathcal{G}(V, E)$ where each hyperedge $e \in E$ has cardinality 2, is a graph, and that every graph $G(V, E)$ where $\bigcup_{e \in E} = V$ is a hypergraph.

Given a hypergraph $\mathcal{G}$, denote by $D(\mathcal{G})$ its diameter. Define *transmission* as the sending of a message along a bus from one node to one or more adjacent nodes. For a question of uniformity, in this paper we use graphs and hypergraphs, with the analogy of processor–node, communication link–edge, and bus–hyperedge.

2.1 Hyperpaths

Topology: A path of size n is defined as $P_n(V, E)$ where: $V = \{0, \ldots, n-1\}$ and $e \in E \Leftrightarrow e = \{v_1, v_2\} \mid v_1 = (v_2 - 1)$.
Generalization: We define hyperpath $\mathcal{P}_{n,m}(V, E)$ of size n and order m by:

$$V = \{0, \ldots, nm - 1\}$$

$$e \in E \Leftrightarrow e = \{v_1, \ldots, v_m, v'_1, \ldots, v'_m\} \text{ where } \begin{cases} v_i = v_j \bmod n \\ v'_i = v'_j \bmod n \\ v_i = (v'_i - 1) \end{cases}$$

2.2 Hypergrids

We define a d-dimensional hypergrid (or simply, d-hypergrid) as $\mathcal{G} = \mathcal{P}^1_{n_1,m_1} \times \ldots \times \mathcal{P}^d_{n_d,m_d}$ where $\mathcal{P}^i_{n_i,m_i}$ are hyperpaths. For this a formal definition of the Cartesian Product for hypergraphs is needed. Below, we define 2-hypergrids, the definition of d-hypergrids being analogous.

Definition 1. Let $\mathcal{P}^1_{n_1,m_1}$ and $\mathcal{P}^2_{n_2,m_2}$ be two hyperpaths as defined above. The 2-dimensional hypergrid $\mathcal{G}(V,E) = \mathcal{P}^1_{n_1,m_1} \times \mathcal{P}^2_{n_2,m_2}$ is as follows.

$V = \{u_k\}, 0 \leq k < n_1 m_1 n_2 m_2$, such that $u_k = (v_i, v'_j)$ with $0 \leq i < n_1 m_1$ and $0 \leq j < n_2 m_2$.

$E = \{e_q\}, \mid e_q \mid = 2m_1$, or $\mid e_q \mid = 2m_2$, such that two vertices (a,b) and (c,d) are adjacent in $\mathcal{G}$ if and only if $a = c$ *and* b, d *are adjacent in* $\mathcal{P}^2_{n_2,m_2}$ or $b = d$ *and* a, c *are adjacent in* $\mathcal{P}^1_{n_1,m_1}$

(Figure 1 shows an example.) Notice that, as for graphs, this is not the only possible definition for the Cartesian Product. It could be interesting to study the impact of other such definitions.

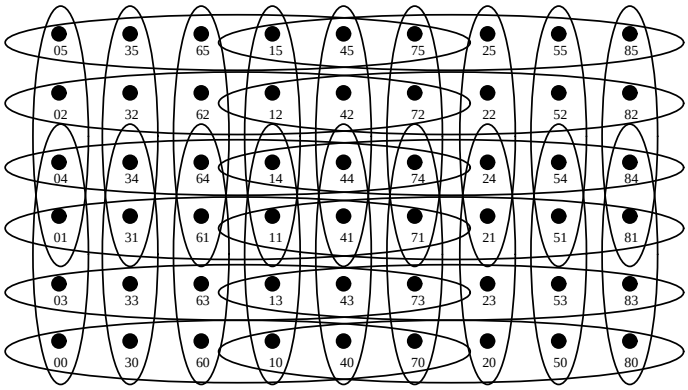


Fig. 1. Hypergrid $\mathcal{G} = \mathcal{P}^1_{3,3} \times \mathcal{P}^2_{3,2}$.

3 Modeling bus access and capacity

We first recall the usual notation used for modeling communications in point-to-point interconnection networks [9]. Given a graph $G(V,E)$, there exists a communication model for edges and another for nodes. Relative to the edges, when data can be transmitted in only one direction in each cycle, we have the *half-duplex* model (H). In the case where information can traverse in both directions simultaneously, we have the *full-duplex* model (F). If a node can use (send/receive) only one of its edges simultaneously, we have the 1-port model; if it can use k edges we have the k-port model. If a node can use all its edges in one same cycle, we have the $*$-port model. Thus, the notation E_k is used, where E is the model for edges and k is the model for nodes (*e.g.*, F_k stands for full-duplex, k-port, and H_1 stands for half-duplex, 1-port).

Our models for BIN's stem from the observation that, as far as VLSI implementations are concerned, an edge being half or full duplex means that a communication link can carry one or two different pieces of information, implying that one or two physical links have to be implemented between the two processors. Therefore, in order to model bus communications, we can assume that

a bus, as a communication link, can carry more than one information. Theoretically speaking, a bus could support as much information transfer as the number of vertices it connects.

Definition 2. Given a BIN $\mathcal{G}(V, E)$, inter-processor communications are modeled as follows.

Bus capacity: if only one information can traverse in a bus in one cycle, we call it the `1-capacity` (or `single-capacity`) model; if k information can traverse at the same time, we call it the `k-capacity` model.

Bus access: if a processor can use (send/receive) simultaneously only one of its busses, we call it the `1-port` model; if it can use k (all) of its busses at the same time, we call it the `k-port` (`*-port`) model.

3.1 Notation

We use the notation $M_{\alpha,\rho}$ where α is the model of bus capacity and ρ is the model for bus access. Given a graph $G(V, E)$ and two adjacent nodes $u, v \in V$, we use the notation $u \leftarrow v$ to denote the transmission of node v to node u. Given a hypergraph $\mathcal{G}(V, E)$, $v \in V$, $e \in E$, with $v \in e$ and given $N \subseteq e$, we denote by $N \Leftarrow v$ a transmission from node v to the set of nodes N.

Due to VLSI feasibility concerns, as well as to main features in global communication techniques, in this paper we focus our attention to gossiping (all-to-all) schemes for BINs, under the model $M_{i,j}, i = 1$ or 2, $j = 1$ or *. Hence, busses can transport at most two informations in a single time step, and nodes are able to use one or all of its busses in a single cycle.

4 Simplification

A parallel computer built according to the model proposed above has to prove its utility in overcoming communication bottlenecks. In the following, we develop a tool to help us carry out standard global communication operations in the previously defined BINs.

The idea is to use a new concept called *simplification* in order to construct a graph, to be called *representative graph*, from the hypergraph defining the BIN, in such a way that it will become easy to describe and perform communication in the hypergraph [4]. Furthermore, we also want to make use of the already known communication algorithms for point-to-point interconnection networks.

The simplification operation is done basically by choosing nodes in the hypergraph, called *representative nodes*, to be the nodes of the representative graph $G(V', E')$. Adjacent representative nodes in the hypergraph will be adjacent in the representative graph. Each representative node u *represents* a set of nodes N_u. The details will be given henceforth.

[4] We remark that "representative graph" does not have here the same meaning as in [2], where it denotes the line graph of a hypergraph.

We will deal with communication in the representative graph as follows. A transmission from a node v to an adjacent node x, in the representative graph, is equivalent to a transmission from the node corresponding to v in the hypergraph, to all the nodes of the hyperedge that contains the nodes corresponding to v and to x. A transmission to a representative node in the representative graph also means transmission to nodes it represents in the hypergraph. That is, $u \leftarrow v$ in the representative graph also implies $N_u \Leftarrow v$ and $N_v \Leftarrow v$ in the hypergraph.

It is important to point out that the simplification operation depends not only on the hypergraph, but also on the communication model adopted. Below we show two of the main forms for simplification of hyperpaths and d-dimensional hypergrids.

4.1 Simplification for the model $M_{j,i}$, $(j \leq i)$

We use the simplification in the hypergraphs in such a way that it will be easy to emulate the communication in the hypergraph under the model $M_{j,i}, (j \leq i)$ through the representative graph under the model H_i, if $j = 1$, or under the model F_i, if $j = 2$. The main idea is to choose a representative node in the intersection of hyperedges.

Representative graph $G(V', E')$ of a hyperpath $\mathcal{P}_{n,m}(V, E)$.
We elect $n-1$ representatives $V' = \{u_1, \ldots, u_{n-1}\}$, with $u_i \in V$ such that:

$$\begin{cases} u_1 \bmod n = 0 \\ u_{i+1} = u_i + 1, \text{ for } 1 \leq i \leq n-2 \end{cases}$$

The nodes of G are adjacent if and only if they are adjacent in $\mathcal{P}_{n,m}$. In this way, we will have one representative for each intersection of hyperedges. In addition to these, we have one representative in *one* of the extremes of $\mathcal{P}_{n,m}$. The sets of represented nodes are such that, for each $u \in V'$, $N_u = \{y_1, y_2, \ldots, y_{m-1}\}$

$$\begin{cases} (y_i - u) \bmod n = 0 \\ y_i \neq u, 1 \leq i \leq m-1 \end{cases}$$

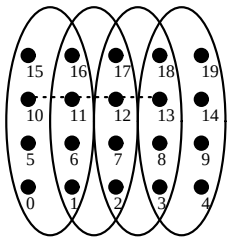


Fig. 2. The representative graph for the hyperpath $\mathcal{P}_{5,4}$ has the nodes $\{10, 11, 12, 13\}$, the edges are in dotted lines. We have also: $N_{10} = \{0, 5, 15\}$, $N_{11} = \{1, 6, 16\}$, $N_{12} = \{2, 7, 17\}$, $N_{13} = \{3, 8, 18\}$ and $W = \{4, 9, 14, 19\}$.

Representative graph $G(V', E')$ of a hypergrid $\mathcal{G} = \mathcal{P}^1_{n_1,m_1} \times \ldots \times \mathcal{P}^d_{n_d,m_d}$.

We perform simplification by dimensions. We first choose the representative nodes for the hyperpath $\mathcal{P}^1_{n_1,m_1}$. Let $G_1(V'^{(1)}, E'^{(1)})$ be its representative graph. By doing the same for the remaining paths, we will have $\mathcal{G}^{(d)} = G_1 \times \ldots \times G_d$.

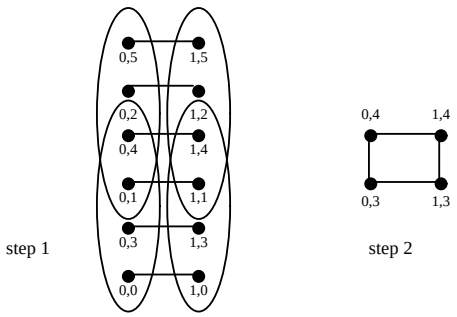


Fig. 3. The 2 steps of simplification for the 2-hypergrid $\mathcal{G} = \mathcal{P}^1_{3,3} \times \mathcal{P}^2_{3,2}$.

4.2 Simplification for the model $M_{2,1}$

In this case we want to model the communication in the hypergraph through a representative graph under the full-duplex model, where each edge can simultaneously carry information in both directions. Instead of picking only one representative node in each intersection of hyperedges in the hyperpath, we now choose two of them. For the d-dimensional hypergrid we have two different possible simplifications: one by choosing two representative nodes per intersection, the other by choosing $2d$ representative nodes in each intersection of hyperedges. The two simplifications will be denominated *type* F_1 and *type* F_*, respectively. For the following definitions of representative graphs, we assume, without loss of generality, that there exist as many nodes in hyperedge intersections as necessary.

We build the *Representative graph* $G(V', E')$ of a hyperpath $\mathcal{P}_{n,m}(V, E)$ as follows. We elect $2(n-1)$ nodes, $V' = \{u_1, \ldots, u_{n-1}, u_n, \ldots, u_{2(n-1)}\}$, such that:

$$\begin{cases} (u_1 \bmod n) = (u_n \bmod n) = 0 \\ u_{i+1} = u_i + 1, 1 \leq i \leq n-2, n \leq i \leq 2n-3 \\ u_1 \neq u_n \end{cases}$$

Nodes of G are adjacent if and only if they are adjacent in $\mathcal{P}_{n,m}$. We define the set of represented nodes N_z, for each $z \in V', N_z = \{y_1, \ldots, y_m\}$, such that $(y_i - z) \bmod n = 0, 1 \leq i \leq m$.

Note that now z belongs to N_z and that for a given $i, N_{u_{i+(n-1)}} = N_{u_i}$. Then, we say in this case that u_i and $u_{i+(n-1)}$ are *equivalent*. Let $N = \bigcup_{u \in V'} N_u$, we define the set of nodes without representative $W = V \backslash N$ (figure 4).

The *Representative graph* $G(V', E')$ of type F_1 of a d-dimensional hypergrid $\mathcal{G} = \mathcal{P}^1_{n_1,m_1} \times \ldots \times \mathcal{P}^d_{n_d,m_d}$ is built similarly, again dimension after dimension. As before, we can define similarly the sets of represented nodes, and the sets of nodes without representative. See Figure 5 for an example.

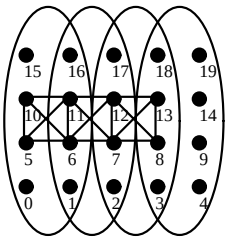
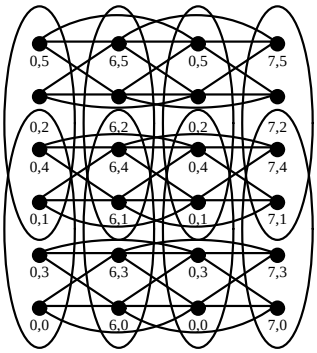


Fig. 4. Simplified graph for the model $M_{2,1}$ of the hyperpath $\mathcal{P}_{5,4}$.

step 1



step 2

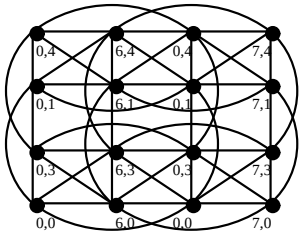


Fig. 5. Simplification for the model $M_{2,1}$ of the 2-hypergrid $\mathcal{G} = \mathcal{P}^1_{2,3} \times \mathcal{P}^2_{2,2}$.

4.3 Reduction

Observe that in the case where we generate the representative graph $G(V', E')$ of $\mathcal{G}(V, E)$ by choosing two representative nodes per intersection of hyperedges, for each node $v \in V'$ there exists a single node $u \in V', u \neq v$, such that u and v are equivalent. Furthermore, if the nodes u and v are equivalent, then they belong to the same hyperedge of $\mathcal{G}(V, E)$.

We can then define the reduced representative graph, or simply *reduced graph*, G^* with regard to $G(V', E')$, where each pair of equivalent nodes in $G(V', E')$ corresponds to a single node in G^*. Two nodes are adjacent in the latter only if they were adjacent in the former.

Let now the relation $(G_1, \lambda) \prec (G_2, \varphi)$ indicate that a communication scheme in graph G_1 under model λ can be implemented within the same number of steps in graph G_2 under model φ. Then, the following holds (proofs are straightforward, hence omitted). (See an example in Figure 6.)

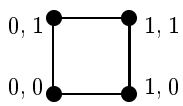


Fig. 6. Reduced graph G^* of the hypergrid $\mathcal{G} = \mathcal{P}^1_{2,3} \times \mathcal{P}^2_{2,2}$.

Claim 1 *If every pair of equivalent vertices have the same information, then* $(G^*, F_1) \prec (G, H_1)$.

Claim 2 *Provided that at most two information are sent within a step, then* $(G, any) \prec (\mathcal{G}, M_{2,1})$.

Claim 3 $(G, H_1) \prec (\mathcal{G}, M_{2,1})$.

The notion of reduced graph can be extended to the case where there are more than two equivalent nodes per hyperedge intersection (representative graph of type F_*). If there exist x equivalent nodes in each hyperedge intersection, we define G^* in a similar way, each of the equivalent nodes corresponding to a single node in G^*. The reduced graph of a representative graph of type F_* for a d-dimensional hypergrid can be obtained analogously, with $x = 2d$. Thus, we have the following claim.

Claim 4 *If the reduced graph* G^* *is obtained from a representative graph of type* F_**, and all vertices that are equivalent have the same information, then* $(G^*, H_x) \prec (\mathcal{G}, M_{1,1})$.

Therefore, from a hyperpath or a d-dimensional hypergrid under model $M_{1,1}$, we can create a simplification that will result in a reduced graph where the communication will work under H_{2d}. It suffices to elect $2d$ representatives per intersection ($d = 1$ for the hyperpath).

5 Communication patterns

Given a hypergraph $\mathcal{G}(V', E')$ and its representative graph $G(V, E)$, we denote by $gossiping_M(G)$ the gossiping operation, under model M, in graph G. Each transmission $v \leftarrow u$ in graph G represents a transmission from u to all the nodes of the hyperedge that contains u and v in $\mathcal{G}$. We denote by $D(G)$ (respectively, $D(\mathcal{G})$) the diameter of G (resp., $\mathcal{G}$). For the description of the algorithms we use the following instructions: for $i = 0$ to $n - 1$, for each $v \in V$, for all $v \in V$, for two $v \in V$.

5.1 Gossiping

We have three distinct phases: *initial, central and final.* The initialization consists of gathering the information from the represented nodes into the representatives. The central phase is the gossiping in the representative graph. Finalization consists of sending off the final information (which is the union of all the information initially present in the nodes) to all the nodes of the hypergraph that still do not contain it.

Hyperpath $\mathcal{P}_{n,m}$.
Model $M_{1,i}$. Let $G(V', E')$ be the representative graph for model $M_{1,i}$, with $N_u, u \in V'$, the set of represented nodes, and W the set of nodes without representative. We have the following algorithm:

Initial	Central	Final
for all $u \in V'$ for each $v \in N_u, u \leftarrow v$ for each $v \in W, u_{n-1} \leftarrow v$	$gossiping_{H_i}(G)$	$W \Leftarrow u_{n-1}$

Analysis. After the Initial phase, every representative has all the information from the represented nodes, with the exception of u_{n-1}, that has all the information from all the vertices in the last bus. Thus, by Claim 4, applying a gossiping scheme on G under model H_i (Central phase above) is almost equivalent to gossiping in the hyperpath under model $M_{1,i}$. Only the Final phase is missing in order to take the information to the non-represented nodes. Notice that a straightforward lower bound stems from the fact that the complete information in an extreme bus takes at least $2m$ steps to reach its neighbour bus, and then it needs to cross all the remaining busses to be sure that every node is informed. This amounts to $2m + D(\mathcal{P}_{n,m}) - 1$.

Theorem 3. *Under model $M_{1,i}$, the number of steps for gossiping in a hyperpath $\mathcal{P}_{n,m}$ is at most $(2m-1)+(D(G)+1)+1 = 2m+D(\mathcal{P}_{n,m})$, which is asymptotically optimal.*

Model $M_{2,1}$. For the model $M_{2,1}$, let $G(V', E')$ and G^* be the representative and reduced graphs, respectively, with representative nodes $V' = \{u_1, \ldots, u_{n-1}, u_n, \ldots, u_{2(n-1)}\}$, $N_u, u \in V'$, the set of represented nodes, and W the set of nodes without representative.

The idea is the same as the one depicted above. First, the information from the represented nodes are sent to their representatives. Then, a gossiping in the representative graph is performed, following the specific simplification used. Finally, the nodes without representatives are informed. In this case, since each bus has capacity 2, the lower bound can be derived in a similar way as above, and is equal to $m + D(\mathcal{P}_{n,m}) - 1$. The algorithm follows.

Initial	Central	Final
for all $u_i \in V' \mid 1 \leq i \leq n-1$ for two $v \in N_{u_i}, \{u_i, u_{i+n-1}\} \Leftarrow v$ for two $v \in W, \{u_{n-1}, u_{2(n-1)}\} \Leftarrow v$	$gossiping_{F_1}(G)$	$W \Leftarrow u_{n-1}$

Analysis. This analysis is analogous to the previous one. The only critical point is that phase 2 has to be "clever", but claim 2 guarantees its correction, yielding the following

Theorem 4. *Under model $M_{2,1}$, the number of steps for gossiping in a hyperpath $\mathcal{P}_{n,m}$ is at most $m + D(G) + 1 = m + D(\mathcal{P}_{n,m})$, which is asymtotically optimal.*

Using an algorithm similar to that of model $M_{1,1}$ with $gossiping_{F_*}(G)$ in the Central phase, we have:

Corollary 5. *Under model $M_{2,2}$, the number of steps for gossiping in a hyperpath $\mathcal{P}_{n,m}$ is at most $m - 1 + (D(G)) + 1 = m + D(\mathcal{P}_{n,m}) - 1$, which is optimal.*

Hypergrid $\mathcal{G} = \mathcal{P}^1_{n_1,m_1} \times \ldots \times \mathcal{P}^d_{n_d,md}$
Model $M_{1,i}$. Gossiping goes as follows. The Initial phase proceeds dimension after dimension, having initialization done for each hyperpath. After completion, all information from the d-dimensional hypergrid have been collected by the nodes in its representative graph. Then, a usual gossiping in a d-dimensional grid is called. In the Final phase, all representative processors have finished gossiping and can send the information back to the represented nodes (notice that this also implies that the vertices without representatives will also get the information, since they belong to the same bus being used). This is done again dimension after dimension, to ensure that all processor get the information, without collisions. Below, we give a somewhat more formal description of the algorithm.

Initial	**Central**	**Final**
for $j = 1$ to d for all $x^{(j)} \in \mathcal{G}^{(j)}$ for each $v \in N_x^{(j)}, x^{(j)} \leftarrow v$ for each $v \in W^{(j)}, \alpha(v) \leftarrow v$	$gossiping_{H_i}(G)$	for $j = d$ to 1 for all $x \in \mathcal{G}^{(j)}, N_x \Leftarrow x$ (transmissions to $W^{(j)}$)

Analysis. We use $\alpha(v)$ to denote the representative node(s) adjacent to v. The Initial phase takes $2m_i - 1$ steps per dimension. The Central phase takes $D(G)$ steps even under the weakest model H_1, provided that $\mathcal{G}$ is large enough [6]. The Final phase takes 1 step per dimension.

Proposition 6. *Under model* $M_{1,i}$*, the number of steps for gossiping in a* d*-dimensional hypergrid* $\mathcal{G} = \mathcal{P}^1_{n_1,m_1} \times \ldots \times \mathcal{P}^d_{n_d,md}$ *is at most* $\sum_{i=1}^{d}(2m_i - 1) + D(G) + d = 2\sum_{i=1}^{d} m_i - d + D(\mathcal{G})$.

Since the Central phase is always optimal for sufficiently large hypergrids [6], and the Final phase takes only one step per dimension, we can only expect to improve the time complexity of the Initial phase, by using more powerful models.
Model $M_{2,1}$. We shall use the representative graph of type F_k (where $k = 1$, or $k = 2d$) and G^*, the corresponding reduced graph, in order to speed up the transmission during the Initial phase.

Initial $M_{2,1}$	**Initial** $M_{2,d}$
for $j = 1$ to d for all $x^{(j)} \in \mathcal{G}^{(j)}$ for two $v \in N_{x^{(j)}}, E(v) \leftarrow v$ for two $v \in W^{(j)}, \alpha(v) \Leftarrow v$	for $j = 1$ to d for all $x^{(j)} \in \mathcal{G}^{(j)}$ for two $v \in N_{x^{(j)}}, x^{(j)} \leftarrow v$ for two $v \in W^{(j)}, \alpha(v) \leftarrow v$

Analysis. We use $E(v)$ to denote the hyperedge that contains the representative of v. The Initial phase now completes in $\sum_{i=1}^{d} m_i$.

Proposition 7. *Under model* $M_{2,1}$*, the number of steps for gossiping in a* d*-dimensional hypergrid* $\mathcal{G} = \mathcal{P}^1_{n_1,m_1} \times \ldots \times \mathcal{P}^d_{n_d,md}$ *is at most* $\sum_{i=1}^{d} m_i + D(\mathcal{G})$.

Model $M_{2,d}$. The number of rounds in the Initial phase remains the same because it depends on the bus capacity, and not on the bus access capability for each processor. The control of the algorithm, however, is much simpler.

Proposition 8. *Under model* $M_{2,d}$*, the number of steps for gossiping in a* d*-dimensional hypergrid* $\mathcal{G} = \mathcal{P}^1_{n_1,m_1} \times \ldots \times \mathcal{P}^d_{n_d,md}$ *is at most* $\sum_{i=1}^{d} m_i + D(\mathcal{G})$.

References

1. A. Aggarwal, Optimal bounds for finding maximum on array of processors with k global busses, *IEEE Transactions on Computers* **35** (1986) 62–64.
2. C. Berge, *Hypergraphs*, (North Holland, 1989).
3. J-C. Bermond, J. Bond and C. Peyrat, Interconnection networks with each node on two busses, in *Parallel Algorithms and Architectures*, eds. M. Cosnard, Y. Robert, P. Quinton and M. Tchunte, (North Holland, 1984) 155–167.
4. J-C. Bermond and F.O. Ergincan, Bus interconnection networks, *Technical Report No. 93-56*, Université de Nice - Sophia Antipolis, September 1993.
5. L.N. Bhuyan and D.P. Agrawal, Generalized hypercube and hyperbus structures for a computer network, *IEEE Transactions on Computers* **33** (1984) 323–333.
6. G. Cybenko, D.W, Krumme, and K.N. Venkataraman, *Gossiping in minimum time*, SIAM Journal on Computing, Vol.21, no.1, pp. 111-139, Feb 1992.
7. A. Ferreira, A. Goldman vel Lejbman, and S. W. Song. Towards simple massively parallel systems: Bus based parallel computers and new communication patterns. Reserach Report 93-18, Laboratoire de l'Informatique du Parallélisme, Ecole Normale Supérieure de Lyon, France, June 1993.
8. C.M. Fiduccia, Bused hypercubes and other pin-optimal networks, *IEEE Transactions on Parallel and Distributed Systems* **3(1)** (1992) 14–24.
9. P. Fraigniaud and E. Lazard, Methods and problems of communication in usual networks, Rapport de Recherche LIP 91-33, LIP/ENS-Lyon, October, 1991, to appear in Discrete Applied Mathematics.
10. H. Jiang and K.C. Smith, PPMB: a partial-multiple-bus multiprocessor architecture with improved cost-effectiveness, *IEEE Transactions on Computers* **41** (1992) 361–366.
11. S. Olariu, J.L. Schwing and J.Y. Zhang, On the power of two-dimensional processor arrays with reconfigurable bus systems, *Parallel Processing Letters* **1** (1991) 29–34.
12. I.D. Scherson, Orthogonal graphs for the construction of a class of interconnection networks, *IEEE Transactions on Parallel and Distributed Systems* **2(1)** (1991) 3–18.
13. T. Szymanski, Graph-theoretic models for photonic networks, *Proceedings of the New Frontiers: A Workshop on Future Directions of Massively Parallel Processing*, McLean, VA, 1992, 85–96.
14. B. Wilkinson, On crossbar switch and multiple bus interconnection networks with overlapping connectivity, *IEEE Transactions on Computers* **41** (1992) 738–746.

This article was processed using the LaTeX macro package with LLNCS style