Broadcasting in bus interconnection networks

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 parallel architectures where the communication means are constituted *solely* by buses. These architectures can use the power of bus technologies, providing a way to interconnect much more processors in a simple and efficient manner. We present the *hyperpath*, *hypergrid*, *hyperring*, and *hypertorus* architectures, which are the bus-based versions of the well used point-to-point interconnection networks. Using (hyper) graph theoretic concepts to model inter-processor communication in such networks, we give optimal algorithms for broadcasting a message from one processor to all the others. For deriving high performance communication patterns we developed a new tool called *simplification*.

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 buses. 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 architecture has the desirable ability to expand.

* 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

*** Same address as above. Supported by FAPESP - Proc. No. 93/0603-1 and CNPq - Proc. No. 306063/88-3 and PROTEM-CC/SP.

Previous work 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* [3], in which an edge is a subset of nodes. Each bus in the network is thus represented by an edge of the hypergraph [4, 18]. Several papers used derived models to propose and study new classes of interconnection networks for massively parallel computers. In [16], Scherson proposes the concept of *orthogonal graphs*, that can describe, for instance, binary m-cubes or meshes with buses in rows and columns. Fiduccia [10] tackles the problem of pin optimality of interconnection networks, concluding that bused n-node hypercubes, where n buses run along the $\log n$ dimensions of the hypercube, are pin-optimal. Hypergraphs are used by Szymanski [18] to model some photonic networks. Although little work has been done concerning algorithms in these new models, Bermond and Ergincan present an excellent survey on structural problems related to BINs [5].

Another approach is to enhance existing point-to-point networks with multiple bus systems. Wilkinson [19], Jiang and Smith [13] consider multiple bus interconnection network, with overlapped connectivity. Other works concentrate on d-dimensional grids with additional buses, Aggarwal [1] consider global buses, to which all the processors are connected. Bar-Noy and Peleg [2] and Stout [17] consider meshes with buses on rows and columns. Meyer auf der Heide and Pham [14] investigate meshes with a tree of planar buses for the computation of associative operations. Olariu, Schwing and Zhang [15] consider two-dimensional mesh with a reconfigurable bus, and use the parity problem to show that such a system can be as powerful as the CRCW (concurrent read concurrent write) shared memory computer. Bhuyan and Agrawal [6] propose a generalization of the hypercube interconnection network.

Main results of this paper In this paper, we propose broadcast schemes for different BINs, modeled by hypergraphs, motivated by the fact that global communication procedures are critical to the efficiency of parallel algorithms in all areas of applications [12]. We focus on the *hyperpath*, the d-dimensional *hypergrid*, the *hyperring*, and the d-dimensional *hypertorus* architectures, which are the bus-based versions of the well known point-to-point interconnection networks. Considering that underlying point-to-point interconnection networks lost most of their interest with the advent of massively parallel computers, we focus our attention on the use of multiple buses as the *only* means to implement inter-processor communication, showing that global communication procedures are very efficient and easy to implement.

For deriving efficient communication patterns 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 use some known communication algorithms for usual networks.

Our paper is organized as follows. In the next Section we discuss some preliminary material on hypergraphs. In Section 3 we suggest the models of commu-

nication 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 present broadcasting schemes for hyperpath, hyperring, d-hypergrid, and d-hypertorus BINs. We close the paper with some concluding remarks and ways for further research in these very attractive architectures.

2 Preliminaries

2.1 Hypertopologies

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 $e_i \subset V \mid \bigcup_{e_i \subset E} = V$.

Given a hypergraph $\mathcal{G}$, denote by $D(\mathcal{G})$ its diameter, defined as for graphs. Define *transmission* as the sending of a message along a bus from one node to one or more adjacent nodes, where two nodes are said to be adjacent if they share a hyperedge. 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.

Hyperpaths We define hyperpath $\mathcal{P}_{n,m}(V, E)$ of size n and order m by:

$$V = \{0, \ldots, nm - 1\};$$
$$e \in E \Leftrightarrow e = \{v_1, \ldots, v_m, v'_1, \ldots, v'_m\} \text{ where } \begin{cases} v_i = v_j \bmod n \\ v'_i = v'_j \bmod n \\ v_i = (v'_i - 1). \end{cases}$$

Hyperrings We define hyperring $\mathcal{R}_{n,m}(V, E)$ of size n and order m (fig. 1) as:

$$V = \{0, \ldots, nm - 1\};$$
$$e \in E \leftrightarrow e = \{v_1, \ldots, v_m, v'_1, \ldots, v'_m\} \text{ where } \begin{cases} v_i = v_j \bmod n \\ v'_i = v'_j \bmod n \\ v_i = (v'_i - 1) \bmod n. \end{cases}$$

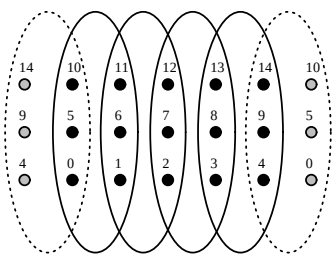

Fig. 1. Hyperring of size 5 and order 3 ($\mathcal{R}_{5,3}$).

Hypergrids Following the definition of grids by means of a product of paths, we can 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) belong to the same edge in $\mathcal{G}$ if and only if
  1. $a = c$ and b, d belong to the same edge in $\mathcal{P}^2_{n_2,m_2}$ *or*
  2. $b = d$ and a, c belong to the same edge in $\mathcal{P}^1_{n_1,m_1}$.

Figure 2 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 on the design of massively parallel systems and algorithms.

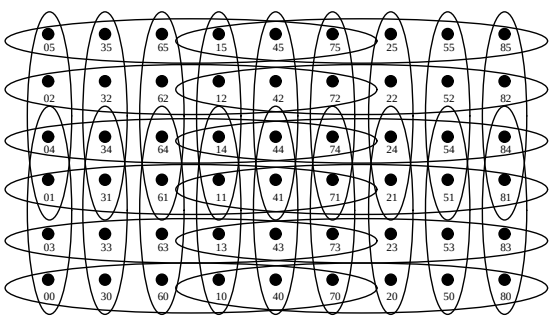

Fig. 2. Hypergrid $\mathcal{G} = \mathcal{P}^1_{3,3} \times \mathcal{P}^2_{3,2}$.

Hypertori Analogously to the generalization of grids to hypergrids, hypertori can be defined as the generalization of tori through the Cartesian product of hyperrings, as follows.

A d-hypertorus $\mathcal{T}^d$ is defined as $\mathcal{R}^1_{n_1,m_1} \times \ldots \times \mathcal{R}^d_{n_d,m_d}$ where $\mathcal{R}^i_{n_i,m_i}$ are hyperrings.

2.2 Comparison with graphs

We compare the defined topologies with the usual graph topologies. In Table 1 $\mathrm{dgr}(v)$ and $\mathrm{dgr}(e)$ are respectively the maximum degree of the nodes and edges in a graph.

Topology	Number of nodes	dgr(v)	dgr(e)	Diameter
$P_n(V,E)$	n	2	2	$n-1$
$R_n(V,E)$	n	2	2	$\lfloor \frac{n}{2} \rfloor$
$G_{n_1,\ldots,n_d}(V,E)$	$n_1 \ldots n_d$	$2d$	2	$\sum_{i=1}^{d}(n_i-1)$
$T_{n_1,\ldots,n_d}(V,E)$	$n_1 \ldots n_d$	$2d$	2	$\sum_{i=1}^{d}(\lfloor \frac{n_i}{2} \rfloor)$
$\mathcal{P}_{n,m}(V,E)$	nm	2	$2m$	$n-1$
$\mathcal{R}_{n,m}(V,E)$	nm	2	$2m$	$\lfloor \frac{n}{2} \rfloor$
$\mathcal{G}^d = \mathcal{P}^1_{n_1,m_1} \times \ldots \times \mathcal{P}^d_{n_d,m_d}$	$\prod_{i=1}^{d}(m_i n_i)$	$2d$	$\max_{i=1}^{d}(2m_i)$	$\sum_{i=1}^{d}(n_i-1)$
$\mathcal{T}^d = \mathcal{R}^1_{n_1,m_1} \times \ldots \times \mathcal{R}^d_{n_d,m_d}$	$\prod_{i=1}^{d}(m_i n_i)$	$2d$	$\max_{i=1}^{d}(2m_i)$	$\sum_{i=1}^{d}(\lfloor \frac{n_i}{2} \rfloor)$

Table 1. Comparative table.

We see that, for example, a topology with N nodes in a ring has diameter $\lfloor \frac{N}{2} \rfloor$. Whereas a topology with N nodes in a hyperring (with hyperedges of size s assuming $N \bmod s = 0$) has diameter s times smaller ($\frac{N}{2s}$).

3 Modeling bus access and capacity

We first recall the usual notation used for modeling communications in point-to-point interconnection networks [11]. 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. The notation E_k is used, where E is the model for edges and k is the model for nodes.

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. A similar idea can be found in [18].

Definition 2. Given a BIN $\mathcal{G}(V,E)$, inter-processor communications can be 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 buses, we call it the `1-port` model; if it can use k (all) of its buses 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 $N \subset e$, 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 broadcast (one-to-all) schemes for BINs, only under the models $M_{i,j}, i = 1$, $j = 1$ or *. Hence, buses can transport at most one information in a single time step, and nodes are able to use one or all of its buses 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, as broadcast, 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. For each $u \in V'$ we will define the set of nodes represented by u, N_u.

Each $y_i \in N_u$ is called a *represented node* with regard to u. Let $N = \bigcup_{u \in V'} N_u$, we define $W = V \backslash (N \bigcup V')$, the set of nodes without representative. The details will be given henceforth.

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 the simplification for the model $M_{i,j}, (i \leq j)$ of hyperpaths, hyperrings, d-dimensional hypergrids and d-dimensional hypertori.

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

We use this simplification in the hypergraphs in such a way that it will be easy to emulate the communication in the hypergraph under the model $M_{i,j}, (i \leq j)$

[4] We remark that "representative graph" does not have here the same meaning as in [3], where it denotes the line graph of a hypergraph.

through the representative graph under the model H_j. 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 $u \in V'$, $N_u = \{y_1, y_2, \ldots, y_{m-1}\}$,

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

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

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

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

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

Representative graph $G(V', E')$ of a d-dimensional 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, with its representative nodes $V'^{(1)} = \{u^{(1)}_1, \ldots, u^{(1)}_{n_1-1}\}$. Let $\mathcal{G}^{(1)} = G_1 \times \mathcal{P}^2_{n_2,m_2} \times \ldots \times \mathcal{P}^d_{n_d,m_d}$. We denote the nodes of $\mathcal{G}^{(1)}$ by *representative nodes of first order*. We choose the representative nodes for the hyperpath $\mathcal{P}^2_{n_2,m_2}$, and we will have the representative graph $G_2(V'^{(2)}, E'^{(2)})$, with its representative nodes $V'^{(2)} = \{u^{(2)}_1, \ldots, u^{(2)}_{n_2-1}\}$. The nodes of $\mathcal{G}^{(2)} = G_1 \times G_2 \times \mathcal{P}^3_{n_3,m_3} \times \ldots \times \mathcal{P}^d_{n_d,m_d}$ will be called *representative nodes of second order*. By doing the same for the remaining paths, we will have $\mathcal{G}^{(d)} = G_1 \times \ldots \times G_d$. Observe that each $G_i, 1 \leq i \leq d$, is a graph, and therefore $\mathcal{G}^{(d)}$ is the representative graph G of $\mathcal{G}$, more specifically a d-dimensional grid. For the sake of uniformity in notation, we define also $\mathcal{G}^{(0)} = \mathcal{G}$.

For each node x of $\mathcal{G}^{(j)}$ we define the set $N^{(j)}_x = \{y^{(j)}_1, \ldots, y^{(j)}_{m_j-1}\}$, that are the nodes in $\mathcal{G}^{(j-1)}$ such that:

$$\begin{cases} (x)_l = (y_i^{(j)})_l, l \neq j, 1 \leq l \leq d \\ (y_i^{(j)} - x)_j \bmod n_j = 0 \\ x \neq y_i^{(j)}, 1 \leq i \leq m_j - 1 \end{cases}$$

where $(\ldots)_j$ means the value in the j-th dimension.

In each hypergraph $\mathcal{G}^{(j)}, 1 \leq j \leq d$, we define the set $W^{(j)}$ of *nodes without representative*, in $\mathcal{G}^{(j-1)}$, that is, $v \in W^{(j)} \Leftrightarrow v \in \mathcal{G}^{(j-1)}, v \notin \mathcal{G}^{(j)}$, and v is not represented by any node in $\mathcal{G}^{(j)}$.

For each j, the set $W^{(j)}$ is formed by the nodes without representative corresponding to the simplification of the $(n_1 - 1) \ldots (n_{j-1} - 1)(n_{j+1} m_{j+1}) \ldots (n_d m_d)$ hyperpaths on this dimension in $\mathcal{G}^{(j-1)}$. Let $W^{(j)}_{i_1,\ldots,i_d}, 0 \leq i_s < n_s - 1, s < j, i_j = 1, 0 \leq i_s < n_s m_s, s > j$, be defined as follows:

$$v \in W^{(j)}_{i_1,\ldots,i_d} \Leftrightarrow \begin{cases} v \in \mathcal{G}^{(j-1)} \\ (v)_s = \begin{cases} u^{(j)}_{i_s+1}, s < j \\ i_s, s > j \end{cases} \\ (v_j \bmod n_j) = n_j - 1. \end{cases}$$

We see that $W^{(j)}$ is the union of the sets $W^{(j)}_{i_1,\ldots,i_d}$. Figure 3 shows an example.

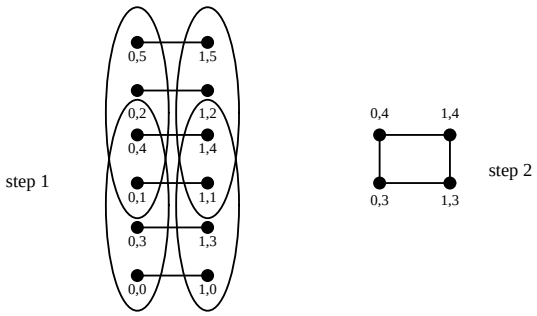

Fig. 3. The 2 steps of simplification for the 2-hypergrid $\mathcal{G} = \mathcal{P}^1_{3,3} \times \mathcal{P}^2_{3,2}$.

In Figure 3 we have steps 1 and 2, that correspond to $\mathcal{G}^{(1)}$ and $\mathcal{G}^{(2)}$. For each node $x^{(1)}$ of $\mathcal{G}^{(1)}$, we have their represented nodes $(N_{x^{(1)}})$ in $\mathcal{G}^{(0)}$, for example, for $x = (0, 5)$ we have $N_{x^{(1)}} = \{(3, 5), (6, 5)\}$. The set of nodes without representative of $\mathcal{G}^{(1)}$ is $W^{(1)} = \bigcup_{i=0}^{5} \{(2, i), (5, i), (8, i)\}$.

It is easy to see that $|W^{(1)}| = m_1(n_2 m_2) = 18$. We see that $W^{(1)} = \bigcup_{i=0}^{5} W_{1,i}$ and $W_{1,i} = \{(2, i), (5, i), (8, i)\}, 0 \leq i \leq 5$.

For each node x of $\mathcal{G}^{(2)}$, we have their represented nodes $(N_{x^{(2)}})$ in $\mathcal{G}^{(1)}$, for example, for $x = (0, 4)$ we have $N_{x^{(2)}} = \{(0, 1)\}$. The set of nodes without representative of $\mathcal{G}^{(2)}$ is $W^{(2)} = \{(0, 2), (0, 5), (1, 2), (1, 5)\}$, $|W^{(2)}| = n_1 m_2 = 4$. We see that $W^{(2)} = \bigcup_{i=0}^{1} W_{i,1}$ and that $W_{0,1} = \{(0, 2), (0, 5)\}$ and $W_{1,1} = \{(1, 2), (1, 5)\}$.

Representative graph of a d-dimensional hypertorus $\mathcal{T}^d = \mathcal{R}^1_{n_1,m_1} \times \ldots \times \mathcal{R}^d_{n_d,m_d}$. We perform simplification by dimensions. We first choose the representative nodes for the hyperring $\mathcal{R}^1_{n_1,m_1}$. Let $G_1(V'^{(1)}, E'^{(1)})$ be its representative graph, with its representative nodes $V'^{(1)} = \{u_1^{(1)}, \ldots, u_{n_1}^{(1)}\}$. Let $\mathcal{T}^{(1)} = G_1 \times \mathcal{R}^2_{n_2,m_2} \times \ldots \times \mathcal{R}^d_{n_d,m_d}$. We denote the nodes of $\mathcal{T}^{(1)}$ by *representative nodes of first order*. We choose the representative nodes for the hyperring $\mathcal{R}^2_{n_2,m_2}$, and we will have the representative graph $G_2(V'^{(2)}, E'^{(2)})$, with its representative nodes $V'^{(2)} = \{u_1^{(2)}, \ldots, u_{n_2}^{(2)}\}$. The nodes of $\mathcal{T}^{(2)} = G_1 \times G_2 \times \mathcal{R}^3_{n_3,m_3} \times \ldots \times \mathcal{R}^d_{n_d,m_d}$ will be called *representative nodes of second order*. By doing the same for the remaining rings, we will have $\mathcal{T}^{(d)} = G_1 \times \ldots \times G_d$. Observe that each $G_i, 1 \leq i \leq d$, is a graph, and therefore $\mathcal{T}^{(d)}$ is the representative graph G of $\mathcal{T}$, more specifically a d-dimensional torus. We define also $\mathcal{T}^{(0)} = \mathcal{T}$.

For each node x of $\mathcal{T}^{(j)}$ we define the set $N_x^{(j)} = \{y_1^{(j)}, \ldots, y_{m_j-1}^{(j)}\}$, that are the nodes in $\mathcal{T}^{(j-1)}$ such that:

$$\begin{cases} (x)_l = (y_i^{(j)})_l, l \neq j, 1 \leq l \leq d \\ (y_i^{(j)} - x)_j \bmod n_j = 0 \\ x \neq y_i^{(j)}, 1 \leq i \leq m_j - 1 \end{cases}$$

Notice that a simplification of the given hypergraphs resulted in a representative graph, of the same family from which the hypergraph was generalized.

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.

Claim 1 *Given a hypergraph $\mathcal{G}$ and its representative graph G under the model $M_{i,j}, i \leq j$. If $i = 1$ then $(G, H_j) \prec (\mathcal{G}, M_{1,j})$. If $i = 2$ then $(G, F_j) \prec (\mathcal{G}, M_{2,j})$.*

5 Communication patterns for broadcast

Given a hypergraph $\mathcal{G}(V', E')$ and its representative graph $G(V, E)$, we denote the broadcast operation, under model M, in graph G by $broadcast_M(G)$. 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 $\mathcal{G}$.

Henceforth, when we define a representative graph, we also implicitly define the sets of representative nodes, represented nodes, and nodes without representative. We denote by $D(G)$ the diameter of G.

Recall that for broadcast it suffices to analyze models $M_{1,i}$ in any hypergraph, for the following reason. There exists, for each step of broadcast, a bipartition of nodes, those which contain the information and those that do not, and the information should be transmitted from one partition to another. Under the model of transmission assumed (transmission time independent of the message size), the capacity of the buses that connect the same two nodes between partitions is meaningless. It follows that the cases in which the buses are k-capacity are included in the ones where buses are single-capacity.

We assume, without loss of generality, that if the node that initially contains the message is candidate to become a representative node, it is selected as one such node during the simplification phase. Furthermore, for all topologies we consider the worst case, that is, we suppose that the source is the node for which the broadcasting time is maximal.

Algorithms for the gossiping problem in hyperpath and hypergrid can be found in [9].

Hyperpath $\mathcal{P}_{n,m}$.

Consider the weakest model $M_{1,1}$, and let $G(V', E')$ be the representative graph, $V' = \{u_1, \ldots, u_{n-1}\}$ the set of representative nodes, $N_u, u \in V'$, the set of represented nodes and W the set of nodes without representative. We have the following theorem.

Theorem 3. *The complexity on the number of steps for broadcasting in a hyperpath* $\mathcal{P}_{n,m}$ *is* $D(\mathcal{P}_{n,m})$*, which is optimal.*

Algorithm:

1. $broadcast_{H_1}(G)$
2. $W \Leftarrow u_{n-1}$

In a number of steps less or equal to $D(G) + 1 = D(\mathcal{P}_{n,m})$ (which is the trivial lower bound) the broadcast will complete.

Hyperring $\mathcal{R}_{n,m}$.

Consider the model $M_{1,x}$, $x = 1, 2$, and let $G(V', E')$ be the representative graph, $V' = \{u_1, \ldots, u_n\}$ the set of representative nodes, $N_u, u \in V'$, the set of represented nodes. The description of the algorithm could not be simpler:
Algorithm:

1. $broadcast_{H_x}(G)$

Theorem 4. *The complexity of the broadcast operation for the hyperring* $\mathcal{R}_{n,m}$*, in models* $M_{1,1}$ *and* $M_{1,2}$*, is optimal.*

Proof: For the model $M_{1,2}$ this is obvious, since the source can send its message over its two buses simultaneously. The time complexity of the algorithm under model $M_{1,1}$ depends on the parity of n. If n is even, then the algorithm completes broadcast in $D(\mathcal{R}_{n,m})$ cycles, implying its optimality. Otherwise, the algorithm takes $D(\mathcal{R}_{n,m}) + 1$ cycles. Notice, however, that n being odd implies that there are two nodes in different hyperedges of $\mathcal{R}_{n,m}$ at distance D (diameter) from the source. Then $D + 1$ is a lower bound in this case.

Hypergrid $\mathcal{G} = \mathcal{P}_{n_1,m_1} \times \ldots \times \mathcal{P}_{n_d,m_d}$.

There will be d phases under the weakest model $M_{1,1}$. First we disseminate the message on the path $\mathcal{P}_{n_1,m_1}$ according to the algorithm already seen, in at most $n_1 - 1$ cycles. Next, all the paths of the second dimension perform a broadcast. In this way, until dimension d, we complete the broadcast in at most $n_1 + \ldots + n_d - d = D(\mathcal{G})$ cycles.

Theorem 5. *The complexity on the number of steps for broadcasting in a hypergrid* $\mathcal{G} = \mathcal{P}_{n_1,m_1} \times \ldots \times \mathcal{P}_{n_d,m_d}$ *is* $D(\mathcal{G})$*, which is optimal.*

Hypertorus $\mathcal{T} = \mathcal{R}_{n_1,m_1} \times \ldots \times \mathcal{R}_{n_d,m_d}$.

As previously, this problem is reduced to broadcast in the representative graph of $\mathcal{T}$, which is in its turn a d-torus. Then, the algorithms from [7, 11] apply directly, as well as their analyses:
Algorithm:

1. $broadcast_{H_x}(G)$

Theorem 6. *The complexity on the number of steps for broadcasting in a* d*-dimensional hypertorus is* $D + \max\{0, k-1\}$*, where* k *is the number of odd dimensions in* $\mathcal{T} = \mathcal{R}_{n_1,m_1} \times \ldots \times \mathcal{R}_{n_d,m_d}$.

The above results imply that for the broadcast operation, only the diameter of the BIN has influence on the communication time. That is, the time to complete a broadcast operation independs on the number of nodes on each bus of the BIN.

6 Conclusion

The need for massively parallel systems in order to deal with real size applications is exposing the weakness of point-to-point interconnection networks, due mostly to VLSI implementation related issues. Bus interconnection networks could be an efficient and reliable way to interconnect thousands of processors, while ensuring that the distances among processors are kept small.

We presented communication patterns for broadcasting on BIN's topologies. Although they are the by-product of this new powerful tool called simplification, the schemes are simple and very efficient, proving that global communication procedures are not a drawback for the utilization of BINs, as corroborated by some investigations we carried out on gossiping in BINs [8].

Below we summarize our results for broadcast on some BINs.

Broadcast	$M_{1,1}$	$M_{1,2}$	diameter
$\mathcal{P}_{n,m}(V,E)$	$n-1$	$n-1$	$n-1$
$\mathcal{R}_{n,m}(V,E)$	$\lceil \frac{n}{2} \rceil$	$\lfloor \frac{n}{2} \rfloor$	$\lfloor \frac{n}{2} \rfloor$
$\mathcal{G}^d(V,E) = \mathcal{P}^1_{n_1,m_1} \times \ldots \times \mathcal{P}^d_{n_d,m_d}$	$\sum_{i=1}^{d}(n_i-1)$	$\sum_{i=1}^{d}(n_i-1)$	$\sum_{i=1}^{d}(n_i-1)$
$\mathcal{T}^d(V,E) = \mathcal{R}^1_{n_1,m_1} \times \ldots \times \mathcal{R}^d_{n_d,m_d}$	$\sum_{i=1}^{d}(\lceil \frac{n_i}{2} \rceil)$	$\sum_{i=1}^{d}(\lfloor \frac{n_i}{2} \rfloor)$	$\sum_{i=1}^{d}(\lfloor \frac{n_i}{2} \rfloor)$

Regarding further research, BINs represent a very interesting way of thinking massive parallelism. Besides designing new algorithms and refining the bus access and capacity models, issues like whether or not the total lack of point-to-point communication will be a major drawback for efficiently using such systems, should be studied. Moreover, as noted in [18], communication algorithms in such a formal model can find applications in the area of emerging optical technologies.

References

1. A. Aggarwal, Optimal bounds for finding maximum on array of processors with k global buses, *IEEE Transactions on Computers* **35** (1986) 62–64.
2. A. Bar-Noy and D. Peleg, Square meshes are not always optimal, *ACM Symposium on Parallel Algorithms and Architectures*, (ACM, 1989) 138–147.
3. C. Berge, *Hypergraphs*, (North Holland, 1989).
4. J-C. Bermond, Le problème des "ouvroirs" (Hypergraph gossip problem), in *Colloques Internacionaux du Centre National de la Recherche Scientifique*, No. 260, Orsay, 1976, 31–34.
5. J-C. Bermond and F.O. Ergincan, Bus interconnection networks, *Technical Report No. 93-56*, Université de Nice - Sophia Antipolis, September 1993.
6. L.N. Bhuyan and D.P. Agrawal, Generalized hypercube and hyperbus structures for a computer network, *IEEE Transactions on Computers* **33** (1984) 323–333.
7. A.M. Farley and S.T. Hedetniemi, *Broadcasting in grid graphs*, In Congressus Numerantium XXI, editor, Proc. 9th S-E conf. combinatorics, graph theory, and computing, pages 275-288, 1978.
8. A. Ferreira, A. Goldman vel Lejbman, and S. W. Song. Towards simple massively parallel systems: Bus based parallel computers and new communication patterns. Research Report 93-18, Laboratoire de l'Informatique du Parallélisme, Ecole Normale Supérieure de Lyon, France, June 1993.
9. A. Ferreira, A. Goldman vel Lejbman, S.W. Song, A viable way for massive parallelism, *PARLE'94 Parallel Architectures and Languages Europe* (Lecture Notes of Computer Science, Springer-Verlag), Athens, July 1994. (to appear)
10. C.M. Fiduccia, Bused hypercubes and other pin-optimal networks, *IEEE Transactions on Parallel and Distributed Systems* **3(1)** (1992) 14–24.
11. P. Fraigniaud and E. Lazard, Methods and problems of communication in usual networks, Rapport de Recherche LIP 91-33, Laboratoire de l'Informatique du Parallélisme, École Normale Supérieure de Lyon, October, 1991, to appear in Discrete Applied Mathematics.
12. K. Hwang, *Advanced Computer Architecture: parallelism, scalability, programmability*, (McGraw-Hill, 1993).
13. 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.
14. F. Meyer auf der Heide and H.T. Pham, On the performance of networks with multiple buses, *STACS'92 9th Annual Symposium on Theoretical Aspects of Computer Science, Lecture Notes in Computer Science*, (Springer-Verlag, 1992) 97–108.
15. 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.
16. 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.
17. Q.F. Stout, Meshes with multiple buses, *Proc. 27th IEEE Symposium on the Foundations of Computer Science* (1986) 264–273.
18. 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.
19. 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