Optimal routing in toroidal networks

Izidor Jerebic and Roman Trobec
Jožef Stefan Institute
Jamova 39, 61 111 Ljubljana, Slovenia

Abstract

In this paper we study routing algorithms for one-to-one communication in multiprocessors, whose interconnection networks have toroidal structure. A toroidal network is an n-dimensional rectangular mesh with additional edge-to-edge connections. Mathematically speaking, these networks are undirected graphs obtained by cartesian product of cycles.

We show, that a broad class of routing algorithms, called homogeneous, is optimal for certain class of such networks. Besides, we propose a transformation, which transforms non-optimal homogeneous routing algorithms into optimal semi-homogeneous routing algorithms. [1]

Keywords: optimal routing, multiprocessors, toroidal networks

1 Introduction

Toroidal interconnection networks are becoming the standard interconnection networks for multiprocessors [6, 5, 7]. Figure 1 shows a two-dimensional toroidal network.

An extensive comparative analysis of the latency in such networks was done by Dally in [7]. An optimal dynamic routing policy was presented by Badr and Podar [1]. A deadlock-free routing algorithm was obtained by Dally and Seitz [2] and later generalized by Linder and Harden [4]. The exact lower bound for load in such networks was given by Heydemann et al. [3].

To use these networks efficiently, one needs good routing algorithms. In our paper we consider static routing algorithms, i.e. routing algorithms, where the path of a message does not depend on the current state of the network. Such routing algorithm is said to be optimal, if it distributes the static communication load equally among all communication channels (for precise definition of the load, see section 2.2 in this paper).

At the moment, the most common routing algorithm calculates the distance vector and decrements its components in some specified order. This algorithm is simple to implement and easy to understand. It is optimal, though only for networks with odd edge-size. The optimality can be deduced from the proof for the exact lower bound for load in [3].

The optimality was thus a consequence of the construction of the algorithm. There were no criteria, concerning the inner structure of the algorithm, showing that the algorithm is

[1] In Information Processing Letters 43, 285-291(1992)

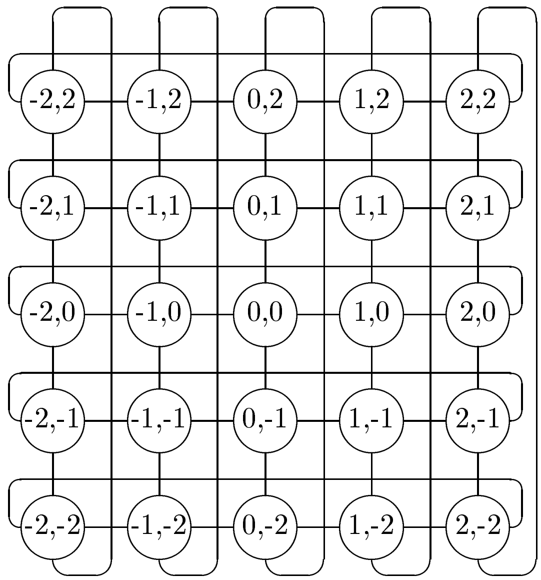

Figure 1: Two-dimensional toroidal network T_5^2.

optimal. The only criterion was the load on edges of a graph, by which the pure optimality of communication was defined.

In this paper we present criteria for optimality, that require regularity of an algorithm to be examined. These criteria define a very large class of optimal algorithms. Also the above-mentioned simple algorithm belongs to this class.

The structure of the paper is as follows. The next section contains formal definitions and notation used in the paper. In section 3 proofs concerning optimality criteria are presented. The next section describes an improvement of a homogeneous routing algorithm for even values of the network's size. Section 5 is an example of the application of our theorems. The last section contains a summary of results presented in the paper.

2 Formal definitions

An *interconnection network* is an undirected graph $IN = (V, E)$, where vertices V represent processors, and edges E represent communication channels between processors. We assume that each edge represents two unidirectional channels. So, bidirectional communication, without any impact of the message flow in one direction on the communication in other direction, is possible.

If the interconnection network is toroidal, we can represent vertices of the graph with integer vectors:

$$V \subset \mathcal{Z} \times \mathcal{Z} \times \ldots \times \mathcal{Z},$$

where $\mathcal{Z}$ denotes the set of all integers. We refer to the vector representation of vertices as *points*.

In our discussion we will assume the symmetry of the network and denote the set of vertices by T_K^n or shortly T, when it is not necessary to know the size of a network:

$$T_K^n = \underbrace{\mathcal{Z}_K \times \mathcal{Z}_K \times \ldots \times \mathcal{Z}_K}_{n},$$

where $\mathcal{Z}_K$ denotes the set of K consecutive integers, starting at $-\lfloor \frac{K-1}{2} \rfloor$.

This simplification doesn't influence the correctness of results, since they can be easily generalized to toroidal networks with different number of vertices in each dimension.

Neighbourhood relation

First, we define the addition $\oplus$ and subtraction $\ominus$ as operations in a cyclical group of a size K, where K is the size of the set to which both operands belong. This means that $\lceil \frac{K-1}{2} \rceil \oplus 1 = -\lfloor \frac{K-1}{2} \rfloor$ and $-\lfloor \frac{K-1}{2} \rfloor \ominus 1 = \lceil \frac{K-1}{2} \rceil$.

For the norm of a vector we will use the vector 1-norm:

$$\|A\| = \sum_{i=1}^{n} |a_i|.$$

Now we can express the neighbourhood relation $\sim$ between two vertices, represented with vectors A and B:

$$A \sim B \iff \|A \ominus B\| = 1.$$

2.1 Communication

A message from the point A comes to the point B over intermediate points X_i. The ordered sequence P of the points is called a *path* from the point A to the point B:

$$P(A,B) = \langle X_1, X_2, \ldots, X_k \rangle \iff (X_i \sim X_{i+1}) \wedge (X_1 = A) \wedge (X_k = B).$$

When a message from A destined for B comes to the intermediate point X_i, it has to be sent to the next point X_{i+1} on the path, except if the point X_i is the destination point B. The decision, to which neighbour this message will be sent, is called a *routing algorithm* R:

$$R : \langle A, B, X_i \rangle \mapsto X_{i+1}.$$

We assume that paths between all points in the network are defined, so messages can be sent from any point to any other point.

We furthermore assume a non-adaptive routing algorithm, what means that there is only one possible path between any two points A and B. All messages from the point A take the same route to the point B.

The *length* L of a path P is the number of components minus one:

$$L(P(A,B)) = k - 1.$$

A unidirectional communication connection from a point X to a neighbouring point Y is called a *channel* and is denoted by $c_{X,Y}$.

A path from A to B can be equivalently expressed as an ordered sequence of channels instead of points:

$$P(A,B) = \langle c_{X_1,X_2}, c_{X_2,X_3}, \ldots, c_{X_{k-1},X_k} \rangle.$$

If a path is expressed with channels, the length of the path is the number of the channels in the path.

2.2 Optimality

The *message density* ρ on a channel $c_{X,Y}$ is the greatest number of messages that can reach the point X and are routed forward to the point Y. This equals the number of all possible paths, defined by a routing algorithm, that contain the channel $c_{X,Y}$:

$$\rho(c_{X,Y}) = |\{P : c_{X,Y} \in P\}|.$$

The *load* α in a network T, when using the routing algorithm R, is the maximum over all message densities:

$$\alpha(R) = \max_{X,Y \in T} \rho(c_{X,Y}).$$

The routing algorithm R^* in a network T is called *optimal* when it minimizes the load:

$$\alpha(R^*) \leq \alpha(R),\ \forall R.$$

3 Optimality criteria

Two paths $P_1 = (X_1, X_2, \ldots, X_p)$ and $P_2 = (Y_1, Y_2, \ldots, Y_q)$ are *congruent* ($P_1 \longleftrightarrow P_2$) iff components of the paths differ by a constant and the numbers of components in the paths are equal:

$$P_1 \longleftrightarrow P_2 \iff (p = q) \wedge (\exists D : X_i = Y_i \oplus D, 1 \leq i \leq p).$$

The routing algorithm R is *homogeneous* iff for a path from any point A to any other point B from a network T holds, that it is congruent to the path from the point $A \ominus B$ to the point $\vec{0} = (0, 0, \ldots, 0)$:

$$\forall A, B \in T : P(A,B) \longleftrightarrow P(A \ominus B, \vec{0}).$$

Lemma 1 If a routing algorithm R in a toroidal network T is homogeneous, then the message densities on channels in the same dimension and direction are the same:

$$\forall X' \sim Y', X'' \sim Y'' :\ X' \ominus Y' = X'' \ominus Y'' \Longrightarrow \rho(c_{X',Y'}) = \rho(c_{X'',Y''}).$$

Proof:

We will prove this assertion by reductio ad absurdum. Suppose the inverse of our claim is true. Therefore there exist two pairs of neighbouring points $A' \sim B'$ and $A'' \sim B''$, such that the following holds:

$$A' \ominus B' = A'' \ominus B'' = D,$$

and

$$\rho(c_{A',B'}) \neq \rho(c_{A'',B''}).$$

This means that for some vectors X and Y the path $P(A' \oplus X, B' \oplus Y)$ contains the pair (A', B'), i.e. $c_{A',B'} \in P(A' \oplus X, B' \oplus Y)$, while the path $P(A'' \oplus X, B'' \oplus Y)$ doesn't contain the pair (A'', B''). This is because the message density is the number of paths containing the particular channel, and if the message densities differ, there must be a path across one channel, which doesn't have a corresponding path across the other channel. But we will show that for *every* path across one channel there exists a corresponding path across the other channel and vice versa.

If we use the definition of a homogeneous algorithm, we get:

a) $(A', B') \in P(A' \oplus X, B' \oplus Y)$:

$$P(A' \oplus X, B' \oplus Y) \longleftrightarrow P(D \oplus X \ominus Y, \vec{0}),$$

and from the above formula we get:

$$(D \ominus Y,\ \ominus Y) \in P(D \oplus X \ominus Y, \vec{0}). \tag{1}$$

b) $(A'', B'') \notin P(A'' \oplus X, B'' \oplus Y)$:

$$P(A'' \oplus X, B'' \oplus Y) \longleftrightarrow P(D \oplus X \ominus Y, \vec{0}),$$

and from the above formula we get:

$$(D \ominus Y,\ \ominus Y) \notin P(D \oplus X \ominus Y, \vec{0}). \tag{2}$$

But (1) and (2) are contradictory. We conclude that the assumption was wrong and there are no such points A', B', A'', B''. □

Lemma 1 assures us that channels in the same dimension and direction have the same message density. What is left to consider is the relation between message densities on channels in different dimensions and directions. This result is given by the following theorem.

Theorem 1 Let T_K^n be a toroidal network and K be an **odd** positive integer. Every homogeneous routing algorithm, using only the shortest paths between two points, is optimal for routing in the network T_K^n. Message densities are the same for all channels and are equal to $\frac{1}{8}(K^2 - 1)K^{n-1}$. This is by definition also the load:

$$\forall X \sim Y \in T_K^n : \rho(c_{X,Y}) = \alpha = \frac{1}{8}(K^2 - 1)K^{n-1}.$$

Proof:

Let's calculate the message density on channels in some arbitrary dimension i. To determine the message density, we project the network in the dimension i.

The network is partitioned in K (n-1)-dimensional sub-networks in the dimension i, i.e. the points with the same i-th component are grouped together. Each group is replaced by a vertex, labeled with the i-th component of the points it represents. The vertices are thus labeled with numbers from $-\frac{K-1}{2}$ to $\frac{K-1}{2}$. All channels between neighbouring points, belonging to different sub-networks, are joined in a single connection.

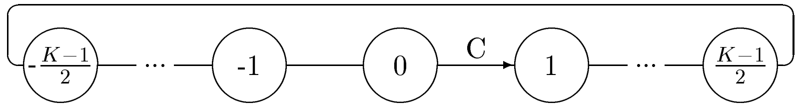

Figure 2: Projection of the T_K^n in dimension i.

The result is a single cycle with K vertices, each representing K^{n-1} points in the original network. Each connection in the cycle represents $2K^{n-1}$ channels in both directions (K^{n-1} in one direction). This structure is seen on Figure 2.

Now we calculate the message flow in one direction between two neighbouring vertices in the projection cycle. Denote the number of possible messages with N.

Since this cycle is symmetric, the result for one connection is valid also for others. We picked points 0 and 1 and direction from 0 to 1, labeled on the Figure 2 with C. Now, look at the table that depicts message flow over C for each sub-network:

source	destinations	# of dest.
0	$1..\frac{K-1}{2}$	$\frac{K-1}{2}$
-1	$1..\frac{K-1}{2}-1$	$\frac{K-1}{2}-1$
$\vdots$	$\vdots$	$\vdots$
$-\frac{K-1}{2}+1$	$1..1$	1

In the first column there are vertices (sub-networks), from which messages are sent. In the second column there are all possible destination vertices, such that the messages cross the connection C. The third column contains the number of such destination vertices.

From this table the following equation ensues:

$$N = \sum_{d=1}^{\frac{1}{2}(K-1)} K^{n-1} \cdot d \cdot K^{n-1}.$$

The variable d is actually following the third column, being multiplied with the number of points in source and destination sub-networks.

And since we know from Lemma 1 that all channels in the same dimension and the same direction have the same message density, we get the following expression for the message density on channels in the dimension i:

$$\rho_i = \frac{N}{K^{n-1}} = \frac{\frac{1}{8}(K^2-1)K^{2n-2}}{K^{n-1}} = \frac{1}{8}(K^2-1)K^{n-1}.$$

Since the procedure is independent of actual value of i and direction, the conclusion is that ρ_i is the message density on all channels. □

We have the theorem about odd K. What happens if K is even? In this case, the additional point in a projection cycle breaks the symmetry. The following theorem formulates this more precisely.

Theorem 2 Let T_K^n be a toroidal network and K be an **even** positive integer. Define furthermore ρ_i^+ as the message density on a channel, where the difference vector between the drain point of the channel and the source point of the channel has $+1$ as i-th component. Define similarly ρ_i^- as the message density on the corresponding channel in the other direction. For every homogeneous routing algorithm, these two values are the same for all channels in the same direction and dimension and are $\frac{1}{8}K^n(K-2)$ and $\frac{1}{8}K^n(K+2)$. Which value corresponds to which direction depends on each particular routing algorithm.
Proof:

The proof is similar to the proof of Theorem 1. In addition, we assume that the routing algorithm sends a message from the point, which has the i-th component 0, destined for the point with the i-th component $\frac{K}{2}$, in direction of decreasing i-th components, i.e. towards the point with the i-th component -1. Such algorithm is called an *i-negative asymmetric* algorithm. Algorithms that take the other direction are called *i-positive asymmetric* algorithms.

Denote the original connection C from Figure 2 with C^+ and its corresponding reverse direction with C^-.

Now observe paths of the length $\frac{K}{2}$. Since our routing algorithm is homogeneous and i-negative asymmetric, no paths of length $\frac{K}{2}$ cross the connection C^+. Obviously the message density ρ_i^- is greater than ρ_i^+.

With the similar approach as in Theorem 1 – counting possible source and destination points – we get the following expressions for ρ_i^- and ρ_i^+ for a homogeneous i-negative asymmetric routing algorithm:

$$\rho_i^+ = \frac{1}{8}K^n(K-2),$$

$$\rho_i^- = \frac{1}{8}K^n(K+2).$$

If a routing algorithm is i-positive asymmetric, the values of ρ_i^- and ρ_i^+ just exchange places. □

We obtained a result, which tells us that homogeneous routing algorithms in even-sized networks are not optimal. But the majority of networks in use today *are* even-sized. The routing algorithms for such networks should not be homogeneous algorithms, since they are not optimal, though they are easy to implement.

4 Optimal routing for even-sized networks

For even-sized networks, homogeneous routing algorithms should be adjusted, in order to remove the asymmetry induced by points at distance $\frac{K}{2}$. This adjustment, however, causes also the algorithms not to be homogeneous any more. But, instead, they are optimal.

There is a very simple adjustment algorithm for networks with K divisible by four.

Let R be a homogeneous routing algorithm in an even-sized toroidal network T_K^n. Define the *semi-homogeneous* routing algorithm R' as the algorithm R with changes in routing at distance $\frac{K}{2}$. For every dimension i pick arbitrary $\frac{K}{4}$ positive numbers between 1 and $\frac{K}{2}$ and denote the set containing them with S_i^+. Denote the set of remaining numbers with S_i^-.

The algorithm R' works as the original algorithm R when the difference vector between source and destination contains no $\frac{K}{2}$. So, if distances in all dimensions are less than $\frac{K}{2}$, routing proceeds without changes.

Denote the i-th component of a source point as s_i. The algorithm R' behaves as an i-positive asymmetric algorithm when s_i or $s_i + \frac{K}{2}$ belong to the set S_i^+. If s_i or $s_i + \frac{K}{2}$ belong to the set S_i^-, the algorithm R' behaves as an i-negative asymmetric algorithm.

If this adjustment algorithm is used in network with K not divisible by 4, for example $K = 6$, the resulting load is not optimal, but it is close to the optimal load. More complicated algorithms are needed for such networks to achieve the optimal load. These algorithms involve combined changes in two or more dimensions and have large impact on the original routing algorithm.

Since the majority of networks used has K a power of 2, we conclude that the simple algorithm satisfies our needs.

Theorem 3 Let T_K^n be a toroidal network and K be a positive integer, divisible by 4. Let R' be a semi-homogeneous routing algorithm. The message densities, when using R', are the same for all channels and equal to $\frac{1}{8}K^{n+1}$. This is also the load:

$$\forall X \sim Y \in T_K^n : \rho(c_{X,Y}) = \alpha = \frac{1}{8}K^{n+1}.$$

Proof:

First, we notice that the routing between points, that have components that differ less than $\frac{K}{2}$, distributes the load equally among channels. This can be easily proven by the same procedure as in Lemma 1.

Next, we again use the projection. We collapse the network in a cycle and observe only paths of the length $\frac{K}{2}$. It can be seen, that the semi-homogeneous routing algorithm distributes the load, generated with the longest paths, equally among positive and negative directions.

With the combining of the two facts we conclude that the message densities, induced by the semi-homogeneous routing algorithm, are the same for all channels. □

5 Example

After all the theory, let's see how we can apply our theorems. Given a two-dimensional network 4×4 and a homogeneous routing algorithm, we want to have an optimal routing with as little changes as possible.

The routing algorithm is the well-known "approach in one dimension first" algorithm. This means that the routing algorithm is decreasing the distance in one dimension first, after that in another dimension, and so on.

In a two-dimensional network, such algorithm routes messages first horizontally and then vertically towards the destination or vice versa. We assume the first choice. The algorithm is also completely (1- and 2-) positive asymmetric. The message densities are 12 on positive connections and 4 on negative connections. The example of routing is shown on the left side of Figure 3.

Since the original algorithm obviously distributes the load unequally, we have to change it. The changes are only at routing where distance is 2, as suggested in the previous section. The change is very simple and easy to implement. When the distance in dimension i is 2, the algorithm behaves as i-positive, if i-th coordinate of the source is odd, and as i-negative, otherwise. Everything else remains unchanged. Message densities induced by the adjusted routing algorithm are 8 on all channels. On the right side of Figure 3 we can see the routing between the same points as on the left, using the adjusted routing algorithm.

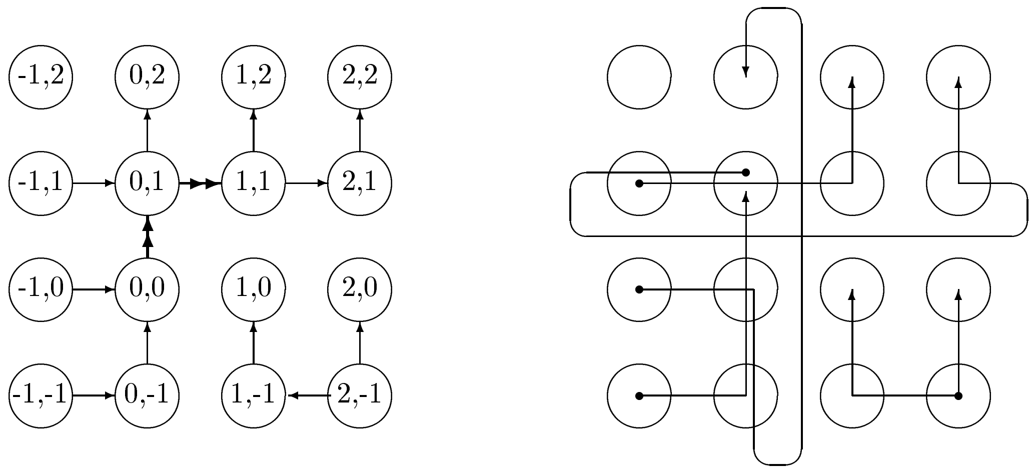

Figure 3: Paths: (-1,-1) → (0,1), (-1,0) → (0,2), (-1,1) → (1,2), (0,1) → (2,2), (2,-1) → (1,0), (2,-1) → (2,0)

6 Conclusion

In the paper, we discussed routing in toroidal networks. These networks are becoming widely used, hence we studied the properties of routing algorithms for such networks.

We introduced the notion of a homogeneous routing algorithm. This actually means that the routing algorithm chooses next direction on the basis of the difference vector between current position and final destination. There is no influence of the original source of the message.

There are two main results, given by the theory in this paper. One is positive, assuring that *any* homogeneous routing algorithm is optimal for odd-sized networks. Other is negative, providing a proof for the claim that *every* homogeneous routing algorithm is non-optimal for even-sized networks.

However, the negative result is alleviated with the Theorem 3, stating that only minor changes are needed to transform a non-optimal homogeneous algorithm to an optimal semi-homogeneous algorithm.

References

[1] H.G. Badr and S. Podar, "An optimal shortest-path routing policy for network computers with regular mesh-connected topologies," IEEE Trans. Comp., vol. C-38, pp.

1362 - 1371, Oct. 1989

[2] W.J. Dally and C.L. Seitz, "Deadlock-free message routing in multiprocessor interconnection networks," IEEE Trans. Comp., vol C-36, pp. 547 - 553, May 1987

[3] M.C. Heydemann, J.C. Meyer and D. Sotteau, "On forwarding indices of networks," Discr. Appl. Math., vol. 23, pp. 103 - 123, 1989

[4] D.H. Linder and J.C. Harden, "An adaptive and fault tolerant wormhole routing strategy for k-ary n-cubes,", IEEE Trans. Comp., vol. C-40, pp. 2 - 12, Jan. 1991

[5] A.J. Martin, "The TORUS: An exercise in constructing a processing surface," Proc. 2nd Caltech conf. VLSI, pp. 527-537, 1981

[6] T.R. Robertazzi, "Toroidal networks," IEEE Comm. Magazine, vol. 26, pp. 45 - 50, June 1988

[7] R. Suaya and G. Birtwistle, "VLSI and parallel computation," Morgan Kaufman Publishers, Inc., 1990

[8] L.D. Wittie, "Communication structures for large networks of microcomputers," IEEE Trans. on Comp., vol C-30, pp. 264-273, Apr. 1981