



Centrum voor Wiskunde en Informatica
REPORT*RAPPORT*

A user's guide to the Software Testpilot

M. L. Kersten, F. Kwakkel

Computer Science/Department of Algorithmics and Architecture

CS-R9343 1993

Regular Layouts of Butterfly Networks in Three Dimensions

Jörg Keller

CWI

P.O. Box 4079, 1009 AB Amsterdam, The Netherlands

Abstract

Physical arrangements of butterfly networks impose severe problems because of wire length. The problem gets even harder if standard technology like printed circuit boards, racks, and cabinets, must be used. We investigate three-dimensional arrangements of butterfly networks. We construct xu -stage butterfly networks from u -stage networks. The sub-networks can be arranged in x cubes of dimension $x - 1$ such that sub-networks in cubes i and $i + 1$ are connected only if they differ in dimension i . This allows for regular wiring and hence for use of standard components. The proposed method generalizes a result by Wise, which is obtained by choosing $x = 2$.

AMS Subject Classification (1991): 68M07, 68M10

Keywords & Phrases: Butterfly network, Interconnection network, Network layout, Parallel architectures

Note: This work was supported by NWO through NFI project ALADDIN under contract no. NF 62-376.

1. INTRODUCTION

Butterfly networks are often used as interconnection networks of parallel machines [3, 4, 6] because of several advantageous properties. A butterfly network with 2^n inputs and outputs and nodes with indegree and outdegree 2 has depth n , which is optimal for a constant-degree network. There is a unique path between each pair of inputs and outputs, routing decisions only need the binary representation of the output number. This simplifies routing logic and reduces node delay. Furthermore, there exist packet routing algorithms that route n -relations¹ in time $O(n)$ with high probability and require only buffers capable of storing a constant number of packets [5, 6]. This allows to use the same network node design for networks of different size.

However, physical arrangements of butterfly networks reveal some difficulties. While networks as three-dimensional meshes directly lead to arrangements with constant wire length, butterfly networks cannot achieve this. Vitányi proves that the average wire length in a butterfly network cannot be a constant even under relaxed conditions (nodes have unit volume and arbitrary shape, wires have zero volume) [7]. Moreover, straight-forward arrangements lead to wiring that is not regular and therefore not suited for an implementation with standard components.

Wise proposes an arrangement based on cutting a $2u$ -stage butterfly network into two parts [8]. Each part now consists of a number of u -stage subnetworks. Each subnetwork is implemented on one printed circuit board (or board for short). Wise shows that each

¹ n packets are fed into each input, n packets are destined to each output.

subnetwork of the first part is connected to each subnetwork of the second part. He arranges the boards as shown in figure 1. The boards of each part are put in a rack vertically, the racks are put one atop another such that boards of different parts are orthogonal. Now all connections between different boards can be done by short vertical wires or via connectors.

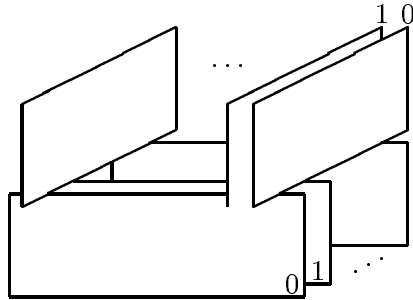


Figure 1: Wise's Arrangement of Boards

Wise's arrangement has the disadvantage that racks with different orientations of boards do not fit into standard cabinets. Vertical wires or connectors between the boards make removal of boards very difficult. Also, as boards are only available up to a size of $50\text{ cm} \times 50\text{ cm}$ and must have a minimum distance of approximately 4 cm from each other to allow cooling, the applicability of this approach is restricted to a small number of stages.

We generalize Wise's approach to overcome these disadvantages. In section 2 we explore the interconnection structure of xu -stage butterfly networks constructed from u -stage butterfly networks. In section 3 we apply our knowledge to derive three arrangements for $x = 2, 3, 4$, respectively. For $x = 2$ we obtain Wise's arrangement.

2. STRUCTURE OF INTERCONNECTIONS

Definition 1 A n -stage (binary) butterfly network is a graph $G = (V, E)$ with $V = \{0, \dots, n-1\} \times \{0, 1\}^{n-1}$. The first index is called stage or row number, the second is called column number. Node (i, v) and (j, w) are connected if and only if $j = i + 1$ and either $v = w$ or v and w differ only in bit i .

Inputs are connected to nodes in stage 0, two inputs per node, outputs are connected to node in stage $n - 1$, two outputs per node.

By a *cut* through a butterfly network G between stages i and $i + 1$ we mean a removal of all edges in G of the form $((i, v), (i + 1, w))$.

By an *arrangement* of a butterfly network G after a cut we mean a graph $G' = (V', E')$, where each node in G' represents a connected component of G after the cut. Nodes in G' are connected if and only if the connected components in G that they represent were connected by an edge that was removed during the cut.

Consider a n -stage butterfly network G , where $n = xu$. If we cut the network between stages $iu - 1$ and iu for $1 \leq i < x$, it falls into x parts.

Lemma 1 *Each part of G consists of $2^{(x-1)u}$ u -stage butterfly networks, hence $V' = \{0, \dots, x-1\} \times \{0, \dots, 2^u - 1\}^{x-1}$.*

Proof: When we cut the n -stage butterfly network in parts with u stages, it is clear from definition 1 that each part consists of u -stage butterfly networks. These networks have 2^{u-1} columns each. As each column in one part is only contained in one u -stage butterfly and as there are 2^{n-1} columns, there must be $2^{n-1}/2^{u-1} = 2^{(x-1)u}$ small butterflies in each part. \square

Theorem 1 *If we arrange G after the cuts described above, then E' consists exactly of all edges of the form*

$$((i, (a_0, \dots, a_{x-2})), (i+1, (a_0, \dots, a_{i-1}, \alpha, a_{i+1}, \dots, a_{x-2}))),$$

for all $\alpha \in \{0, \dots, 2^u - 1\}$.

Proof: Consider the u -stage butterfly networks in part i . We number them according to the smallest column they contain. It is easy to see that butterfly $(b_0, \dots, b_{(x-1)u-1})$ contains columns

$$(b_0, \dots, b_{iu-1}, \beta_0, \dots, \beta_{u-2}, b_{iu}, \dots, b_{(x-1)u-1})$$

for all $(\beta_0, \dots, \beta_{u-2}) \in \{0, 1\}^{u-1}$. Vice versa we can say that column (c_0, \dots, c_{n-2}) is contained in butterfly $(c_0, \dots, c_{iu-1}, c_{(i+1)u-1}, \dots, c_{n-2})$. We denote this by

$$\text{butt}_i(c_0, \dots, c_{n-2}) = (c_0, \dots, c_{iu-1}, c_{(i+1)u-1}, \dots, c_{n-2}) \quad . \quad (2.1)$$

We order the butterflies of a part in a $(x-1)$ -dimensional cube by using u bits as the binary representation of a component:

$$b_{ju}, \dots, b_{(j+1)u-1} = \text{bin}(a_j) \quad , \quad (2.2)$$

where $\text{bin}(a_j)$ denotes the u -bit binary representation of a_j .

Node $x = (x_0, \dots, x_{n-2})$ in the last stage of part i , which is stage $(i+1)u-1$, is connected to nodes $x' = (x_0, \dots, x_{n-2})$ and

$$x'' = (x_0, \dots, x_{(i+1)u-2}, 1 - x_{(i+1)u-1}, x_{(i+1)u}, \dots, x_{n-2})$$

in the first stage of part $i+1$ by definition 1.

Therefore butterfly $X = \text{butt}_i(x)$ in part i is connected to $X' = \text{butt}_{i+1}(x')$ and $X'' = \text{butt}_{i+1}(x'')$ in part $i+1$.

By equation (2.1) we get

$$\begin{aligned} X &= (x_0, \dots, x_{iu-1}, x_{(i+1)u-1}, \dots, x_{n-2}) \\ X' &= (x_0, \dots, x_{(i+1)u-1}, x_{(i+2)u-1}, \dots, x_{n-2}) \\ X'' &= (x_0, \dots, x_{(i+1)u-2}, 1 - x_{(i+1)u-1}, x_{(i+2)u-1}, \dots, x_{n-2}) \end{aligned}$$

As bits $x_{iu}, \dots, x_{(i+1)u-2}$ can be freely chosen in X' and X'' for fixed X and X' and X'' differ also in the next bit, we see that X is connected to all butterflies of the form

$$(x_0, \dots, x_{iu-1}, \gamma_0, \dots, \gamma_{u-1}, x_{(i+2)u-1}, \dots, x_{n-2}) \quad (2.3)$$

for arbitrary $(\gamma_0, \dots, \gamma_{u-1}) \in \{0, 1\}^u$.

We apply the ordering (2.2) to the butterflies of form (2.3) and choose α such that $(\gamma_0, \dots, \gamma_{u-1}) = \text{bin}(\alpha)$. \square

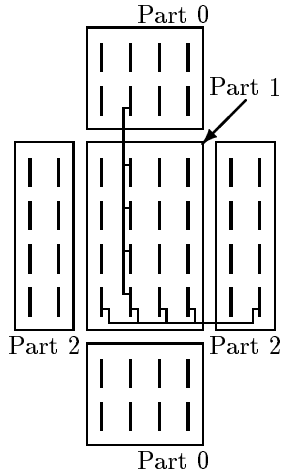


Figure 2: Arrangement of Boards for $u = 6$ and $x = 3$

3. APPLICATIONS

We implement each u -stage butterfly network on a printed circuit board, then Theorem 1 directly leads to several useful applications for $x = 2, 3, 4$.

3.1 Arranging 2u-stage butterfly networks

For $x = 2$, we obtain $V' = \{0, 1\} \times \{0, \dots, 2^u - 1\}$ and $E' = \{((0, i), (1, j)) \mid 0 \leq i, j < 2^u\}$. This is exactly the arrangement proposed by Wise.

3.2 Arranging 3u-stage butterfly networks

For $x = 3$, we obtain $V' = \{0, 1, 2\} \times \{0, \dots, 2^u - 1\}^2$ and

$$E' = \{((0, (i, k)), (1, (j, k))) \mid 0 \leq i, j, k < 2^u\} \\ \cup \{((1, (k, i)), (2, (k, j))) \mid 0 \leq i, j, k < 2^u\}.$$

We arrange the boards of each part in a square, and arrange the squares as shown in figure 2. The squares for parts zero and two are split into two pieces each to make the arrangement symmetric.

All interconnections between boards are horizontal or vertical, the boards can be put into standard racks and cabinets. Two example wirings are shown in the figure. Each board can be removed as soon as the links connected to it are removed. If links are narrow, then they can leave the boards via connectors to the backplane, the wiring can be made on the rear side of the cabinet. If links are wide, then links can leave boards via connectors at the front of the cabinet. The links form wiring channels between the boards.

The arrangement can also be applied if u is not a multiple of three. Then boards in part two consist of two $(u-1)$ -stage or four $(u-2)$ -stage butterfly networks, instead of one u -stage butterfly network.

This arrangement was described for $u = 7$ in [1].

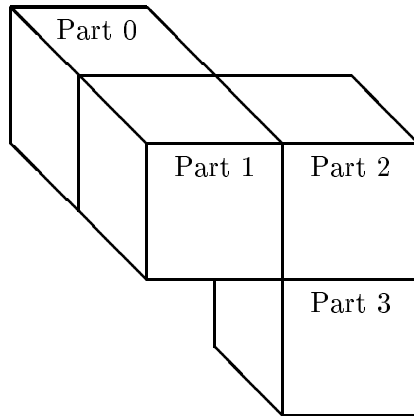


Figure 3: Arrangement of Boards for $x = 4$

3.3 Arranging 4u-stage butterfly networks

For $x = 4$ we arrange the boards of each part as a cube with a length of side 2^u , and arrange the four parts as shown in figure 3. Connections between boards all run orthogonal to the plane separating two cubes.

4. WIRE LENGTHS

We compute the length of wires in arrangements of $3u$ -stage butterfly networks.

A wire is at most as long as the horizontal wire in Fig. 2, crossing the complete part 1 and half of part 2 (or 0 for a vertical wire). We assume that the boards we use have height h_y and are apart by distance h_x . Let further the width of each wiring channel between rows or columns of boards be w . Then the length of a wire is at most

$$l_i = w2^u + h_i(2^u + 2^u/2) \quad (4.4)$$

where $i = y$ for vertical wires and $i = x$ for horizontal wires. The first term represents the 2^u wire channels in part 1, the second term represents the boards the wire has to pass in part 1 (2^u) and in part 0 or 2 ($2^u/2$).

The channel width w can be expressed in terms of the thickness of a single wire w_0 . Each board in part 0 has 2^u output wires, and 2^u boards form a column in part 0. Hence $w = 2^{2u}w_0$.

The boards' height h_y can be given by the number of wires leaving the board, hence $h_y = 2^u w_1$ where w_1 is the length of a connector needed to attach a wire to the board. The boards' distance h_x is a constant w_2 depending on technology.

Unfortunately, for printed circuit boards $h_y \gg h_x$, resulting in different lengths for horizontal and vertical wires. We therefore give the arrangement equal height and width. If $h_y = 4h_x$, then we merge consecutive rows of boards two by two. We reduce the height by a factor two and increase the width by a factor of two, resulting in "pseudo heights" $h'_y = h'_x = h_y/2 = 2h_x$. If in general $h_y = 4^i h_x$ or $\log_4(h_1/h_2) = i$, then we merge $[2^i]$ consecutive rows of boards into one and obtain $h'_y = h'_x = \sqrt{h_y h_x}$ as "pseudo" height and width.

We now can transform Eq. 4.4 to

$$l_i = 2^{3u}w_0 + 3 \cdot 2^{u-1}\sqrt{2^uw_1w_2}.$$

We see that l_i is linear in the number of inputs of the butterfly network, but with a very small constant w_0 . By grouping each wire channel into u layers of $2^{2u}/u$ wires each, the wire length can be reduced to $O(2^{3u}/u)$ which is equivalent to the best 2-dimensional solution which is optimal [2].

5. CONCLUSIONS

Although butterfly networks have no obvious regular implementation with short wires, their structure can be exploited to obtain arrangements with regular wiring that allow for use of standard components. The approach for $x = 3$ is most suited for current technology with racks and cabinets.

REFERENCES

1. F. Abolhassan, R. Drefenstedt, J. Keller, W. J. Paul, and D. Scheerer. On the physical design of PRAMs. In J. Buchmann, H. Ganzinger, and W. J. Paul, editors, *Informatik — Festschrift zum 60. Geburtstag von Günter Hotz*, pages 1–19. Teubner Verlag, 1992.
2. R. Beigel and C. P. Kruskal. Processor Networks and Interconnection Networks without Long Wires. In *Proc. Symp. on Parallel Algorithms and Architectures*, pages 42–51, 1989.
3. *Butterfly Product Overview*, BBN Advanced Computers, Inc., Cambridge, MA, 1987.
4. A. Gottlieb, R. Grishman, C. P. Kruskal, K. McAuliffe, L. Rudolph, and M. Snir. The NYU ultracomputer – designing an MIMD shared memory parallel computer. *IEEE Trans. Comput.*, C-32(2):175–189, Feb. 1983.
5. F. Leighton, B. Maggs, and S. Rao. Universal packet routing algorithms. In *Proc. 29th Symp. on Foundations of Computer Science*, pages 256–269, 1988.
6. A. G. Ranade. How to emulate shared memory. *J. Comput. System Sci.*, 42(3):307–326, 1991.
7. P. M. B. Vitányi. Locality, communication and interconnect length in multicomputers. *SIAM J. Comput.*, 17(4):659–672, Aug. 1988.
8. D. S. Wise. Compact layouts of Banyan/FFT networks. In H. T. Kung, B. Sproull, and G. Steele, editors, *Proc. CMU Conference on VLSI Systems and Computations*, pages 186–195, 1981.