Circuit-switched Broadcast in Multi-port 2D Tori*

San-Yuan Wang, Yu-Chee Tseng, Sze-Yao Ni, and Jang-Ping Sheu

Department of Computer Science and Information Engineering
National Central University, Chung-Li, 32054, Taiwan

Abstract. This paper studies the *one-to-all broadcast* in a *circuit-switched* 2D torus of any size with α-*port* capability. This is a generalization of the one-port and all-port models. Existing results, as compared to ours, can only solve very restricted sizes of tori, and use more numbers of steps.

1 Introduction

One primary communication in an interconnection network is the *one-to-all broadcast*, where a source node needs to send a message to all other nodes in the network. This paper studies the scheduling of one-to-all broadcast in a 2D circuit-switched torus. The network is assumed to use the α-*port* communication model, in which a node can send up to α messages and simultaneously receive up to α messages at a time, where $1 \leq \alpha \leq 4$. This is a generalization of the *one-port* model ($\alpha = 1$) and the *all-port* model ($\alpha = 2k$). Following the formulation in many works [2, 3, 4, 5], this is achieved by a sequence of *steps*, where a step consists of a set of congestion-free communication paths each indicating a message delivery. The goal is to minimize the total number of steps used.

One-to-all broadcast has been studied for meshes and torus based on different port models and switching models [1, 2, 3, 4, 5, 6]. Some of these results are compared and summarized in Table 1. Our result improves over existing results [1, 2, 4] in both the number of broadcast steps used and its applicability in network sizes.

A 2D torus of size $n_1 \times n_2$ is denoted as $T_{n_1 \times n_2}$. Each node is denoted as p_{x_1,x_2}, $0 \leq x_i < n_i$. The torus is assumed to have α ports, in that a node can send up to α messages along any α of its 4 ports.

We map $T_{n\times n}$ into a modulo Euclidean integer space $\mathbb{Z}^2$, where $\mathbb{Z} = \{0, \ldots, n-1\}$. Let elementary vectors $\boldsymbol{e}_1 = (1,0), \boldsymbol{e}_2 = (0,1), \boldsymbol{e}_{-i} = (-1,0), \boldsymbol{e}_{-i} = (0,-1)$. For simplicity, we may write $\boldsymbol{e}_{1,2} = \boldsymbol{e}_1 + \boldsymbol{e}_2$ and $\boldsymbol{e}_{1,-2} = \boldsymbol{e}_1 - \boldsymbol{e}_2$.

Lemma 1. *In a 2D α-port torus $T_{n_1\times n_2}$, a lower bound on the number of steps to perform one-to-all broadcast is $\lceil \log_{\alpha+1} n_1 n_2 \rceil$.*

Definition 1. *In $\mathbb{Z}^2$, given a node x, an m-tuple of vectors $B = (\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_m)$, and an m-tuple of integers $N = (n_1, n_2, \ldots, n_m)$, we define the* span of x by vectors B and distances N *as a set of nodes* $\mathrm{SPAN}(x, B, N) = \{x + \sum_{i=1}^{m} a_i \boldsymbol{b_i} | 0 \leq a_i < n_i$.

* This work is supported by the National Science Council of the Republic of China under Grant #NSC88-2213-E-008-014 and #NSC88-2213-E-008-027.

Table 1. Comparison on the solvable network sizes and required numbers of steps, assuming a $T_{n_1 \times n_2}$. ($\text{LB}(2)_\alpha$ = the lower bound for α-port 2D tori in Lemma 1)

Algorithm	Ours	Lee-Lee[1]	Park-Choi[2]	Peters[4]
Port Model	α-port	α-port	all-port	all-port
Size	$n_1 \times n_2$	$n_1 = n_2$	$n_1 = n_2 = 5^p$	$n_1 = n_2 = 5^p$ or 2×5^p
Steps	$\begin{cases} \text{LB}(2)_\alpha + 1 & \text{if } n_1 = n_2, \\ \text{LB}(2)_\alpha + 4 & \text{otherwise} \end{cases}$	$\text{LB}(2)_\alpha + 2$	$\text{LB}(2)_4$	$\text{LB}(2)_4$

2 Broadcasting in a Square 2-D Torus

When $\alpha = 4$. The approach is similar to that in [6] except that there is more flexibility as circuit switching is used. There are two stages to achieve broadcast. In the first stage, we recursively spread the broadcast message to the main diagonal of the torus, as shown in Fig. 1(a). In the second stage, we recursively distribute the broadcast message to other diagonals parallel to the main diagonal. This is illustrated in Fig. 1(b).

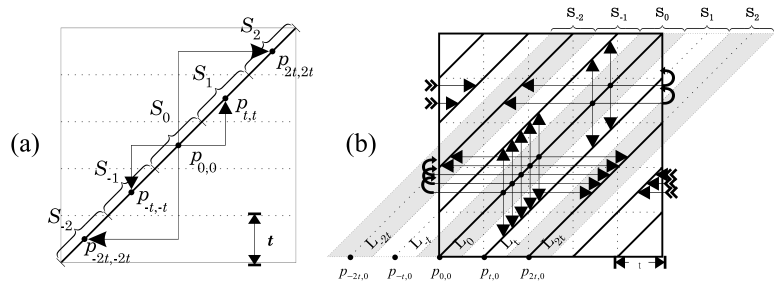

Fig.1. Broadcasting in a square 2-D torus: (a) stage 1 and (b) stage 2.

When $\alpha \leq 3$. The recursion goes is a slower manner. For instance, in stage 2, the number of diagonals getting the broadcast message will be multiplied by $\alpha + 1$ after each step. The modification is straight-forward.

Theorem 1. *In a circuit-switched α-port $T_{n \times n}$ torus, broadcast can be done in $2\lceil \log_{\alpha+1} n \rceil$ steps, which number of steps is at most 1 steps more than the lower-bound Lemma 1.*

3 Broadcasting in a Non-square 2D Torus

When $\alpha = 4$. This case can be solved by modifying the result in [6] for circuit switching. In an all-port $T_{n_1 \times n_2}$ such that $n_1 < n_2$, broadcast can be done within $\lceil \log_5 n_1 \rceil + \lceil \log_5 \frac{n_1}{2} \rceil + \left\lceil \log_5 \frac{n_2}{n_1} \right\rceil + c$ steps, where $c = 1$ (resp., 2) when n_1 is even (resp. odd), which number of steps is at most 3 (resp. 4) more than the lower bound in Lemma 1.

When $\alpha = 3$. We assume that n_1 is even, with the understanding that one more step is sufficient if n_1 is odd. To avoid the tedium of using floor and ceiling functions, we assume that n_2 is a multiple of n_1.

Definition 2. *Given a non-square torus $T_{n_1 \times n_2}$ such that $n_1 < n_2$, the dilated torus induced by $T_{n_1 \times n_2}$, denoted as $\hat{T}_{n_1 \times n_2}$, is an $n_1 \times n_1$ torus consisting of nodes from the following four $\frac{n_1}{2} \times \frac{n_1}{2}$ tori:*

$$T_{0,0} = \mathrm{SPAN}(p_{0,0}, B_2, N_2), \quad T_{1,0} = \mathrm{SPAN}(p_{1,0}, B_2, N_2),$$
$$T_{0,1} = \mathrm{SPAN}(p_{0,\frac{4}{3}\frac{n_2}{n_1}}, B_2, N_2), \; T_{1,1} = \mathrm{SPAN}(p_{1,\frac{4}{3}\frac{n_2}{n_1}}, B_2, N_2),$$

where $B_2 = (2\boldsymbol{e}_1, \frac{2n_2}{n_1}\boldsymbol{e})$ and $N_2 = (\frac{n_1}{2}, \frac{n_1}{2})$. $\hat{T}_{n_1 \times n_2}$ has n_1^2 nodes which are denoted by $\hat{p}_{i,j}$, for $i, j = 0..n_1 - 1$.

Intuitively, the dilated torus in Definition 2 is partitioned into four sub-tori. For instance, Fig. 2(a) shows the four dilated tori in an $n_1 \times n_2$ torus ($n_1 = 6$). Note that we now do not have "straight" diagonals as in the earlier cases.

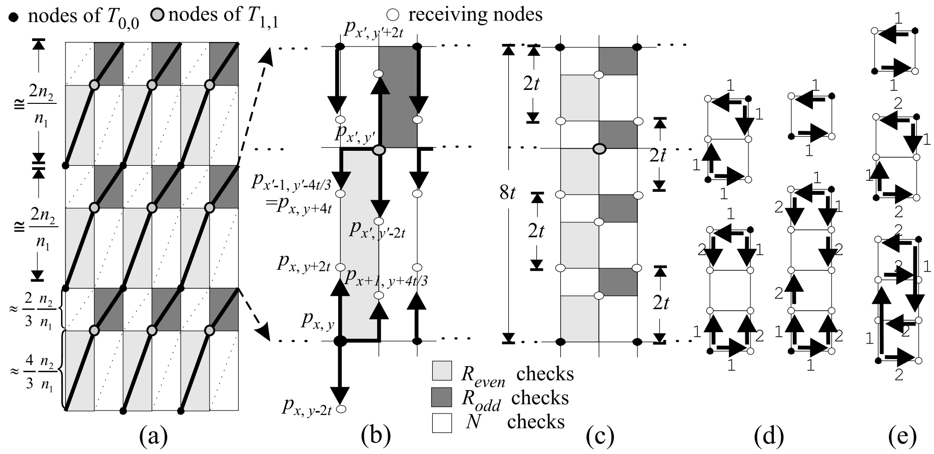

Fig.2. (a) A 3-port $T_{n_1 \times n_2}$ regarded as four dilated $\frac{n_1}{2} \times \frac{n_1}{2}$ tori ($n_1 = 6$). The lines in bold are the alternative diagonals used in Stage 2. (b) The communication pattern in stage 3. (c) The new checkerboard with four smaller rectangles after (b). (d) Stage 4 for R_{even} of heights 1,2,3,and 4. (e) R_{odd} of heights 1, 2, and 3.

Stage 1: Spread M to $\mathrm{SPAN}(\hat{p}_{0,0}, (\boldsymbol{e}_{1,2}), (n_1))$, by applying the stage 1 in Section 2. This takes $\lceil \log_4 n_1 \rceil$ steps.
Stage 2: Spread M to $\frac{n_1}{2}$ diagonals, $\mathrm{SPAN}(\hat{p}_{2i,0}, (\boldsymbol{e}_{1,2}), (n_1)), 0..\frac{n_1}{2} - 1$, by applying the stage 2 of Section 2. This takes $\lceil \log_4 \frac{n_1}{2} \rceil$ steps. Now, nodes of $T_{0,0}$ and $T_{1,1}$ already have M.
Stage 3: We first regard the torus $T_{n_1 \times n_2}$ as a checkerboard and classify checks therein as follows (see Fig. 2(a) for an example).

Definition 3. *[6] In $T_{n_1 \times n_2}$, each smallest submesh in which the lower-left and upper-right corner nodes are the only two nodes that have received the broadcast message is regarded as a check marked by R (received). Excluding the R-marked checks, the rest of the checkerboards are considered as a number of checks marked by N (non-received).*

Definition 4. *A check marked by R is classified as R_{even} if its lower-left node's index along the x-axis is even, and classified as R_{odd} otherwise.*

The recursion should proceed as long as the sum of the heights of two consecutive R_{even} and R_{odd} is ≥ 8. Let's consider the four consecutive checks in Fig. 2(b). For ease of presentation, let the height h of the rectangle be a multiple of eight, $h = 8t$. We perform the following communications:

- for each $p_{x,y}$ located at the lower-left corner of a R_{even}-check, $p_{x,y}$ sends three messages to $p_{x,y+2t}$, $p_{x,y-2t}$, and $p_{x+1,y+\frac{4}{3}t}$, and
- for each $p_{x,y}$ located a the lower-left corner of a R_{odd}-check, $p_{x,y}$ sends three messages to nodes $p_{x,y+2t}$, $p_{x,y-2t}$, and $p_{x-1,y+\frac{4}{3}t}$.

After this step, the rectangle will be partitioned into 4 smaller rectangles as shown in Fig. 2(c). The recursion maintains an important invariant:

I1: The ratio of the height of R_{even}-checks and the height of R_{odd}-checks is (or close to) $2:1$.

The above recursion is repeated until the height of every rectangle is less than 8. As the initial height of the first rectangle is upper-bounded by $\lceil \frac{2n_2}{n_1} \rceil$ and the rectangle height is reduced by a factor of 4 after each recursion, this stage will take $\lceil \log_4 \frac{2n_2}{n_1} \rceil - 1$ steps.

Stage 4: At the end of Stage 3, it is possible to manage the height of each R_{even} and R_{odd} checkd not exceeding 4 and 3, respectively. For each possible height, we show one possible solution in Fig. 2(d) and (e) to send M to nodes in R_{even} and R_{odd} checks. Note how the 3-port model is observed in the communication.

Theorem 2. *In a circuit-switched non-square 3-port $T_{n_1 \times n_2}$ torus such that $n_1 < n_2$, broadcast can be done within $\lceil \log_4 n_1 \rceil + \lceil \log_4 \frac{n_1}{2} \rceil + \left\lceil \log_4 \frac{2n_2}{n_1} \right\rceil + c$ steps, where $c = 1$ (resp. 2) when n_1 is even (resp. odd), which number of steps is at most 4 (resp. 5) steps more than the lower bound in Lemma 1.*

When $\alpha = 1$ and $= 2$. As our scheme follows a dimension-by-dimension approach, a simple recursive doubling/tripling on rows and columns will do the job.

References

[1] S.-K. Lee and J.-Y. Lee. Optimal broadcast in α-port wormhole-routed mesh networks. In *Int'l Conf. on Paral. and Distrib. Sys.*, pages 109–114, 1997.

[2] J. L. Park and H.-A. Choi. Circuit-switched broadcasting in tori and meshes networks. *IEEE Trans. on Paral. and Distrib. Sys.*, 7(2):184–190, Feb. 1996.

[3] J. L. Park et al. Circuit-switched broadcasting in d-dimensional torus and mesh networks. In *Int'l Parallel Processing Symp.*, pages 26–29, 1994.

[4] J. G. Peters and M. Syska. Circuit-switched broadcasting in torus networks. *IEEE Trans. on Paral. and Distrib. Sys.*, 7(3):246–255, March 1996.

[5] Y.-J. Tsai and P. K. McKinley. A broadcasting algorithm for all-port wormhole-routed torus networks. *IEEE Trans. on PDS*, 7(8):876–885, Aug. 1996.

[6] Y.-C. Tseng. A dilated-diagonal-based scheme for broadcast in a wormhole-routed 2d torus. *IEEE Trans. on Comput.*, 46:947–952, Aug. 1997.