Circuit-Switched Gossiping in the 3-Dimensional Torus Networks

Olivier Delmas, Stéphane Pérennes

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Circuit-switched gossiping in the 3-dimensional torus networks

Olivier Delmas, Stéphane Perennes

N° 2930

Juillet 1996

THÈME 1

apport de recherche

ISSN 0249-6399

Circuit-switched gossiping in the 3-dimensional torus networks

Olivier Delmas, Stéphane Perennes

Thème 1 — Réseaux et systèmes
Projet SLOOP

Rapport de recherche n˚2930 — Juillet 1996 — 20 pages

Abstract: in this paper we describe, in the case of short messages, an efficient gossiping algorithm for 3-dimensional torus networks (wrap-around or toroidal meshes) that uses synchronous circuit-switched routing. The algorithm is based on a recursive decomposition of a torus. The algorithm requires an optimal number of rounds and a quasi-optimal number of intermediate switch settings to gossip in an $7^i \times 7^i \times 7^i$ torus.

Key-words: circuit-switching, gossiping, torus network, toroiral mesh, linear code theory.

(Résumé : tsvp)

Version of July 2, 1996. A short version of this paper will appear in *Proceedings of the EuroPar'96 Conference* (Lyon - France - August 26-29, 1996).

Email : `{delmas, sp}@unice.fr`

Unité de recherche INRIA Sophia-Antipolis
2004 route des Lucioles, BP 93, 06902 SOPHIA-ANTIPOLIS Cedex (France)
Téléphone : (33) 93 65 77 77 – Télécopie : (33) 93 65 77 65

Diffusion en mode commutation de circuits dans les tores de dimension 3

Résumé : dans cet article nous décrivons, pour des messages courts, un algorithme de diffusion efficace dans les réseaux toriques (grilles toriques) de dimension 3 qui utilise un routage par commutation de circuits synchone. L'algorithme est basé sur une décomposition récursive du tore. L'algorithme requiert un nombre d'étape optimal et un nombre de commutateur intermédiaire quasi-optimal pour diffuser dans le tore $7^i \times 7^i \times 7^i$.

Mots-clé : commutation de circuits, diffusion, réseau torique, grille torique, théorie des codes linéaires.

1 Introduction

DISTRIBUTED memory multicomputer systems in which the processors communicate by exchanging messages over an interconnection network are a very popular method for achieving cost-effective high-performance computing. The performance of a message-passing system is strongly dependent on the topology of the interconnection network and on the routing mechanism that is used to move information around the network. Multi-dimensional tori (wrap-around meshes, toroidal meshes, k-ary n-cube) are currently popular interconnection networks because their low degrees permit efficient layouts and construction with standard components. The large diameters of tori are a disadvantage when store-and-forward routing is used because communication time for store-and-forward routing is proportional to the diameter of the network. Store-and-forward routing has been displaced by circuit-switched routing in many recent multicomputer systems such as the Intel Paragon, IBM SP2, Cray-T3D or more recently the new Cray-T3E. Since the cost of circuit-switched routing is less dependent than store-and-forward on the diameter of a network, torus networks with circuit-switched routing are a practical choice.

Gossiping is an all-to-all information dissemination problem in which a distinct message originating at each node of a network must be distributed to all other nodes of the network. The gossiping problem has been studied for many different network topologies and routing strategies. So far, research on gossiping mainly concentrated on unit cost store-and-forward models in which each message transmission travels along one communication link and takes one unit of time. Much of the recent research on gossiping has used linear cost models in which the propagation time of a message is proportional to the length of the message. Linear costs models have been used to study both store-and-forward routing [9, 17, 29] and various types of circuit-switched routing including direct connect [25], virtual cut-through [18], and wormhole routing [24]. See [10] for a survey of store-and-forward routing under both unit cost and linear cost model.

In this paper, we present a new gossiping algorithm for 3-dimensional torus networks which uses circuit-switched routing and a linear cost model. We prove that our algorithm is optimal in terms of number of rounds and quasi-optimal in term of number of intermediate switch settings when the messages are short or when the time to initiate a message transmission is much larger than the unit propagation time of a message along a link. This is the situation in many current multiprocessor networks.

We will first describe in section 2 the different hypothesis of communication used in this paper. Then in section 3 we give some definitions and notations before to establish in section 4 lower bounds for gossiping protocols in circuit-switched routing. In section 5 we will study the gossiping protocol in the $3-$dimensional torus network. We give the protocol for the $7^i \times 7^i \times 7^i$ torus network. Note that we study the case of power of 7 because in our model this is the extremal case in terms of numbers of processors which can receive a message originating at one processor in a given number of rounds. The proof uses recursion and is based on a detailed study of the $7 \times 7 \times 7$ torus network. The number of of rounds of the protocols matches the lower bound established in section 4. In the last part we discuss the case of large messages.

2 Models of communication

2.1 Basic models

First we assume that all links have the same bandwidth. In the case of a basic routing (one-to-one) problem, we briefly describe two of the most used protocols of communications as following.

- When store-and-forward routing (without pipe-line techniques) is used to send a message along a path of d links, the message is stored in buffers at intermediate nodes on the path. An intermediate node does not begin to send the message to the next node on the path until it has received the entire message.

  Thus, the transmission time for a message of length L to be sent at distance d in the linear cost model is $d(\beta + L\tau)$ where β is the time to initiate a message transmission and $1/\tau$ is the bandwidth of the communication links.

- When circuit-switched routing is used, a header containing the destination address is sent through the network to "build" a path. At each intermediate node on the path, the input and output ports used by the header are connected. When the destination node receives the header, it sends an acknowledgment back to the source node establishing a direct electrical physic connection between source and destination. The bytes of the message are then sent in pipeline fashion. Since the message is switched through intermediate nodes, there is no need to buffer the entire message. The links of the path can be released as the last byte passes through each node or by an acknowledgment from the destination when the last byte is received. The former case is known as direct connect. In a wormhole implementation of circuit-switching, the header establishes a path to the destination in the same way as in a direct connect implementation, but an acknowledgment is not sent back to the source node. Instead, the remaining bytes immediately follow the header in pipeline fashion with the last byte releasing the switches as it passes through. The Torus Routing Chip described in [4] uses wormhole routing for point-to-point (i.e., one-to-one) communications and can be used to build multidimensional tori [3].

  In the linear cost model, the transmission time for a message of length L to be sent at distance d is $\alpha + d\delta + L\tau$, where α is the time to initiate a new message transmission, δ is the time to switch an intermediate node, and $1/\tau$ is the bandwidth of the communications links.

  In most current machines, message transmissions are initiated in software and switching is done in hardware, so δ is usually much smaller than α. Furthermore, α is usually much larger than τ.

Remark 2.1 Store-and-forward routing can be simulated by circuit-switched routing by restricting d to be 1 for all transmissions, so the store-and-forward upper bounds are also true for circuit-switched system with $\beta = \alpha + \delta$.

2.2 Complementary hypothesis

In this paper we will consider the circuit-switched routing model. If we want to study global communication problems like gossiping, we need to precise the model.

- We will use the link-bounded [10] (or shouting [11] or all-port) model of communication in which a processor can use all of its communications links simultaneously. In contrast, the processor-bounded (or whispering) model permits the use of only one link for each node at any given time.
- We also assume that the communication links are full-duplex so that messages can travel in both directions simultaneously.
- Finally, we assume that each node has an initial distinct message, but all these messages have the same length L. Furthermore we allow messages to be concatenated with negligible cost.

The reader will find more complementary informations on the models of communication in [28, 31] or for wormhole routing in [22, 23] and in the forthcoming survey [2].

Remark 2.2 When any type of circuit-switched routing is used, and communication patterns are arbitrary, deadlock is possible and many papers on wormhole routing are more concerned with deadlock avoidance than efficiency [5, 20, 8, 7, 21]. The most common deadlock avoidance method is the use of virtual channels which uses multiplexing to share physical links. Here we will design synchronous algorithms decomposed into rounds in which is no possibility of deadlock. Indeed, we will enforce that during a round all the dipaths corresponding to the communications are arc-disjoint. Efficient global communication problems with the arc(edge)-disjoint hypothesis have been also studied (see for example [13, 27]).

Under these hypothesis, the protocol given in this paper is a minimum phase circuit-switched gossiping algorithm for link-bounded systems.

3 Definitions

In this article, we will use the following definitions and notations.

- Z_q will denote the set of integers modulo q, addition of elements in Z_q will always occur in Z_q.
- G will denote a **digraph** with vertex set $V(G)$ and arc set $A(G)$.
- N will denote the **number of vertices** of G, that is $N = |V(G)|$.
- $d_G(x, y)$ will denote the **distance** from a vertex $x \in V(G)$ to a vertex $y \in V(G)$ that is the length of the shortest dipath from x to y.
- $D(G)$ will denote the **diameter** of a digraph G, that is the maximum of the distances between every couple of vertices: $D(G) = max_{(x,y)\in V^2(G)} d_G(x, y)$.
- In a symmetric digraph G, $\Delta(G)$ (or shortly Δ) will denote **maximum in-degree (or out-degree)** of G, that is the maximum over the in-degrees (or out-degrees) of all vertices $V(G)$.

- C_N will denote the symmetric circuit of order N.

Definition 3.1 *The* **cartesian sum** *(also called* **cartesian product** *or* **box product***) denoted by* $G \Box G'$ *of two digraphs* $G = (V, A)$ *and* $G' = (V', A')$*, is the digraph whose vertices are the pairs* (x, x') *where* x *is a vertex of* G *and* x' *is a vertex of* G'*. Two vertices* (x, x') *and* (y, y') *of* $G \Box G'$ *are adjacent if and only if* $x = y$ *and* $[x', y']$ *is an arc of* G'*, or if* $x' = y'$ *and* $[x, y]$ *is an arc of* G*.*

Definition 3.2 **The k-dimensional torus** *is the cartesian sum of* k *symmetric circuits of orders* $p_1, p_2, \ldots, p_k$ *and is denoted by* $TM(p_1, p_2, \ldots, p_k) = C_{p_1} \Box C_{p_2} \Box \cdots \Box C_{p_k}$*.*

Remark 3.3 If all $p_i \geq 3$, it is a regular digraph of degree $\Delta = 2k$. Its order is $p_1 \times p_2 \times \cdots \times p_k$, the number of arcs is $2kN$ and its diameter is $\sum_{i=1}^{k} \lfloor \frac{p_i}{2} \rfloor$.

Notation 3.4 *When* $p_1 = p_2 = \cdots = p_k$*, we will use the abbreviated notation* $TM(p)^k$ *and suppose in what follows that* $p \geq 3$*.*

Remark 3.5 **The k-dimensional torus** can be defined as a Cayley digraph on an abelian group. It is a vertex-transitive digraph where each vertex x of the $TM(p_1, p_2, \cdots, p_k)$ digraph can be seen as a k-tuple vector $(x_1, x_2, \cdots, x_k) \in Z_{p_1} \times Z_{p_2} \times \cdots \times Z_{p_k}$. Each vertex $(x_1, x_2, \cdots, x_k)$ (or k-tuple vector) is joined by an arc to the $2k$ vertices (or k-tuple vectors) $(x_1, x_2, \cdots, x_i \pm 1, \cdots, x_k) \in Z_{p_1} \times Z_{p_2} \times \cdots \times Z_{p_i} \times \cdots \times Z_{p_k}$ for $1 \leq i \leq k$.

Definition 3.6 *A* **routing function** R *for a digraph* G *is a set of* $N(N-1)$ *dipaths* $R = \{R(x, y) | x, y \in V(G)\}$*, where* $R(x, y)$ *is a dipath in* G *from* x *to* y*.*

Definition 3.7 *Given a routing function* R *for the digraph* G*, the* **load of an arc** $u \in A(G)$*, denoted by* $\pi(G, R, u)$*, is the number of dipaths of* R *going through* u*.*

Definition 3.8 *Given a routing function* R *for the digraph* G*, the* **arc-forwarding index of** (G, R)*, denoted by* $\pi(G, R)$*, is the maximum number of dipaths of* R *going through any arc of* G*, that is* $\pi(G, R) = max_{u \in A(G)}\ \pi(G, R, u)$*.*

Definition 3.9 *The* **arc-forwarding index of the digraph** G*, denoted by* $\pi(G)$*, is defined as* $\pi(G) = min_R\ \pi(G, R)$*.*

Notation 3.10 *The total time necessary to achieve a gossiping protocol in a digraph* G *will be denoted by* $g(G) = g_\alpha(G)\alpha + g_\delta(G)\delta + g_\tau(G)\tau$*:* $g_\alpha(G)$ *is the number of rounds,* $g_\delta(G)$ *the sum of the maximum communication distances of each couples of processors implicated in each round of the gossiping protocol and* $g_\tau(G)$ *measure the flow of information.*

As said before, here we don't use pipe-line techniques and consider only short messages (or equivalently suppose $\tau << \alpha$ and $\tau << \delta$). So we are mainly interested in determining the optimal $g_\alpha(G)$ and $g_\delta(G)$ parameters.

For $g_\delta(G)$ a trivial lower-bound is the diameter $D(G)$. In the next section we give lower-bound for $g_\alpha(G)$.

4 Lower bounds

We first give some additional definitions and notations.

Definition 4.1 *Given a gossip protocol in a digraph G, we will say that the couple $(x, y) \in V(G) \times V(G)$* **uses the arc** $u \in A(G)$ **at round** t *if a message originated from x and finally reaching y goes through the arc u at round t of the gossip protocol.*

Definition 4.2 *Given a gossip protocol in a digraph G, the* **gossip load** *of an arc $u \in A(G)$ at round t denoted $\mathcal{GL}(u, t)$ is the number of couples $(x, y) \in V^2(G)$ using the arc u at round t of the gossip protocol.*

Definition 4.3 *Given a gossip protocol in a digraph G completed at round $g_\alpha(G)$, the* **total gossip load** *of an arc $u \in A(G)$ noted $\mathcal{TGL}(u)$ is defined as $\mathcal{TGL}(u) = \sum_{t=0}^{g_\alpha(G)} \mathcal{GL}(u, t)$.*

Proposition 4.4 *For a gossip protocol in a digraph G, there exists an arc u such that $\mathcal{TGL}(u) \geq \pi(G)$.*

Proof. Any gossip algorithm constructs finally at least a dipath from each vertex to each other one. By choosing for each couple (x, y) one of these dipaths, we can associate to a gossip algorithm a routing R for the digraph G. For any dipath created from x to y by the gossip algorithm and for any arc u of the dipath there exists a round t such that the couple (x, y) uses the arc u at round t. Hence the total gossip load of any arc u is at least the load of u for the routing R. It remains to observe that there exists an arc of G whose load for the routing R is at least $\pi(G)$. $\square$

Let G be a digraph of maximum degree Δ, we calculate an upper bound of the total gossip load of an arc $u \in A(G)$, for any gossip protocol with $g_\alpha(G)$ rounds.

Proposition 4.5 *Let G be a digraph with maximum degree Δ and order N, let $t_0 = \lceil \log_{\Delta+1}(N) \rceil$, if $g_\alpha(G) \leq 2t_0$ then*

$$\forall u \in A(G),\ \mathcal{TGL}(u) \leq N[(\frac{2}{\Delta} + 2t_0 - g_\alpha(G))(\Delta+1)^{g_\alpha(G)-t_0} - \frac{2}{\Delta}]$$

Proof. The total gossip load of an arc u is the sum of the the gossip load of u at round t for all the rounds $t \in \{0 \ldots g_\alpha(G)\}$. At the end of the round $t-1$ a vertex knows at most $min((\Delta+1)^{t-1}, N)$ informations and so at round t at most $min((\Delta+1)^{t-1}, N)$ informations can go throught arc u. After round t the information which has gone throught u will be able to reach at most $min((\Delta + 1)^{g_\alpha(G)-t}, N)$ nodes. Hence the gossip load of u at round t cannot exceed $min((\Delta+1)^{t-1}, N) \cdot min((\Delta+1)^{g_\alpha(G)-t}, N)$. So, as $(\Delta+1)^{t_0} \geq N$ and $(\Delta+1)^{t_0-1} < N$ (and we suppose that $g_\alpha(G) \leq 2t_0$, that is $g_\alpha(G) - t_0 \leq t_0$), we have the following array:

Round	$[0, \cdots, g_\alpha(G) - t_0]$	$[g_\alpha(G) - t_0 + 1, \cdots, t_0]$	$[t_0 + 1, \cdots, g_\alpha(G)]$
$g_\alpha(G) - t$	$(\Delta+1)^{g_\alpha(G)-t} \geq N$	$(\Delta+1)^{g_\alpha(G)-t} < N$	$(\Delta+1)^{g_\alpha(G)-t} < N$
$t-1$	$(\Delta+1)^{t-1} < N$	$(\Delta+1)^{t-1} < N$	$(\Delta+1)^{t-1} \geq N$
$\mathcal{GL}(u,t)$	$N(\Delta+1)^{t-1}$	$(\Delta+1)^{g_\alpha(G)-1}$	$N(\Delta+1)^{g_\alpha(G)-t}$

Hence, the number of the dipaths on the arc u is:

$$\begin{aligned}
\mathcal{TGL}(u) &\leq N \sum_{i=0}^{g_\alpha(G)-t_0-1} (\Delta+1)^i + (2t_0 - g_\alpha(G))(\Delta+1)^{g_\alpha(G)-1} + N \sum_{i=0}^{g_\alpha(G)-t_0-1} (\Delta+1)^i \\
&\leq 2N \sum_{i=0}^{g_\alpha(G)-t_0-1} (\Delta+1)^i + (2t_0 - g_\alpha(G))(\Delta+1)^{g_\alpha(G)-1} \\
&\leq 2N \frac{(\Delta+1)^{g_\alpha(G)-t_0} - 1}{\Delta} + (2t_0 - g_\alpha(G))(\Delta+1)^{g_\alpha(G)-t_0}(\Delta+1)^{t_0-1} \\
&\leq N((\Delta+1)^{g_\alpha(G)-t_0}(\frac{2}{\Delta} + 2t_0 - g_\alpha(G)) - \frac{2}{\Delta}), \text{ as } (\Delta+1)^{t_0-1} < N
\end{aligned}$$

□

Consequently for any digraph we have the following theorem.

Theorem 4.6 *Let G be a digraph with maximum degree Δ and order N, let $t_0 = \lceil \log_{\Delta+1}(N) \rceil$, if $g_\alpha(G) \leq 2t_0$ then*

$$g_\alpha(G) \geq t_0 + \log_{\Delta+1}(\frac{\pi(G)}{N}) - O(\log_{\Delta+1}\log_{\Delta+1}(N))$$

Proof. By the proposition 4.4 we have,

$$\begin{aligned}
N((\Delta+1)^{g_\alpha(G)-t_0}(\frac{2}{\Delta} + 2t_0 - g_\alpha(G)) - \frac{2}{\Delta}) &\geq \pi(G) \\
(\Delta+1)^{g_\alpha(G)-t_0}(\frac{2}{\Delta} + 2t_0 - g_\alpha(G)) &\geq \frac{\pi(G)}{N}
\end{aligned}$$

When we take the logarithm in base $\Delta + 1$ we have

$$g_\alpha(G) - t_0 + \log_{\Delta+1}(\frac{2}{\Delta} + 2t_0 - g_\alpha(G)) \geq \log_{\Delta+1}(\frac{\pi(G)}{N})$$

$$g_\alpha(G) \geq t_0 + \log_{\Delta+1}(\frac{\pi(G)}{N}) - \log_{\Delta+1}(\frac{2}{\Delta} + 2t_0 - g_\alpha(G))$$

$$\text{then } g_\alpha(G) \geq t_0 + \log_{\Delta+1}(\frac{\pi(G)}{N}) - O(\log_{\Delta+1}(t_0))$$

□

In order to exploit the bound above we can use the lower bound on the arc-forwarding index.

Proposition 4.7 (M.C. Heydemann, J.C. Meyer, D. Sotteau [12]) *The arc-forwarding index for the $TM(n)^k$ symmetric digraph is $\pi(TM(n)^k) = \frac{n^{k-1}}{2}\lfloor\frac{n^2}{4}\rfloor$.*

Now, we are able to state the following corollary.

Corollary 4.8 *Given a gossip protocol in the $TM(n)^k$ digraph of degree $\Delta = 2k$, the number of rounds necessary to achieve this protocol is:*

$$g_\alpha(G) \geq (k+1)\log_{2k+1}(n) - O(\log_{2k+1}\log_{2k+1}(n))$$

Proof. This corollary is correct as in [2, 6] it has been shown that the total number of rounds to achieve a broadcasting protocol in the $TM(n)^k$ digraph is t_0. Then a trivial gossiping protocol will be the concatenation of 2 broadcasting protocols, so $g_\alpha(G) \leq 2t_0$. □

Remark 4.9 There exists an analogy between the arc-forwarding index and the vertex bisection, since in [19] R. Klasing establish a lower bound of the same kind (in a model named two-way vertex-disjoint paths mode) by using vertex bisection width.

5 Gossiping in the 3-dimensional torus TM$(7^i)^3$

5.1 Case of TM$(7)^3$

The idea of this section come from the original study from J.G. Peters and M. Syska [26] for the circuit-switched broadcast in the 2-dimensional torus. Note that the gossiping in the 2-dimensional torus was studied in [1] and other results on circuit-switched structured communications on torus can be found in [30]. Gossiping can be derived from the broadcast protocol given in [6] and the forthcoming paper [2]. We will recall briefly the principle of this broadcast protocol. In the following, we will make use of the following notations and definitions.

- Here G denote the $TM(7)^3$ symmetric digraph.
- According remark 3.5 we will consider vertices of G as elements of the 3-dimensional vector-space Z_7^3, with canonical base $\{e_1, e_2, e_3\}$. Vertex $(x_1e_1 + x_2e_2 + x_3e_3)$ will be denoted as $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$.
- M will be a 3-dimensional matrix, given a set of vectors U, MU will denote the image of U by $M : \{Mx \mid x \in U\}$.
- The sum of two sets of vectors U_1 and U_2 will be $U_1 + U_2 = \{x \mid x = u_1 + u_2, u_1 \in U_1, u_2 \in U_2\}$.

Definition 5.1 *$\mathcal{E}$ is the set of the 3-dimensional vector-space Z_7^3 associated to all the vertices of G that is $\forall x \in V(G)$.*

Definition 5.2 *We will denote by B_1 the set $\{e_1, e_2, e_3, 0, -e_1, -e_2, -e_3\}$.*

Remark 5.3 Note that B_1 is the sphere of radius 1 centered on zero of Z_7^3 for the Lee distance. $x+B_1$ is the sphere of radius 1 centered on x, it also contains the neighbours of x in G union x (see figure 1).

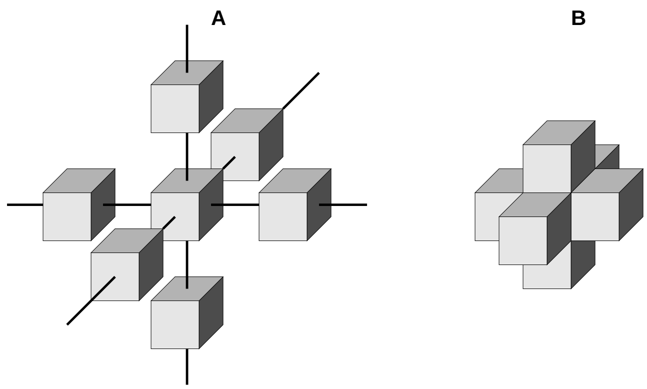

Figure 1: *A Lee sphere B_1. Cubes represent vertices and the central cube is a vertex of the code $\mathcal{C}$ (see definition 5.4). On figure 1-A we represent the links (a link represents symmetric arcs) and on figure 1-B the representation without links. It's this pattern that we must pack to cover the entire torus (see remark 5.7).*

Definition 5.4 *The* **code** $\mathcal{C}$ *is the set of vertices such that* $\mathcal{C} = \{(x_1, x_2, x_3) \in Z_7^3 | x_1 + 2x_2 + 3x_3 = 0\}$

Definition 5.5 *Let* $M_0 = \begin{pmatrix} -2 & 0 & 1 \\ 1 & 2 & 3 \\ 0 & 1 & 0 \end{pmatrix}$ *and note that* $M_0^2 = \begin{pmatrix} -3 & 1 & -2 \\ 0 & 0 & 0 \\ 1 & 2 & 3 \end{pmatrix}$.

The broadcast algorithm relies on the following properties:

Property 5.6 $\mathcal{C} = M_0^2 B_1 + M_0 B_1$, and $\mathcal{E} = \mathcal{C} + B_1$.

Proof. See [2] □

Remark 5.7 $\mathcal{C}$ is a linear code of length 3 defined over Z_7. As $\mathcal{C} + B_1 = \mathcal{E}$, then $\mathcal{C}$ is a perfect Lee code. Note that $\mathcal{C}$ has 49 elements like $(0,0,0), (-2,1,0), (2,-3,-1), \cdots$.

Proposition 5.8 *There exists a gossiping protocol on the symmetric digraph* $G = TM(7)^3$ *with time* $g(G) = 4\alpha + 12\delta + (1 + 7 + 7^2 + 7^3)L\tau$.

To describe the gossiping protocol of proposition 5.8 we introduce an additional notation.

Notation 5.9 *Let x be a vertex of a digraph G and let $\mathcal{A} \subset V(G)$ a subset of the set of vertices of G. The notation $x \to \mathcal{A}$* (resp. $x \leftarrow \mathcal{A}$) *is used when the vertex x sends his message towards all the vertices of $\mathcal{A}$ (*resp. all the vertices of $\mathcal{A}$ send their own message toward the vertex x*).*

5.1.1 Description of the gossiping algorithm in $TM(7)^3$

The algorithm that we present here is similar to a method designing so-called *"three-phase algorithms"* [14, 15, 16] (this method uses an accumulation, a gossip and a broadcast phase).

Begin ———— *Gossiping Algorithm* ———— *in $TM(7)^3$.*

All the information of the digraph G is concentrated (more exactly equally distributed) on the set of vertices of the code $\mathcal{C}$. This round uses dipaths of length 1 and messages of length L.

▷ **Round 1** ◁ **Concentration**	
Scheme of communication	*Cost*
$\forall x \in \mathcal{C}, x \leftarrow \{x + B_1\}$	$\alpha + \delta + L\tau$

Here the scheme of communications is decomposed into two rounds. We will describe later the arc-disjoint dipaths used to perform these two rounds, and prove the cost of these rounds.

▷ **Step 2** ◁ **Gossiping between the vertices of $\mathcal{C}$**	
Round 2-a	
Scheme of communication	*Cost*
$\forall x \in \mathcal{C}, x \to \{x + M_0 B_1\}$	$\alpha + 5\delta + 7L\tau$
Round 2-b	
Scheme of communication	*Cost*
$\forall x \in \mathcal{C}, x \to \{x + M_0^2 B_1\}$	$\alpha + 5\delta + 7^2 L\tau$

Each vertex of the code $\mathcal{C}$ send his information to its 6 direct neighbours. This round uses dipaths of length 1 and messages of length $7^3 L$.

▷ **Round 3** ◁ **Final broadcasting**	
Scheme of communication	*Cost*
$\forall x \in \mathcal{C}, x \to \{x + B_1\}$	$\alpha + \delta + 7^3 L\tau$

End ———— *Gossiping Algorithm* ———— *in $TM(7)^3$.*

5.1.2 Analysis of the algorithm

- **Round 1:** As $\mathcal{C}$ is a perfect code, the first round equally distributes the entire information of G on the set of the vertices of $\mathcal{C}$.
- **Step 2:** During second step, each vertex $x \in \mathcal{C}$ has sent his information to $x + M_0 B_1 + M_0^2 B_1 = x + \mathcal{C} = \mathcal{C}$, thus step 2 performs a gossiping between the vertices of $\mathcal{C}$. Note that at the end of the round 1, vertices of $\mathcal{C}$ have a message of length $L_0 = 7L$. We will show in section 5.1.3 that rounds 2-a and 2-b use dipaths of length at most 5, hence step 2 is completed in time $(\alpha + 5\delta + L_0\tau) + (\alpha + 5\delta + 7L_0\tau)$.

After step 2 each vertex x of $\mathcal{C}$ has received the whole information initially distributed on G.

- **Round 3:** As $\mathcal{E} = \mathcal{C} + B_1$, this last round enables us to achieve the gossiping of the digraph.

5.1.3 Dipaths used by the algorithm

Now the problem is to exhibit a set of dipaths in G realizing each round of communication of the algorithm. As rounds 1 and 3 raise no problem, we give now the dipaths associated to step 2.

Round 2-a (resp. 2-b) uses communications of the kind $\forall x \in \mathcal{C}$, $x \rightarrow \{x + A\}$ (resp. $x \rightarrow \{x + A'\}$) with $A = M_0 B_1$ (resp. $A' = M_0^2 B_1$). In both cases, A and A' are symmetric, that is of the kind: $u_1, u_2, u_3, 0, -u_1, -u_2, -u_3$. We will describe for each u_i the dipath associated, the dipath associated to $-u_i$ being the opposite one. As there are many possible dipaths from x to $x + u_i$, we will describe them by decomposing the associated dipaths with $u_i = a_1 + a_2 + \ldots + a_l$, where a_i is proportional to some vector of the base, meaning that from x we first go to $x + a_1$ along vector a_1 then to $x + a_1 + a_2$ following a_2 and so on.

Dipaths used in the round 2-a In that case the vectors u_1, u_2, u_3 are

$$\left\{ \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix} \right\} \text{ with } \begin{cases} \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \begin{pmatrix} -2 \\ 0 \\ 0 \end{pmatrix} \\ \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix} + \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix} \\ \begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 3 \\ 0 \end{pmatrix} \end{cases}$$

The maximum length of these dipaths is $max(|+1|+|-2|; |-1|+|+2|+|+2|; |+1|+|+3|) = 5$ and the length of the messages is $7L$. Indeed after the first round each vertex $x \in \mathcal{C}$ has accumulated 7 informations . The cost of this round is $\alpha + 5\delta + 7L\tau$. *Figure* 2-(2a) shows clearly that the scheme of communications described above is arc-contention free.

Dipaths used in the round 2-b In that case the vectors u_1, u_2, u_3 are

$$\left\{ \begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}, \begin{pmatrix} -2 \\ 0 \\ 3 \end{pmatrix} \right\} \text{ with } \begin{cases} \begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -3 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \\ \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \\ \begin{pmatrix} -2 \\ 0 \\ 3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 3 \end{pmatrix} + \begin{pmatrix} -2 \\ 0 \\ 0 \end{pmatrix} \end{cases}$$

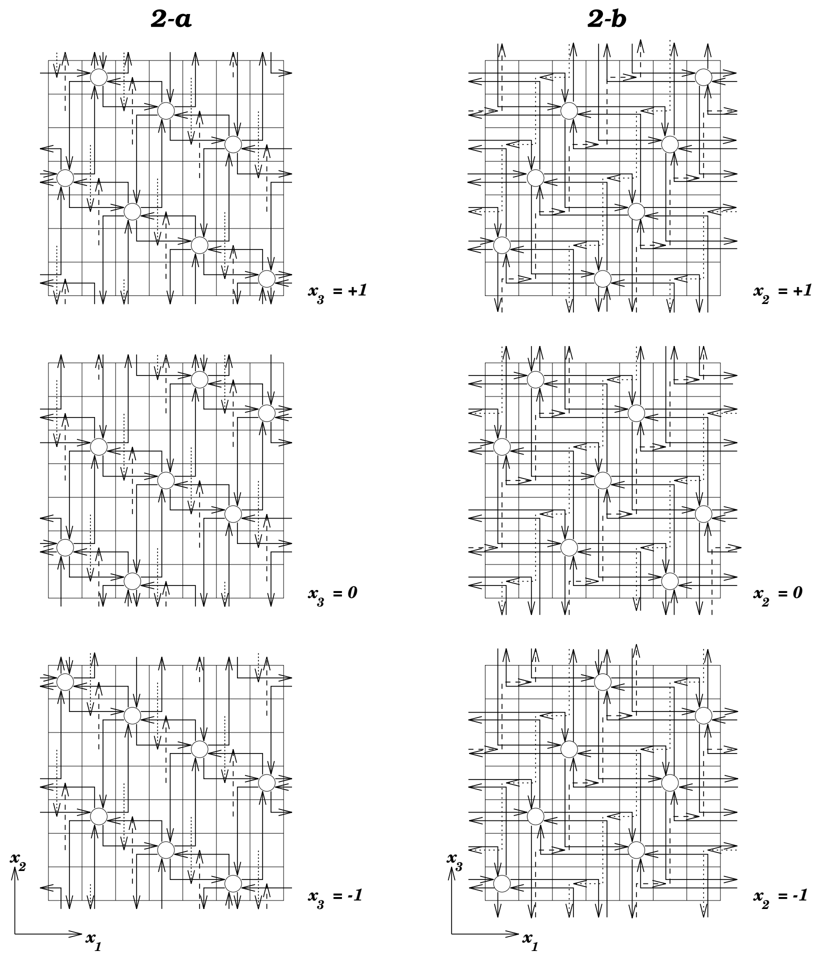

Figure 2: *Communication patterns for step 2 of the gossiping algorithm in the* $TM(7)^3$ *digraph. This step is decomposed into two rounds 2-a and 2-b. In the left* (resp. right) *drawing the torus is displayed with layers for a fixed* x_3 (resp. x_2). *Here, only three layers are displayed. In these figures the solid arcs correspond to the vectors* $u_1, -u_1, u_3, -u_3$*, the dashed arcs to vectors* u_2 *and the dotted arcs to vectors* $-u_2$ *for each round 2-a and 2-b.*

The maximum length of the dipaths is $max(|-3|+|+1|;|-1|+|+2|+|+1|+|+1|;|+3|+|-2|) = 5$ and the length of the messages is 7^2L. Indeed, after round 1 and 2-a each vertex of the code has 7×7 informations. The cost of this round is $\alpha + 5\delta + 7^2L\tau$. *Figure* 2-(2b) shows clearly that the scheme of communications used in this round creates no-arc-contention.

5.2 Generalization for torus $TM(7^i)^3$

The main idea of this section is to use recursively the gossiping protocol designed for the torus $TM(7)^3$ to the torus $TM(7^i)^3$.

Notation 5.10 *Here G_i denotes the symmetric digraph $TM(7^i)^3$.*

Definition 5.11 *The* **code** *$\mathcal{C}_i$ is the set of the vertices of G_i defined as $\mathcal{C}_i = \{(x_1, x_2, x_3) \in Z^3_{7^i} | x_1 + 2x_2 + 3x_3 \equiv 0$ (mod 7)$\}$.*

Remark 5.12 This code is once again a perfect code for the Lee distance. Indeed, spheres of radius 1 centered on each vertex of the code $\mathcal{C}_i$ cover completely the digraph G_i. That is $V(G_i) = Z^3_{7^i} = B_1 + \mathcal{C}_i$. The code $\mathcal{C}_i$ has 7^{3i-1} elements.

Notation 5.13 *Let U_0 be a sub-group of $Z^3_{7^i}$ defined as $U_0 \equiv 0$ (mod 7).*

Remark 5.14 U_0 is clearly isomorphic to $Z^3_{7^{i-1}}$. U_0 of vector $x = (x_1, x_2, x_3)$ such that $x_1 \equiv 0$ (mod 7), $x_2 \equiv 0$ (mod 7) and $x_3 \equiv 0$ (mod 7) has $7^{i-1} \times 7^{i-1} \times 7^{i-1}$ vertices (or vectors).

Definition 5.15 *The* **family** *associated to a vector (or vertex) x denotes the set of vectors (or vertices) defined as $x + U_0$.*

Lemma 5.16 *In the $TM(7^i)^3$ symmetric digraph the vertices of the code $\mathcal{C}_i$ form 7^2 disjoint dilated sub-torus of $TM(7^{i-1})^3$ symmetric digraph with a dilated factor of 7.*

Proof. The distance between two vertices of U_0 is a multiple of 7. If we join by a dipath of length 7 any couple of vertices at distance exactly 7, we obtain a sub-graph H_0 of G_i which is a dilated sub-torus (obtained from the torus G_{i-1} by dilating each arc in a dipath of length 7). Now we can partition the vertices of the code $\mathcal{C}_i$ into 7^2 disjoint families. Indeed, any vertex (x_1, x_2, x_3) of the code $\mathcal{C}_i$ belongs to the family $(a, b, c) + U_0$ where $x_1 \equiv a$ (mod 7), $x_2 \equiv b$ (mod 7) and $x_3 \equiv c$ (mod 7) with $-3 \leq a \leq 3$, $-3 \leq b \leq 3$ and $-3 \leq c \leq 3$. So, we have 7^2 possible choices for (a, b, c). Indeed, c is determined as soon as we fixe a and b. If we consider the subgraph $H_{a,b}$ generated by the vertices of the family associated to (a, b, c), two vertices at distance 7 being joined by a dipath of length 7, $H_{a,b}$ is isomorphic to H_0, the dilated sub-torus of $TM(7^{i-1})^3$. Note that by definition any two different $H_{a,b}$ have no arc in common. So we can in a given round do concurrently communications on each $H_{a,b}$ (or families). $\square$

Proposition 5.17 *There exists a gossiping protocol on the symmetric digraph $G_i = TM(7^i)^3$ with time $g(G_i) = 4i\alpha + \frac{12}{9}D(G_i)\delta + [\frac{57}{49}(7^{i-1} - 1) + \frac{7^3}{7^{2i}} - \frac{1}{7^{3i}}]\frac{NL}{6}\tau$.*

We describe now the gossiping protocol of this proposition.

5.2.1 Description of the gossiping algorithm in $TM(7^i)^3$

Here the gossiping algorithm in $TM(7^i)^3$ is similar to gossiping algorithm in $TM(7)^3$.

Begin ________ *Gossiping Algorithm* ________ *in* $TM(7^i)^3$.

All the information of the digraph G_i is concentrated (more exactly equitably distributed) on the set of vertices of the code $\mathcal{C}_i$. This round uses dipaths of length 1 and messages of length L.

$\triangleright$ **First Round** **Concentration** $\triangleleft$	
Scheme of communication	*Cost*
$\forall x \in \mathcal{C}_i,\, x \leftarrow \{x + B_1\}$	$\alpha + \delta + L\tau$

Vertices of the code $\mathcal{C}_i$ are partitioned into families. The vertices of the same family perform the gossiping protocol on the dilated sub-torus $TM(7^{i-1})^3$ recursively. We will describe later the cost of this phase.

$\triangleright$ **Recursive phase** $\triangleleft$
Scheme of communication
Recursive gossiping protocol on G_{i-1}
Cost
$g_\alpha(G_{i-1})\alpha + 7g_\delta(G_{i-1})\delta + 7g_\tau(G_{i-1})L\tau$

For this step, the communications are similar to communications used in the step 2 of the gossiping protocol for the vertices of the code on Z_7^3. At the beginning of this step the length of messages is $\frac{N}{7^2}L$.

$\triangleright$ **Local Gossiping Step** **Gossiping between the vertices of** $\mathcal{C}_i$ $\triangleleft$	
Round a	
Scheme of communication	*Cost*
$\forall x \in \mathcal{C}_i,\, x \rightarrow \{x + M_0 B_1\}$	$\alpha + 5\delta + \frac{N}{7^2}L\tau$
Round b	
Scheme of communication	*Cost*
$\forall x \in \mathcal{C}_i,\, x \rightarrow \{x + M_0^2 B_1\}$	$\alpha + 5\delta + \frac{N}{7}L\tau$

Each vertex of the code $\mathcal{C}_i$ send his information at its 6 direct neighbours. This round uses dipaths of length 1 and messages of length $7^{3i}L$.

$\triangleright$ **Last Round** **Final broadcasting** $\triangleleft$	
Scheme of communication	*Cost*
$\forall x \in \mathcal{C}_i,\, x \rightarrow \{x + B_1\}$	$\alpha + \delta + 7^{3i}L\tau$

End ________ *Gossiping Algorithm* ________ *in* $TM(7^i)^3$.

5.2.2 Analysis of the algorithm and recursion

- The **first round** raises no problem.
- At the beginning of the **recursive phase** the length of messages is $7L$. As this phase performs a gossiping protocol on the dilated sub-torus G_{i-1}, then the distances of the communications have a dilated-factor of 7 and this recursive phase can be applied on each family of a same partition in the same time because these families are disjoint.

- As this **local gossiping step** used the same scheme of communication that the step 2 on $TM(7)^3$, then each vertex of a family receives an information from a vertex of each other families. So, the cost of this step is $(\alpha + 5\delta + \frac{N}{7^2}L\tau) + (\alpha + 5\delta + 7\frac{N}{7^2}L\tau)$.
- This **last step** raises no problem. Just note that at this step each vertex of the code $\mathcal{C}_i$ knows all the information of G_i, so the length of the messages is $7^{3i}L$.

From this algorithm and recursion we obtain (with $N = 7^{3i}$),

$$\begin{array}{rclcl} g_\alpha(G_i) & = & g_\alpha(G_{i-1}) + 4 & = & 4i \\ g_\delta(G_i) & = & 7g_\delta(G_{i-1}) + 12 & = & \frac{12}{9}D(G_i) \\ g_\tau(G_i) & = & 7g_\tau(G_{i-1}) + 1 + (1 + \frac{1}{7} + \frac{1}{7^2})N & = & [\frac{57}{49}(7^{i-1} - 1) + \frac{7^3}{7^{2i}} - \frac{1}{7^{3i}}]\frac{N}{6} \end{array}$$

6 Case of large messages

In the previous parts we presented the case for short messages as we focused only on the parameters $\alpha(G)$ and $\delta(G)$. But our protocols are not efficient if we consider large messages, indeed the predominant parameter becomes $\tau(G)$. So our previous algorithms implie a maximal parameter $\tau(G)$ at the last round (final broadcasting). In this part we discuss the case of large messages and will show that it is possible to decrease the parameter $\tau(G)$ using classical methods, if we authorize to increase the parameter $\alpha(G)$ (see [1]). Here we use the following definition and lemma.

Definition 6.1 *Let x be a vertex (or vector) of the $TM(7^i)^3$ symmetric digraph, then $c(x)$ will denote the color of the vertex x defined by $c(x) = (x_1 + 2x_2 + 3x_3)$ mod 7, so $c(x) \in \Omega = \{-3, -2, -1, 0, 1, 2, 3\}$.*

Lemma 6.2 *The vertices of the $TM(7^i)^3$ digraph can be partitioned with 7 different colors such as a vertex x and his direct neighbours have all a distinct different color.*

Proof. Just use the $c(x)$ colors and note that the direct neigbours of the vertex x are the vertices with ± 1 on only one coordonate x_1 or x_2 or x_3, then x and his direct neighbours have all a distinct different value. □

6.1 Case of $TM(7)^3$

As saw in section 5.1, at the beginning of the last round (final broadcasting) of the gossiping protocol on the symmetric digraph $G = TM(7)^3$, all the vertices of the code $\mathcal{C}$ know the entire information of G. During this last round each vertex of the code send the total information to their direct neighbours using their out-links (out-arcs) and no out-arc of the vertices which are not in the code are used. The idea is to split this last round into two using out-arcs of the vertices which are not in the code to discrease the data flow on the arcs by involving smaller messages.

Proposition 6.3 *There exists a gossiping protocol on the symmetric digraph $G = TM(7)^3$ with time $g(G) = 5\alpha + 13\delta + (1 + 7 + 3 \cdot 7^2)L\tau$.*

Proof. We split the last round (final broadcasting) into two.

Just before step 3 the message is splitted into 7 pieces and during these rounds the vertices of the code send only a piece of the total message (that is $1/7th$ of the total length of the initial message). During round 3-b the vertices which are not in the code send the message they have received at round 3-a toward their direct neighbours.

▷ **Step 3** ◁ **Final broadcasting**	
Round 3-a	
Scheme of communication	*Cost*
$\forall x \in \mathcal{C}, x \xrightarrow{\frac{1}{7}} \{x + B_1\}$	$\alpha + \delta + \frac{7^3}{7} L\tau$
Round 3-b	
Scheme of communication	*Cost*
$\forall x \in \mathcal{C}, x \xrightarrow{\frac{1}{7}} \{x + B_1\}$ and $\forall x \notin \mathcal{C}, x \to \{x + B_1\}$	$\alpha + \delta + \frac{7^3}{7} L\tau$

At the beginning of round 3-a, all the vertices of the code $\mathcal{C}$ which know the entire information of G, partition their total message m of length 7^3 into 7 pieces (as for the concatenation we forget the time necessary to partition the message). Now we give a different color from the set Ω for each piece of the message and we will denote each one by its color that is $m = m_{-3}, m_{-2}, \cdots, m_2, m_3$ (this operation is exactly the same for all the vertices of the code). Then, during step 3-a each vertex of the code sends the piece of message m_i to y $\forall y \in \{x + B_1\}$ if and only if $c(y) = i$. So, at the end of this round x $\forall x \notin \mathcal{C}$ knows the piece of message $m_{c(x)}$.

Now, the lemma 6.2 assure that the round 3-b achieves the gossiping protocol. Just note that during this last round the vertices of the code send the piece of message m_0. □

6.2 Case of $TM(7^i)^3$

Now we are able to state the generalization for the symmetric digraph $G_i = TM(7^i)^3$.

Proposition 6.4 *There exists a gossiping protocol on the symmetric digraph $G_i = TM(7^i)^3$ with time $g(G_i) = 5i\alpha + \frac{13}{9}D(G_i)\delta + [\frac{22}{49}(7^{i-1} - 1) + \frac{19 \cdot 7}{7^{2i}} - \frac{1}{7^{3i}}]\frac{NL}{6}\tau$.*

Proof. As the lemma 6.2 is true for G_i, just split the last round of each recursive phase and the last round (final broadcasting) of the gossiping protocol in $TM(7^i)^3$ by the 2 rounds described in the proof of proposition 6.3. □

Remark 6.5 If we are not concerned by the $g_\alpha(G)$ parameter, it is possible to discrease many more the data flow, as we can simulate the store-and-forward model with the wormhole model. Such technics in the store-and-forward model are described in the survey [10] and P. Fraignaud presents in [9] a protocol which discreases significantly the data flow as he obtains for the graph $TM(p)^k$ of order N, $g(TM(p)^k) \leq k\lfloor \frac{p}{2} \rfloor(\alpha + \delta) + \frac{(N-1)}{2k}\tau$.

Acknowledgments

The authors are grateful to J-C. Bermond and M. Syska for helpful discussions and remarks.

References

[1] C. Calvin, S. Perennes, and D. Trystram. Gossiping in Torus with Wormhole-Like Routing. In *Proceedings of the "7^{th} IEEE Symposium on Parallel and Distributed Processing"*, 1995.

[2] Jean de Casanice. An overview of wormhole routing. I3S - CNRS URA 1376 - Manuscript, 1996.

[3] W.J. Dally. Performance analysis of $k-$ary $n-$cube interconnection networks. *IEEE Trans. Computers*, C-39(6):775–785, 1990.

[4] W.J. Dally and C.L. Seitz. The torus routing chip. *J. Distributed Computing*, 1(3):187–196, 1986.

[5] W.J. Dally and C.L. Seitz. Deadlock-free message routing in multiprocessor interconnection networks. *IEEE Trans. Computers*, C-36(5):547–553, May 1987.

[6] O. Delmas and S. Perennes. Diffusion en mode commutation de circuits. In *RenPar'8*, pages 53–56, Bordeaux, France, 20-24 May 1996. Edition spéciale, présentation des activités du GDR-PRC Parallélisme, Réseaux et Systèmes, Richard Castanet and Jean Roman.

[7] J. Duato. A necessary and sufficient condition for deadlock-free adaptive routing in wormhole networks. *IEEE Transactions on Parallel and Distributed Systems*, 6(10), October 1995.

[8] J. Duato. A theory of deadlock-free adaptive multicast routing in wormhole networks. *IEEE Transactions on Parallel and Distributed Systems*, 6(9), September 1995.

[9] P. Fraigniaud. *Communications intensives dans les architectures à mémoire distribuée et algorithmes parallèles pour la recherche de racines de polynômes*. PhD thesis, Université de Lyon I, E.N.S.L, 1990.

[10] P. Fraigniaud and E. Lazard. Methods and problems of communication in usual networks. *Discrete Applied Mathematics*, 53:79–133, 1994.

[11] S.M. Hedetniemi, S.T. Hedetniemi, and A.L. Liestman. A survey of gossiping and broadcasting in communication networks. *Networks*, 18:319–349, 1986.

[12] M.C. Heydemann, J.C. Meyer, and D. Sotteau. On forwarding indices of networks. *Discrete Applied Mathematics*, 23:103–123, 1989.

[13] C.T. Ho and M.Y. Kao. Optimal broadcast in all-port wormhole-routed hypercube. *IEEE Transactions on Parallel and Distributed Systems*, 6(2):200–204, February 1995.

[14] J. Hromkovič, R. Klasing, and E.A. Stöhr. Gossiping in vertex-disjoint paths mode in interconnection networks. In Springer Verlag, editor, *Proc. 19th Int. Workshop on Graph-Theoretic Concepts in Computer Science (WG'93)*. LNCS. to appear.

[15] J. Hromkovič, R. Klasing, E.A. Stöhr, and H. Wagener. Gossiping in vertex-disjoint paths mode in $d-$dimensional grids and planar graphs. In Springer Verlag, editor, *Proc. First Annual European Symposium on Algorithms (ESA'93)*, pages 200–211. LNCS 726, 1993.

[16] J. Hromkovič, R. Klasing, W. Unger, and H. Wagener. Optimal algorithms for broadcast and gossip in the edge-disjoint paths modes. In Springer Verlag, editor, *Proc. 4th Scandinavian Workshop on Algorithm Theory (SWAT'94)*. LNCS. to appear.

[17] S.L. Johnsson and C.T. Ho. Optimum broadcasting and personalized communication in hypercubes. *IEEE Trans. Computers*, 38(9):1249–1268, 1989.

[18] P. Kermani and L. Kleinrock. Virtual cut-through: a new computer communication switching technique. *Computers Networks*, 3:267–286, 1979.

[19] R. Klasing. The relationship between gossiping in vertex-disjoint paths mode and bisection width. In *Proc. 19th Int. Symp. on Mathematical Foundations of Computer Science*, 1994.

[20] X. Lin, P.K. McKinley, and A-H. Esfahanian. Adaptative multicast wormhole routing in 2D mesh multicomputers. *Journal of Parallel and Distributed Computing*, 28(1):19–31, July 1995.

[21] D.H. Linder and J.C. Harden. An adaptive and fault tolerant wormhole routing strategy for $k-$ary $n-$cubes. *IEEE Trans. Computers*, 40(1):2–12, January 1991.

[22] P.K. McKinley, Y-J. Tsai, and D.F. Robinson. A survey of collective communication in wormhole-routed massively parallel computers. Technical Report MSU-CPS-94-35, Michigan state University, June 1994.

[23] P.K. McKinley, Y-J. Tsai, and D.F. Robinson. Collective communication trees in wormhole-routed massively parallel computers. Technical Report MSU-CPS-95-6, Michigan state University, March 1995.

[24] L.M. Ni and P.K. McKinley. A survey of wormhole routing techniques in direct networks. *Computers*, 26(2):62–76, feb 1993.

[25] S.F. Nugent. The iPSC/2 direct-connect technology. In G.C. Fox, editor, *Proceedings of 3rd Conference on Hypercube Concurrent Computers and Applications*, pages 51–60. ACM, 1988.

[26] J.G. Peters and M. Syska. Circuit-switched broadcasting in torus networks. *IEEE Transactions on Parallel and Distributed Systems*, 7(3):246–255, March 1996.

[27] D.F. Robinson, D. Judd, P.K. McKinley, and B.H.C. Cheng. Efficient multicast in all-port wormhole-routed hypercubes. *Journal of parallel and distributed computing*, 31:126–140, 1995.

[28] Jean de Rumeur. *Communication dans les réseaux de processeurs*. Collection Etudes et Recherches en Informatique. Masson, 1994. (English version to appear).

[29] Y. Saad and M.H. Schultz. Data comunication in parallel architectures. *Parallel Computing*, 11:131–150, 1989.

[30] C. Spencer. *Circuit-switched structured communications on toroidal meshes*. PhD thesis, Simon Fraser University, February 1994.

[31] M. Syska. *Communications dans les architectures à mémoire distribuée*. PhD thesis, Université de Nice - Sophia Antipolis, I3S, 1992.

Table of Contents

Unité de recherche INRIA Lorraine, Technopôle de Nancy-Brabois, Campus scientifique,
615 rue du Jardin Botanique, BP 101, 54600 VILLERS LÈS NANCY
Unité de recherche INRIA Rennes, Irisa, Campus universitaire de Beaulieu, 35042 RENNES Cedex
Unité de recherche INRIA Rhône-Alpes, 46 avenue Félix Viallet, 38031 GRENOBLE Cedex 1
Unité de recherche INRIA Rocquencourt, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex
Unité de recherche INRIA Sophia-Antipolis, 2004 route des Lucioles, BP 93, 06902 SOPHIA-ANTIPOLIS Cedex

Éditeur
INRIA, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex (France)
ISSN 0249-6399