Gossiping in bus interconnection networks

A. Ferreira*

CNRS - Laboratoire de l'Informatique du Parallélisme

École Normale Supérieure de Lyon

46, Allée d'Italie

69364 Lyon Cédex 07 - France

e-mail: ferreira@lip.ens-lyon.fr Tel. +33 72 72 80 00 - Fax + 33 72 72 80 80

A. Goldman vel Lejbman and S. W. Song†

Department of Computer Science

Institute of Mathematics and Statistics

University of São Paulo

Rua do Matão, 1010 - São Paulo, SP 05508-900 - Brazil

e-mail: gold@ime.usp.br, song@ime.usp.br Tel. +55 11 818 6141 - Fax +55 11 814 4135

Abstract

Several architectures have been proposed to enhance point-to-point architectures with the addition of multiple bus systems. In particular, we consider an architecture for a massively parallel system where processors are connected 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. In this paper we model the so-called *bus interconnection networks* (BINs) by a hypergraph. We consider the gossip operation in the various proposed architectures. We focus on the hyperpath, the d-dimensional hypergrid, the hyperring, and the d-dimensional hypertorus architectures and we give corresponding algorithms for the gossiping operation. Some lower bounds are also derived.

Key-words: Massively parallel architectures, multiple bus systems, global communication, communication models, hypergraphs and applications, gossiping, total exchange.

*Partially supported by the PRC C^3 and ANM of the French CNRS.

†Supported by FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo) - Proc. No. 93/0603-1 and 94/4544-2, CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) - Proc. No. 523112/94-7 and PROTEM II, and Commission of the European Communities through Project ITDC-207.

1 Introduction

Recently several architectures have been proposed to enhance point-to-point architectures with the addition of multiple bus systems [12, 14, 18, 21]. The goal is to improve the communication efficiency, since interprocessor communication frequently constitutes serious bottlenecks. In particular, in [10, 11], we have proposed an architecture for a massively parallel system where processors are connected *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. 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. For instance, 8100 processors can compose a 2-dimensional hypergrid, as defined in a later section, where each processor is connected to at most four buses, each bus has 30 processors, and the distance between the farthest apart processors is as low as 10. Notice that this is better than a usual hypercube of the same size, and easier to build, since it requires only 900 buses.

Usually networks are modeled by a graph where processors are nodes and communication links are edges. In this paper we model the so-called *bus interconnection networks* (BINs, for short) 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 [5, 4, 20]. We consider the gossip operation in the various proposed architectures. Gossiping is a very important communication operation that exchanges information among processors [15]. The broadcast operation has been treated in separate paper [11]. We focus on the *hyperpath*, the d-dimensional *hypergrid*, the *hyperring*, and the d-dimensional *hypertorus* architectures.

We use the new tool called *simplification* [10, 11]. 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.

The following results are obtained.

We generalize the linear network and design a BIN with the following features. Let $\mathcal{P}_{n,m}$ be a parallel computer having nm processors, $n-1$ buses, with $2m$ processors per bus, each processor being connected to at most two buses. We show that the gossiping operation can be done in at most $2m+n$ steps.

We generalize the ring network, to be called $\mathcal{R}_{n,m}$, having nm processors, n buses, with $2m$ processors per bus, each processor being connected to exactly two buses. We show how to do the gossiping in at most $m+n/2+2$ steps.

We start from a 2-d grid, the 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 buses, each bus with $2m_1$ or $2m_2$ processors, where each processor is connected to at most four buses. We show that, in $\mathcal{G}_2$, we can perform the gossiping operation in $m_1+m_2+n_1+n_2$ steps.

Finally, if we start from a 2-d torus, we will generate a $\mathcal{T}_2 = \mathcal{R}^1_{n_1,m_1} \times \mathcal{R}^2_{n_2,m_2}$. This is a parallel computer with $n_1m_1n_2m_2$ processors, n_1n_2 buses, each bus with $2m_1$ or $2m_2$ processors, where each processor is connected to four buses. In $\mathcal{T}_2$, we can perform broadcasting in no more than $\lceil n_1/2 \rceil + \lceil n_2/2 \rceil$ steps, if n_1 or n_2 is even. We show how to do the gossiping in

less than $m_1 + m_2 + n_1/2 + n_2/2 + 2$ steps.

All these schemes are optimal since they asymptotically match the lower bound given by the diameter of the network.

Our paper is organized as follows. In Section 1.1 we discuss some preliminary material on hypergraphs. In Section 2 we discuss the models of communication best suited for BINs. Section 3 we describe the main tool, called simplification, to deal with global communication in such massively parallel systems. Then, in Section 4, we present gossiping schemes for hyperpath, hyperring, d-hypergrid, and d-hypertorus BINs. We close the paper with some concluding remarks.

1.1 Previous works

Modeling of bus networks by hypergraphs has been considered in [5, 4, 20]. In [18], Scherson proposes the concept of *orthogonal graphs*, that can describe binary m-cubes or meshes with buses in rows and columns. Fiduccia [12] 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 [20] 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 present an excellent survey on structural problems related to BINs [6].

Some researchers use the approach to enhance existing point-to-point networks with multiple bus systems. Wilkinson [21], Jiang and Smith [14] consider multiple bus interconnection network, with overlapped connectivity. Many other works concentrate on d-dimensional grids with additional buses. Bokhari [8] and Aggarwal [1] consider global buses, to which all the processors are connected. Bar-Noy and Peleg [2] and Stout [19] consider meshes with buses on rows and columns. Meyer auf der Heide and Pham [16] investigate meshes with a tree of planar buses for the computation of associative operations. Olariu, Schwing and Zhang [17] consider two-dimensional mesh with a reconfigurable bus. Bhuyan and Agrawal [7] propose a generalization of the hypercube interconnection network.

1.2 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 $\mathcal{H}(V, E)$, 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 $\mathcal{H}$, denote by $D(\mathcal{H})$ 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.

The next definitions of hypertopologies are the same as in [11].

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

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

Two examples are shown in Figure 1.

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

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

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. Figure 1 shows an example.

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 Modeling bus access and capacity

Recall the usual notation used for modeling communications in point-to-point interconnection networks [13]. 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 (*e.g.*, F_k stands for full-duplex, k-port, and H_1 stands for half-duplex, 1-port).

Definition 1 *Given a BIN $\mathcal{H}(V, E)$, 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.*

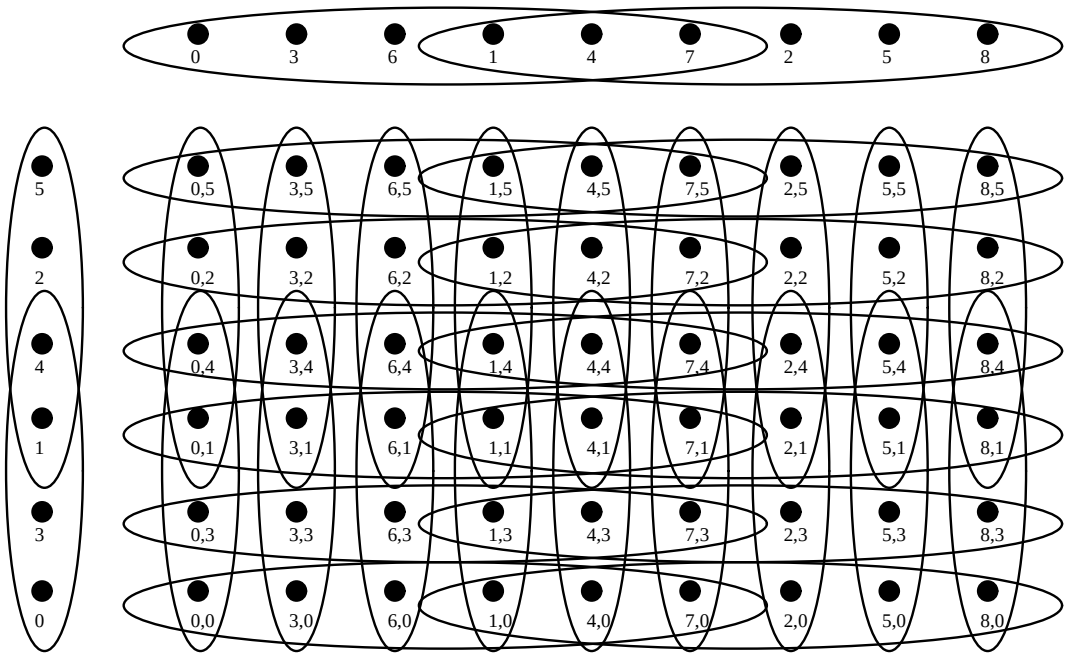

Figure 1: Hyperpath $\mathcal{P}^1_{3,3}$ (horizontal), $\mathcal{P}^2_{3,2}$ (vertical). Hypergrid $\mathcal{G} = \mathcal{P}^1_{3,3} \times \mathcal{P}^2_{3,2}$.

2.1 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 $\mathcal{H}(V, E)$, $v \in V$, $e \in E$, 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.

2.2 Motivation

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

In addition to these models, we will also consider the models $*M_{i,j}$, to be defined later. In such models we try to make a better use of the characteristics of the hypergraph, where there are many nodes in each edge that can communicate simultaneously with the adjacent edges.

We will consider four models, namely, $M_{i,j} (i \leq j)$, $M_{2,1}$, $*M_{1,1}$, and $*M_{2,1}$. In particular,

in the first one, $M_{i,j}(i \leq j)$, we give algorithms for only the cases $i = 1, 2$. The limit of 2 in the bus capacity seems to be feasible and is motivated by the analogous bi-directional transmission on full-duplex channels. On the other hand, for the bus access, we study only the cases `1-port` and `*-port`, as not to make the paper to long.

3 Simplification

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

The idea is to use a 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[1].

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. Let $N = \bigcup_{u \in V'} N_u$, we use $W = V \backslash (N \bigcup V')$ to denote the set of nodes without representative.

A transmission from a node to another in the representative graph also means transmission to the nodes that they represent 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.

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 through the representative graph under the model F, or H.

We show the four main forms for simplification of hyperpaths, hyperrings, d-dimensional hypergrids and d-dimensional hypertori. In the following table we can see the number of representative nodes to be chosen in each form of simplification, and the emulation model of the representative graph. Details will be shown later.

Model	# representative nodes per intersection of hyperedges	model of the representative graph
$M_{i,j}(i \leq j)$	1	H_j if $i = 1$, F_j if $i = 2$
$M_{2,1}$	2	F_1
$*M_{1,1}$	$2\times$ (number of dimensions)	H_*
$*M_{2,1}$	$2\times$ (number of dimensions)	F_*

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

- **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:

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

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

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 (nodes of degree 1) of $\mathcal{P}_{n,m}$.

The sets of represented nodes are such that, for each $u \in V', N_u = \{y_1, y_2, \ldots, y_{m-1}\}$,

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

An example can be seen in Figure 2.

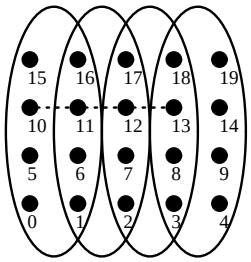

Figure 2: The representative graph for the hyperpath $\mathcal{P}_{5,4}$ has the nodes $\{10, 11, 12, 13\}$, the edges are in dotted lines. We have also: $N_{10} = \{0, 5, 15\}$, $N_{11} = \{1, 6, 16\}$, $N_{12} = \{2, 7, 17\}$, $N_{13} = \{3, 8, 18\}$ and $W = \{4, 9, 14, 19\}$.

- **Representative graph $G(V', E')$ of a 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 the 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 $V'^{(2)} = \{u_1^{(2)}, \ldots, u_{n_2-1}^{(2)}\}$. 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 the 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 $N_x^{(j)} = \{y_1^{(j)}, \ldots, y_{m_j-1}^{(j)}\}$ of the nodes in $\mathcal{G}^{(j-1)}$ such that:

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

where $(\ldots)_j$ means the value in the j-th dimension. The vertices in $N_x^{(j)}$ are equal to x in all coordinates, except coordinate j, and in coordinate j those vertices are equal to x $(\mathrm{mod} n_j)$.

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 all 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

$$W^{(j)}_{i_1,\ldots,i_d} \begin{cases} 0 \leq i_s < n_s - 1, s < j \\ i_j = 1 \\ 0 \leq i_s < n_s m_s, s > j, \end{cases}$$

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 $W^{(j)}_{i_1,\ldots,i_d}$.

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 its represented nodes $(N_{x^{(1)}})$ in $\mathcal{G}^{(0)}$ (Figure 1). For example, for $x^{(1)} = (0, 5)$ 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 $|W^{(1)}| = m_1(n_2 m_2) = 18$.

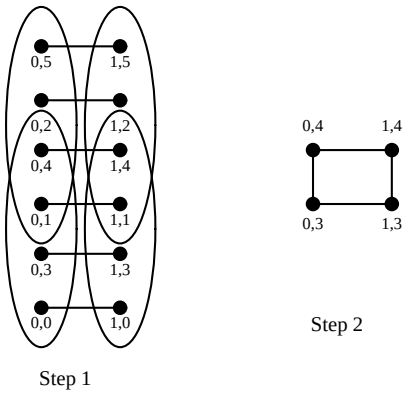

Figure 3: The 2 steps of simplification for the 2-hypergrid $\mathcal{G} = \mathcal{P}^1_{3,3} \times \mathcal{P}^2_{3,2}$.

For each node x of $\mathcal{G}^{(2)}$, we have their represented nodes ($N_{x^{(2)}}$) in $\mathcal{G}^{(1)}$, for example, for $x^{(2)} = (0,4)$ we have $N_{x^{(2)}} = \{(0,1)\}$. The set of nodes without representative of $\mathcal{G}^{(2)}$ is $W^{(2)} = \{(0,2),(0,5),(1,2),(1,5)\}$, $|W^{(2)}| = (n_1 - 1)m_2 = 4$. We see 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 $V'^{(1)} = \{u^{(1)}_1, \ldots, u^{(1)}_{n_1}\}$. 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 the 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 $V'^{(2)} = \{u^{(2)}_1, \ldots, u^{(2)}_{n_2}\}$. 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 the second order*. By doing the same for the remaining hyperrings, 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. For the sake of uniformity in notation, we define also $\mathcal{T}^{(0)} = \mathcal{T}$.

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

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

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

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, excluding the trivial communication of a node to itself, 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 $\mathcal{H}$ and its representative graph G under the model $M_{i,j}, i \leq j$. If $i = 1$ then $(G, H_j) \prec (\mathcal{H}, M_{1,j})$. If $i = 2$ then $(G, F_j) \prec (\mathcal{H}, M_{2,j})$.*

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

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 need two nodes in each intersection, in the model $M_{2,1}$, and two times the number of dimensions in the models $*M_{1,1}$ and $*M_{2,1}$.

- **Representative graph $G(V', E')$ of a hyperpath $\mathcal{P}_{n,m}(V, E)$.**

  We elect $2(n-1)$ representative nodes, $V' = \{u_1, \ldots, u_{n-1}, u_n, \ldots, u_{2(n-1)}\}$, such that:

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

  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 $(y_i - z) \bmod n = 0, 1 \leq i \leq m$. Figure 4 shows an example.

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

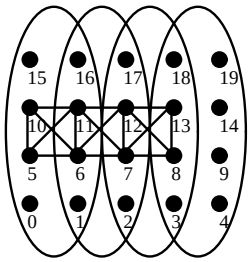

Figure 4: Representative graph for the model $M_{2,1}$ of the hyperpath $\mathcal{P}_{5,4}$.

- **Representative graph $G(V', E')$ of a hyperring $\mathcal{R}_{n,m}(V, E)$.** We elect $2n$ representative nodes, $V' = \{u_1, \ldots, u_n, u_{n+1}, \ldots, u_{2n}\}$, such that:

  $$\begin{cases} (u_1 \bmod n) = (u_n \bmod n) = 0 \\ u_{i+1} = u_i + 1 \text{ where } 1 \leq i \leq n-1 \text{ or } n < i \leq 2n-1 \\ u_1 \neq u_n. \end{cases}$$

  We define analogously the sets $N_u, u \in V'$, of the represented nodes, in the same manner as for the hyperpath.

- **Representative graph $G(V', E')$ of a d-dimensional hypergrid** $\mathcal{G} = \mathcal{P}^1_{n_1,m_1} \times \ldots \times \mathcal{P}^d_{n_d,m_d}$ is constructed 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.
- **Representative graph of a d-dimensional hypertorus** $\mathcal{T}^d = \mathcal{R}^1_{n_1,m_1} \times \ldots \times \mathcal{R}^d_{n_d,m_d}$. Analogous to the previous one.

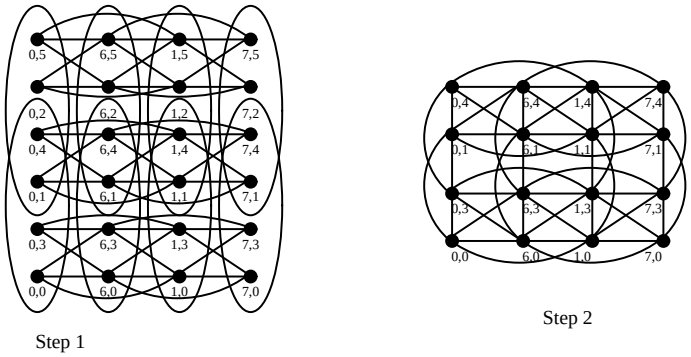

Figure 5: Simplification for the model $M_{2,1}$ of the 2-hypergrid $\mathcal{G} = \mathcal{P}^1_{3,3} \times \mathcal{P}^2_{3,2}$.

3.3 Reduction

Observe that in the case where we generate the representative graph $G(V', E')$ of $\mathcal{H}(V, E)$ 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 $\mathcal{H}(V, E)$.

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 each dimension of $G(V', E')$ corresponds to a single node in the same dimension of G^*. Two nodes are adjacent in the latter only if they were adjacent in the former.

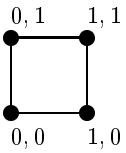

Figure 6: Reduced graph G^* of the hypergrid $\mathcal{G} = \mathcal{P}^1_{3,3} \times \mathcal{P}^2_{3,2}$.

Proposition 1 *Consider a hypergraph H and its corresponding representative graph G and reduced graph G^*. If every pair of equivalent nodes have the same information, then $(G^*, F_1) \prec (G, H_1)$.*

Proof idea:

(Consider $\mathcal{H}$ as the hypergraph $\mathcal{G} = \mathcal{P}^1_{3,3} \times \mathcal{P}^2_{3,2}$.) A simultaneous transmission $(0,1) \leftarrow (1,1)$, $(1,1) \leftarrow (0,1)$ in the reduced graph of Figure 6 is realized in graph $\mathcal{G}^{(d)}$ (Figure 5) through the following transmissions: $(0,1) \leftarrow (1,1), (7,1) \leftarrow (6,1), (0,4) \leftarrow (1,4), (7,4) \leftarrow (6,4)$, for such transmissions in $\mathcal{G}^{(d)}$ correspond to the following transmissions, in $\mathcal{G}$.

$\mathcal{G}^{(d)}$	$\mathcal{G}$
$(0,1) \leftarrow (1,1)$	$\{(0,1),(6,1)\} \Leftarrow (1,1)$
$(7,1) \leftarrow (6,1)$	$\{(1,1),(7,1)\} \Leftarrow (6,1)$
$(0,4) \leftarrow (1,4)$	$\{(0,4),(6,4)\} \Leftarrow (1,4)$
$(7,4) \leftarrow (6,4)$	$\{(1,4),(7,4)\} \Leftarrow (6,4)$

Proof: Suppose that initially the nodes $\{v_1, \ldots, v_{2d}\} (\subset \mathcal{H})$ that correspond to each node v of G^* contain the same information as v.

Given a cycle T in G^*, in model F_1, let $t_1, \ldots, t_p$ be the transmissions of T. Since we are in model F_1 the edges envolved in each transmission are t_i, pairwise disjoint, $(t_i \bigcap t_j = \emptyset, i \neq j)$. We will show that a transmission, between two adjacent nodes u, v, with $u \leftarrow v$ and $v \leftarrow u$, can be realized in $\mathcal{H}$.

Let (u_i, u_i') (resp. (v_i, v_i')), $1 \leq i \leq 2(d-1)$, be the pairs of equivalent nodes that gave rise to u (resp. v), on the dimension where the transmission occurs, such that u_i and v_i have a common edge in $\mathcal{H}$.

A simultaneous transmission $u \leftarrow v, v \leftarrow u$, is equivalent to transmission $u_i \leftarrow v_i, v_i' \leftarrow u_i', 1 \leq i \leq 2(d-1)$, since, as v_i, v_i', u_i, u_i' are on the same hyperedge, $u_i \leftarrow v_i$ implies $\{u_i, u_i'\} \Leftarrow v_i$ and $v_i' \leftarrow u_i'$ implies $\{v_i, v_i'\} \Leftarrow u_i'$. Observe that after the transmission the initial assumption continues valid, that is, every pair of equivalent nodes have the same information. □

Proposition 2 $(G, H_1) \prec (\mathcal{H}, M_{2,1})$.

Proof: Given two adjacent nodes u, v in $\mathcal{H}$ it is easy to verify that the isolate transmission $u \leftarrow v$ can be realized in $\mathcal{H}$. Suppose that there exists a cycle T realizable in G in model H_1, that cannot be executed in $\mathcal{H}$ in model $M_{2,1}$. We can decompose T into $t_1, \ldots, t_d$, where t_i contains the transmissions on dimension i.

It follows that there exists a $j, 1 \leq j \leq d$, such that t_j cannot be realized in $\mathcal{H}$ in model $M_{2,1}$. Observing that on dimension j, we have two nodes of G per hyperedge of $\mathcal{H}$, and that each edge can carry out two transmissions, we come to a contradiction. □

3.4 *-Simplification for model $M_{1,1}$

With this *-simplification operation, as well as the next one, we make better use of the characteristics of hypergraphs. For this purpose, we choose more than one representative per intersection.

- **Representative graph $G(V', E')$ of hyperpath $\mathcal{P}_{n,m}(V, E)$.**

  The representative graph $G(V', E')$ is the same as in the simplification for the model $M_{2,1}$. Let G^* be its reduced graph.

Proposition 3 $(G^*, H_*) \prec (\mathcal{P}_{n,m}, M_{1,1})$.

Proof: As the limitation of the model is relative to the sending of messages, we will analyze the case in which a node v of G^* sends messages to its two neighbors u, w. Let v_1 and v_2 be the equivalent nodes corresponding to v in G, and N_u and N_w the sets of the represented nodes, respectively, by nodes corresponding to u and w. It is easy to see that the cycle $\{u \leftarrow v, w \leftarrow v\}$ is equivalent to $\{N_u \Leftarrow v_1, N_w \Leftarrow v_2\}$.

Notice that for this proof we assumed that every node of v of graph G^* contains the same information as the nodes corresponding to v, in G. □

- **Representative graph $G(V', E')$ of a hyperring $\mathcal{R}_{n,m}(V, E)$.** Analogous to the previous one.
- **Representative graph of a d-dimensional hypergrid** $\mathcal{G}(V, E) = \mathcal{P}^1_{n_1,m_1} \times \ldots \times \mathcal{P}^d_{n_d,m_d}$. Let $G_1(V'_1, E'_1)$ be the representative graph of $\mathcal{P}^1_{n_1,m_1}$. As in the following, let the representative nodes be $V'_1 = \{u^{(1)}_1, \ldots, u^{(1)}_{(n_1-1)2d}\}$, such that:

$$\begin{cases} u_{1+j(n_1-1)} \bmod n_1 = 0 \\ u_{i+1+j(n_1-1)} = u_{i+j(n_1-1)} + 1 \qquad 1 \leq i \leq n-1 \text{ and } 0 \leq j \leq 2d-1 \\ u_{1+l(n_1-1)} \neq u_{1+k(n_1-1)}, \text{ if } k \neq l. \end{cases}$$

where the nodes of G_1 are adjacent, if and only if they are adjacent in $\mathcal{P}^1_{n_1,m_1}$. In the same way we define the representative graphs $G_i(V'_i, E'_i)$ corresponding to the hyperpaths $\mathcal{P}^i_{n_i,m_i}$, $2 \leq i \leq d$.

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 $\mathcal{G}^{(1)}$ as representative nodes of the first order. 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 denoted as representative nodes of the second order. Continuing like this, we will have $\mathcal{G}^{(d)} = G_1 \times \ldots \times G_d$, that will be a graph, whose nodes will be the representative nodes of the d-th order (Figure 7).

For each $x \in \mathcal{G}^{(j)}$ we define the set $N_x = \{y^{(j)}_1, \ldots, y^{(j)}_{m_j}\}$ belonging to $\mathcal{G}^{(j-1)}$ such that:

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

We also define the sets $W^{(j)}_{i_1,\ldots,i_d}$, $1 \leq j \leq d$, of the nodes without representative analogously to the definition for d-hypergrid in model $M_{1,1}$. From the graph $\mathcal{G}^{(d)}$ we define the reduced graph of G^* (Figure 8).

Following the same reasoning used for the hyperpath, we have:

Claim 2 *If every set of equivalent vertices have the same information, then* $(G^*, H_*) \prec (\mathcal{G}^{(d)}, H_1)$.

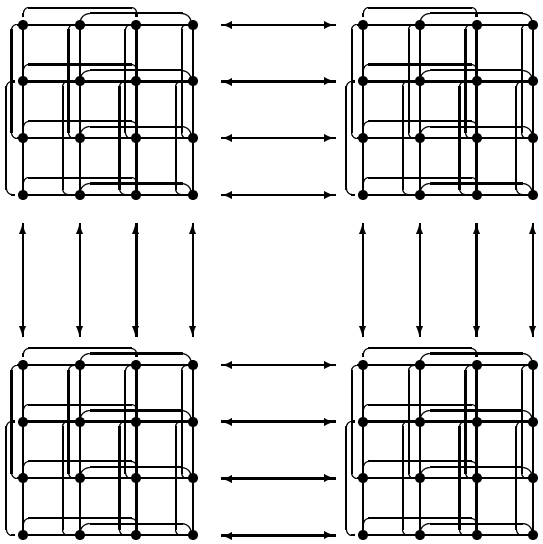

Figure 7: representative graph (under *-simplification) of a two-dimensional hypergrid $\mathcal{G}^2 = \mathcal{P}_{3,m_1} \times \mathcal{P}_{3,m_2}$, $(m_1, m_2 \geq 4)$. The arrows indicate that all the nodes of a row or column are adjacent.

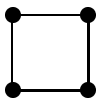

Figure 8: Reduced graph of a two-dimensional hypergrid $\mathcal{G}^2 = \mathcal{P}_{3,m_1} \times \mathcal{P}_{3,m_2}$, $(m_1, m_2 \geq 4)$.

Idea: A transmission from a node $v \in G^*$ to all its neighbors (in a two-dimensional grid) is realized according to Figure 9 by the nodes that gave rise to v in $\mathcal{G}^{(d)}$ (The arrows indicate the direction of transmission). □

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

- The procedures for the hypertorus are analogous.

3.5 *-Simplification for model $M_{2,1}$

Following the same steps to derive the representative graph G and reduced graph G^* of a hypergraph $\mathcal{H}$ in the *-simplification for model $M_{1,1}$, we have:

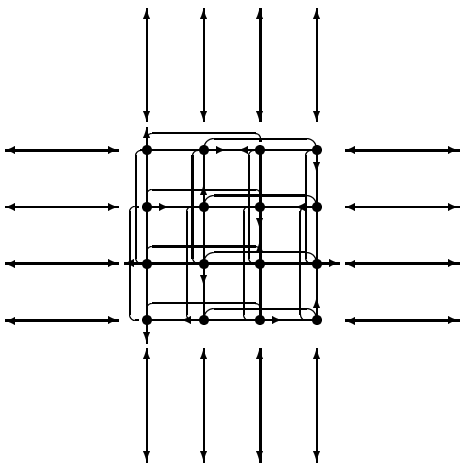

Figure 9: Simulation of a transmission in model H_* in the representative graph of $\mathcal{G}^{(d)}$. The arrows indicate the direction of transmission.

Claim 4 *A communication algorithm in G^*, in model F_* can be simulated in $\mathcal{G}^{(d)}$ in model F_1.*

Claim 5 *A communication algorithm in $\mathcal{G}^{(d)}$, in model F_1, can be simulated in $\mathcal{G}$ in model $M_{2,1}$.*

4 Communication patterns

Given a hypergraph $\mathcal{H}(V, E)$ and its representative graph $G(V', E')$, let $gossiping_M(G)$ denote 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 $\mathcal{H}$.

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 $D(G)$ (respectively, $D(\mathcal{H})$) the diameter of G (resp., $\mathcal{H}$).

We use the following for-instruction types in the algorithms.

- for $i = 0$ to $n - 1$
- for each $v \in V$ - execute sequentially for each element in the set.
- for all $v \in V$ - execute in parallel for each element in the set.
- for two $v \in V$ - execute in parallel two elements of the set in each iteration of the for.

4.1 Gossiping

The algorithms consist of three phases, *initial, central* and *final*. The initialization consists of gathering the information from the represented nodes into the representatives. The central

phase is to realize the gossiping in the representative, or reduced, 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 $M_{1,i}$ ($i = 1$ or 2), with $N_u, u \in V'$, the set of represented nodes, and W the set of nodes without representative. Claim 1 leads us to the following algorithm, that takes $2m + D(\mathcal{P}_{n,m}) + (2 - i)$ steps.

1. Initial ($2m - 1$ steps)
   (a) for all $u \in V'$
   for each $v \in N_u, u \leftarrow v$
   (b) for each $v \in W, u_{n-1} \leftarrow v$
2. Central ($n + 1 - i$ steps)
   $gossiping_{H_i}(G)$
3. Final (1 step)
   $W \Leftarrow u_{n-1}$

Analysis. Notice, however, that a straightforward lower bound stems from the fact that the complete information in an extreme bus takes at least m steps to reach its neighbor bus, and then it needs to cross all the remaining buses to be sure that every node is informed. This amounts to $m + D(\mathcal{P}_{n,m}) - 1$ steps, implying the non-optimality of the algorithm above. An optimal one can be designed for the case where $\lfloor m/2 \rfloor \leq \lfloor n/2 \rfloor - 1$, taking advantage of the independence of the buses. Only for the sake of compactness of the description of the algorithm, suppose that n is odd and that the nodes without representative are the ones in the intersection of the middle buses.

1. Initial
   for all $u \in V'$
   for each $v \in N_u, u \leftarrow v$
2. Central
   do in parallel
   - $gossiping_{H_2}(G \cup v_{(n-1)/2})$
   - $u_{(n-1)/2-1} \leftarrow v_{in+(n-1)/2}$, $1 \leq i \leq m/2$, and
   - $u_{(n-1)/2+1} \leftarrow v_{in+(n-1)/2}$, $m/2 < i \leq m$

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

Proof: The first phase can be done in $m - 1$ steps, the second in $D(\mathcal{P}_{n,m}) + 1$ steps and the lower bound is $m + D(\mathcal{P}_{n,m}) - 1$ steps.

Models $M_{2,1}$ and $M_{2,2}$. 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. 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 $\lceil m/2 \rceil + D(\mathcal{P}_{n,m}) - 1$. The algorithm is the same as above, with the exception that data concentration in the Initial phase is done two by two, and that we can use $gossiping_{F_2}(G^* \cup v_{(n-1)/2})$ in the Central phase.

Theorem 2 *Under models $M_{2,1}$ and $M_{2,2}$, the number of steps for gossiping in a hyperpath $\mathcal{P}_{n,m}$, where $\lfloor m/4 \rfloor \leq \lfloor n/2 \rfloor - 1$, is at most $\lceil m/2 \rceil + D(\mathcal{P}_{n,m}) + 1$, which is asymptotically optimal.*

Proof: Supposing again that the nodes without representative are the ones in the intersection of the middle buses, we have the following algorithms for the models $M_{2,j}, j =$ 1, or 2.

Model $M_{2,1}$ ($\lceil m/2 \rceil + D(\mathcal{P}_{n,m}) + 1$ steps)

1. Initial
   for all $u \in V'$
   for two $v \in N_u, \{u_i, u_{i+n-1}\} \leftarrow v$
2. Central
   do in parallel
   - $gossiping_{F_1}(G \cup v_{(n-1)/2})$
   - $u_{(n-1)/2-1} \leftarrow v_{in+(n-1)/2}$, $1 \leq i \leq m/4$
   - $u_{(n-1)/2+1} \leftarrow v_{in+(n-1)/2}$, $m/4 < i \leq m/2$
   - $u_{(n-1)/2-1} \leftarrow v_{in+(n-1)/2}$, $m/2 \leq i \leq 3m/4$
   - $u_{(n-1)/2+1} \leftarrow v_{in+(n-1)/2}$, $3m/4 < i \leq m$

Model $M_{2,2}$ ($\lceil (m-1)/2 \rceil + D(\mathcal{P}_{n,m})$ steps)

1. Initial
   for all $u \in V'$
   for two $v \in N_u, u \leftarrow v$
2. Central
   do in parallel
   - $gossiping_{F_*}(G \cup v_{(n-1)/2})$
   - $u_{(n-1)/2-1} \leftarrow v_{in+(n-1)/2}$, $1 \leq i \leq m/4$
   - $u_{(n-1)/2+1} \leftarrow v_{in+(n-1)/2}$, $m/4 < i \leq m/2$
   - $u_{(n-1)/2-1} \leftarrow v_{in+(n-1)/2}$, $m/2 < i \leq 3m/4$
   - $u_{(n-1)/2+1} \leftarrow v_{in+(n-1)/2}$, $3m/4 < i \leq m$

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

Lower bound

Definition 2 *A hypergraph $\mathcal{H}$ with diameter D, is eccentric if for each node $v \in V$, there exists a node u, such that the distance between v and u, $d(v,u)$, is equal to D.*

Theorem 3 *Given a eccentric hypergraph $\mathcal{H}(V,E)$ with $|V|$ nodes and n hyperedges, the time for gossiping $g_{M_{i,j}}(\mathcal{H})$ is*

$$\geq \frac{|V|}{n} + d(\mathcal{H}) - 1.$$

Proof: In the gossiping all the nodes have to exchange their information, meaning that each node has to write at least one time in an edge.

Let P be an algorithm, let v be the last node to write on a hyperedge for the first time. Let t be the step in what v wrote (first time) on a hyperedge. We have that $t \geq \frac{|V|}{ni}$, because at each step no more than ni nodes can write on a hyperedge.

Let u be a node such that $d(u,v) = \max_{k \in V} d(k,u)$, the information of v will reach u in at least $d(u,v) - 1 = d(\mathcal{H}) - 1$ cycles.

It follows that $g_M(\mathcal{H}) \geq \frac{|V|}{ni} + d(\mathcal{H}) - 1$.

Claim 6 *The hyperring is eccentric.*

The lower bound for gossiping on the hyperring $\mathcal{P}_{n,m}$ in the model $M_{i,j}$ is $\frac{mn}{n} + \lfloor \frac{n}{2} \rfloor - 1$ $= m + \lfloor \frac{n}{2} \rfloor - 1$.

Model $M_{1,i}$.

Proposition 4 *For the model $M_{1,i}(i = 1, 2)$, the number of steps of the best algorithm for gossiping in the hyperring $\mathcal{R}_{n,m}$ is no bigger than $m + \lfloor \frac{n}{2} \rfloor + 2$*

Proof:

Algorithm ($M_{1,2}$)

1. Initial ($m-1$ steps)
   for all $u \in V'$
   for each $v \in N_u, u \leftarrow v$
2. Central ($\lfloor \frac{n}{2} \rfloor + 2$ steps)
   $gossiping_{H_i}(G)$

This is a $m + \lfloor \frac{n}{2} \rfloor + 1$ steps algorithm. Using the simplification * for the model $M_{1,1}$, let $G(V', E')$ be the representative graph and $G^*(V^*, E^*)$ be the reduced graph. Algorithm:

1. Initial (m steps)
   for all $u_i \in V' \mid 1 \leq i \leq n$
   for each $v \in N_{u_i}, \{u_i, u_{i+n}\} \Leftarrow v$
2. Central ($\leq \lfloor \frac{n}{2} \rfloor + 2$ steps)
   $gossiping_{H_*}(G^*)$

This leads to $m + \lfloor \frac{n}{2} \rfloor + 2$ steps.

Model $M_{2,1}$.

Proposition 5 *In the model $M_{2,1}$, we have $g(\mathcal{R}_{n,m}) \leq \lceil \frac{m}{2} \rceil + \lfloor \frac{n}{2} \rfloor + 2$.*

Proof: Let $G(V', E')$ be the representative graph in the model $M_{2,1}$, let $G^*(V^*, E^*)$ be the reduced graph of $G(V', E')$. Algorithm:

1. Initial ($\lceil \frac{m}{2} \rceil$ steps)
   for all $u_i \in V' \mid 1 \leq i \leq n$
   for two $v \in N_{u_i}, \{u_i, u_{n+i}\} \Leftarrow v$
2. Central ($\leq \lfloor \frac{n}{2} \rfloor + 2$ steps)
   $gossiping_{F_1}(G^*)$

Model $M_{2,2}$.

Proposition 6 *In the model $M_{2,2}$, $g(\mathcal{R}_{n,m}) = \lceil \frac{m-1}{2} \rceil + \lfloor \frac{n}{2} \rfloor$.*

Proof: Let $G(V', E')$ be the representative graph in the model $M_{2,2}$. Using the initial phase of model $M_{1,i}$ and the central phase in the model F_* we have the following algorithm.

1. Initial ($\lceil \frac{m-1}{2} \rceil$ steps)
   for all $u \in V'$
   for two $v \in N_u, u \leftarrow v$
2. Central ($\lceil \frac{n}{2} \rceil$ steps)
   $gossiping_{F_*}(G^*)$

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

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 nodes 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 processors get the information, without collisions. Below, we give a somewhat more formal description of the algorithm.

Proposition 7 *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 $\sum_{i=1}^{d} (2m_i - 1) + D(G) + d = 2\sum_{i=1}^{d} m_i - d + D(\mathcal{G})$.*

Proof: Algorithm

1. Initial
   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 all $W^{(j)}_{i_1,\ldots,i_d}$
   for each $v \in W^{(j)}_{i_1,\ldots,i_d}$
   $v_{i_1,\ldots,i_{j-1},u^{(j)}_{n_j-1},i_{j+1},\ldots,i_d} \leftarrow v$
2. Central
   $gossiping_{H_i}(G)$
3. Final
   for $j = d$ to 1
   for all $x^{(j)} \in \mathcal{G}^{(j)}, N_{x^{(j)}} \Leftarrow x^{(j)}$

Analysis. 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 [9]. The Final phase takes 1 step per dimension.

Since the Central phase is always optimal for sufficiently large hypergrids [9], 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.

Proposition 8 *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 $\sum_{i=1}^{d} m_i + D(\mathcal{G})$.*

Proof: Algorithm

1. for $j = 1$ to d
   for all $x^{(j)} \in \mathcal{G}^{(j)}$
   for two $v \in N_{x^{(j)}}, E \Leftarrow v$
   where E is the edge that has the representatives of v in dimension j.
   for all $W^{(j)}_{i_1,\ldots,i_d}$
   for two $v \in W^{(j)}_{i_1,\ldots,i_d}$
   $E \leftarrow v$
   where E is the hyperedge that has v in dimension j.
2. $gossiping_{F_k}(G^*)$

3. for $j = d$ to 1
   for all $x^{(j)} \in \mathcal{G}^{(j)}, N_{x^{(j)}} \Leftarrow x^{(j)}$
   for all $W^{(j)}_{i_1,\ldots,i_d}$
   $W^{(j)}_{i_1,\ldots,i_d} \Leftarrow v_{i_1,\ldots,i_j,u^{(j)}_{n_j-1},i_{j+1},\ldots,id}$

Analysis. The Initial phase takes m_i steps per dimension. The Central phase takes $D(G^*) = D(\mathcal{G}) - d$ steps. The Final phase takes d steps.

Model $M_{2,d}$. The number of steps 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 9 *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 $\sum_{i=1}^d m_i + D(\mathcal{G})$.*

Proof: Algorithm

1. 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 all $W^{(j)}_{i_1,\ldots,i_d}$
   for two $v \in W^{(j)}_{i_1,\ldots,i_d}$
   $v_{i_1,\ldots,i_{j-1},u^{(j)}_{n_j-1},i_{j+1},\ldots,i_d} \leftarrow v$
2. $gossiping_{F_*}(G)$
3. Same as step 3 of the previous algorithm.

Hypertorus By using the same ideas for the hypergrid, similar algorithms can be derived for hypertorus. We state without proof the following. Given a hypertorus $\mathcal{T}^d(V,E) = \mathcal{R}^1_{n_1,m_1} \times \ldots \times \mathcal{R}^d_{n_d,m_d}$. The number of steps for the gossiping is smaller than or equal to

- $\sum_{i=1}^d (m_i - 1) + g_{H_i}(G) + d$, in model $M_{1,i}$
- $\sum_{i=1}^d (\lceil \frac{m_i}{2} \rceil) + g_{F_1}(G) + d$, in model $M_{2,1}$
- $\sum_{i=1}^d (\lceil \frac{m_i - 1}{2} \rceil) + g_{F_*}(G) + d$, in model $M_{2,*}$

where G and G^* are respectively the representative graph and the reduced graph in the correspondent model.

The details are omitted since they do not constitute new ideas.

In the conclusion section we show these values as function of the diameter.

5 Conclusion

Due mostly to VLSI implementation issues, point-to-point interconnection networks are inadequate for dealing with real size applications in massively parallel systems. We argue that 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 gossiping 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 [10, 11]. BINs represent a very interesting way of dealing with massive parallelism.

Below we summarize our results for gossip on some BINs. For the sake of simplicity, we use D to denote the corresponding hypertopology diameter.

Gossiping	$M_{1,1}$	$M_{1,2}$
$\mathcal{P}_{n,m}(V,E)$	$2m+D+1$	$2m+D$
$\mathcal{R}_{n,m}(V,E)$	$m+D+2$	$m+D+1$
$\mathcal{G}(V,E)=\mathcal{P}^1_{n_1,m_1}\times\ldots\times\mathcal{P}^d_{n_d,m_d}$	$2\sum_{i=1}^{d}(m_i)-d+D$	$2\sum_{i=1}^{d}(m_i)-d+D$
$\mathcal{T}^d(V,E)=\mathcal{R}^1_{n_1,m_1}\times\ldots\times\mathcal{R}^d_{n_d,m_d}$	$\sum_{i=1}^{d}(m_i)+D+18d+39$	$\sum_{i=1}^{d}(m_i)+D+2d$

Gossiping	$M_{2,1}$	$M_{2,2}$
$\mathcal{P}_{n,m}(V,E)$	$\lceil\frac{m}{2}\rceil+D+1$	$\lceil\frac{m-1}{2}\rceil+D$
$\mathcal{R}_{n,m}(V,E)$	$\lceil\frac{m}{2}\rceil+D+2$	$\lceil\frac{m}{2}\rceil+D$
$\mathcal{G}(V,E)=\mathcal{P}^1_{n_1,m_1}\times\ldots\times\mathcal{P}^d_{n_d,m_d}$	$\sum_{i=1}^{d}(m_i)+D$	$\sum_{i=1}^{d}(m_i)+D$
$\mathcal{T}^d(V,E)=\mathcal{R}^1_{n_1,m_1}\times\ldots\times\mathcal{R}^d_{n_d,m_d}$	$\sum_{i=1}^{d}(\lceil\frac{m_i}{2}\rceil)+3d+D$	$\sum_{i=1}^{d}(\lceil\frac{m_i-1}{2}\rceil)+d+D$

6 Acknowledgment

We would like to thank the referees for their helpful comments that improved the paper considerably.

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, J. Bond and C. Peyrat, Interconnection networks with each node on two buses, in *Parallel Algorithms and Architectures*, eds. M. Cosnard, Y. Robert, P. Quinton and M. Tchunte, (North Holland, 1984) 155–167.

[6] J-C. Bermond and F.O. Ergincan, Bus interconnection networks, *Technical Report No. 93-56*, Université de Nice - Sophia Antipolis, September 1993.

[7] L.N. Bhuyan and D.P. Agrawal, Generalized hypercube and hyperbus structures for a computer network, *IEEE Transactions on Computers* **33** (1984) 323–333.

[8] S.H. Bokhari, Finding maximum on an array processor with a global bus, *IEEE Transactions on Computers* **33** (1984) 133–139.

[9] 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.

[10] 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.

[11] A. Ferreira, A. Goldman vel Lejbman, S.W. Song, Broadcasting in bus interconnection networks, *CONPAR 94 / VAPP VI*, (Lecture Notes of Computer Science, Springer-Verlag), September 1994.

[12] C.M. Fiduccia, Bused hypercubes and other pin-optimal networks, *IEEE Transactions on Parallel and Distributed Systems* **3(1)** (1992) 14–24.

[13] 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.

[14] 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.

[15] D.W. Krumme, G. Cybenko, and K.N. Venkataraman. Gossiping in minimal time. *SIAM Journal of Computing*, Vol. 21, No. 1, pp. 111 – 139.

[16] F. Meyer auf der Heide and H.T. Pham, On the performance of networks with multiple buses, in *STACS'92 9th Annual Symposium on Theoretical Aspects of Computer Science*, eds. A. Finkel and M. Jantzen, *Lecture Notes in Computer Science*, (Springer-Verlag, 1992) 97–108.

[17] 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.

[18] 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.

[19] Q.F. Stout, Meshes with multiple buses, *Proc. 27th IEEE Symposium on the Foundations of Computer Science* (1986) 264–273.

[20] 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.

[21] B. Wilkinson, On crossbar switch and multiple bus interconnection networks with overlapping connectivity, *IEEE Transactions on Computers* **41** (1992) 738–746.