K-WAY BITONIC SORT

Bruce W. Arden
College of Engineering and Applied Science
University of Rochester

Toshio Nakatani
Department of Computer Science
Princeton University

CS-TR-040-86

May, 1986

Index Terms: Bitonic sort, parallel processing, parallel sorting.

K-Way Bitonic Sort

Bruce W. Arden

College of Engineering and Applied Science
University of Rochester
Rochester, N.Y. 14627

Toshio Nakatani

Department of Computer Science
Princeton University
Princeton, N.J. 08544

ABSTRACT

The paper presents k-way bitonic sort, which is the generalization of Batcher's bitonic sort. K-way bitonic sort is based on the k-way decomposition scheme instead of two-way decomposition. We prove that Batcher's bitonic sequence decomposition theorem still holds with multi-way decomposition. This leads to the applications of the sorting network with bitonic sorters of arbitrary or mixed sizes.

May, 1986

K-Way Bitonic Sort

Bruce W. Arden

College of Engineering and Applied Science
University of Rochester
Rochester, N.Y. 14627

Toshio Nakatani

Department of Computer Science
Princeton University
Princeton, N.J. 08544

1. Introduction

Batcher's bitonic sort (Batcher[1968] and Knuth[1973])) has been studied extensively. Stone has described the bitonic sort on the single-stage shuffle-exchange network (Stone[1971]). He has also implemented it on the STAR vector processor (Stone[1978]). Orcutt has implemented the bitonic sort on the ILLIAC-IV (Orcutt[1976]). Thompson and Kung have shown improved time complexity for the bitonic sort on a mesh-connected processor by adopting the shuffle-row-major ordering (Thompson and Kung [1977]). Nassimi and Sahni have also made a different adaptation of the bitonic sort on a mesh-connected processor achieving the same time complexity based on the row-major ordering (Nassimi and Sahni[1979]). Meertens has studied the bitonic sort on the Ultracomputer (Meertens[1979]). Jayanata and Hsiao have used the bitonic sort for a data-base machine design (Jayanata and Hsiao[1979]). Chung, Luccio, and Wong have studied the bitonic sort for magnetic bubble memory systems (Chung, Luccio, and Wong[1980]).

For VLSI implementation, Preparata and Vuillemin have studied the adaptation of the bitonic sort on the cube-connected-cycles network (Preparata and Vuillemin[1981]). Nath, Maheshwari, and Bhatt have adapted the bitonic sort to the orthogonal tree (or mesh-of-trees) (Nath, Maheshwari, and Bhatt[1983]). Bonuccelli and Pagli[1984] have taken a similar approach on the mesh-of-trees for external sorting (Bonuccelli and Pagli[1984]). Bilardi and Preparata have designed an optimal VLSI architecture for the bitonic sort (Bilardi and Preparata[1984]). Thompson has made an extensive survey for parallel sorting including various bitonic sorting schemes in terms of VLSI complexity (Thompson[1983]). Loui has studied the bitonic sort in the context of distributed computing (Loui[1984]). More recently, several people have implemented the bitonic sort using reduced hardware (Ja' Ja' and Owen[1984], Owen and Ja' Ja'[1985], Hsiao and Shen[1985], and Tseng, Hwang, and kumar[1985]).

All of these efforts are based on the two-way decomposition of the original Batcher's bitonic sort. In this paper, we present k-way bitonic sort, which is the generalization of Batcher's bitonic sort. K-way bitonic sort is based on a k-way decomposition scheme instead of two-way decomposition. We prove that Batcher's bitonic sequence decomposition theorem still holds with multi-way decomposition. This leads to the application of the sorting network with bitonic sorters of arbitrary or mixed sizes. In section 2, we describe mathematical notation and definitions. In section 3, we prove the main theorem and several corollaries.

2. Mathematical Notation and Definitions

In this section, notation and definitions are described for later use.

Definition 1: A *sequence* of real numbers is denoted as $\mathbf{a}=\{a_1,a_2, \ldots, a_n\}$. The *length* of a sequence $\mathbf{a}$ is denoted as $|\mathbf{a}|=n$. A *subsequence* of the sequence is a subset of the sequence with preserved order. Especially, a part of the sequence is denoted as $\mathbf{a}_{1,t}=\{a_1,a_2,\ldots a_t\}$ and $\mathbf{a}_{t,n}=\{a_t,a_{t+1}, \ldots, a_n\}$ for $1 \leq t \leq n$. For the subsequence $\mathbf{b}$ of a sequence $\mathbf{a}$, *the complement* of $\mathbf{b}$ in $\mathbf{a}$ is denoted as $\mathbf{a}-\mathbf{b}$. A *concatenation* of the two sequences $\mathbf{a}=\{a_1, \ldots, a_n\}$ and $\mathbf{b}=\{b_1, \ldots, b_m\}$ is denoted as $\mathbf{a}.\mathbf{b}=\{a_1, \ldots, a_n, b_1, \ldots, b_m\}$.

Definition 2: If a sequence is sorted in ascending order, then it is denoted as $\mathbf{a}\uparrow=\{a_1,a_2, \ldots, a_n \mid a_1 \leq a_2 \leq \cdots \leq a_n\}$. If it is sorted in descending order, then it is denoted as $\mathbf{a}\downarrow=\{a_1,a_2, \ldots, a_n \mid a_1 \geq a_2 \geq \cdots \geq a_n\}$. If it is sorted in either ascending or descending order, then it is called *monotonic* and denoted as $\vec{\mathbf{a}}=\{a_1,a_2, \ldots, a_n \mid a_1 \leq a_2 \leq \cdots \leq a_n \text{ or } a_1 \geq a_2 \geq \cdots \geq a_n\}$.

Definition 3: If any element of the sequence $\mathbf{a}$ is no larger than any element of the sequence $\mathbf{b}$, then it is denoted as $\mathbf{a} \leq \mathbf{b}$. Similarly, if any element of the sequence $\mathbf{a}$ is no smaller than the sequence $\mathbf{b}$, then it is denoted as $\mathbf{a} \geq \mathbf{b}$.

Definition 4: A sequence of real numbers $\mathbf{a}$ is *bitonic* if

1) it is a concatination of a monotonically increasing sequence $\mathbf{b}\uparrow$ and a monotonically decreasing sequence $\mathbf{c}\downarrow$, that is $\mathbf{a}=\mathbf{b}\uparrow.\mathbf{c}\downarrow$, where either $\mathbf{b}\uparrow$ or $\mathbf{c}\downarrow$ can be empty: or if

2) the sequence $\mathbf{a}$ can be shifted cyclically so that condition 1 is satisfied.

A bitonic sequence is denoted $\hat{\mathbf{a}}$.

Fact 1: A subsequence of a bitonic sequence is bitonic.

Fact 2: $\mathbf{a}\downarrow.\mathbf{b}\uparrow$ is bitonic.

Fact 3: $\mathbf{a}\downarrow.\mathbf{b}\downarrow$ and $\mathbf{a}\uparrow.\mathbf{b}\uparrow$ are not necessarily bitonic.

Fact 4: If $\mathbf{a} \geq \mathbf{b}$, then $\mathbf{a}\downarrow.\mathbf{b}\downarrow$ is bitonic. Similarly, if $\mathbf{a} \leq \mathbf{b}$, then $\mathbf{a}\uparrow.\mathbf{b}\uparrow$ is bitonic.

Fact 5: If $\mathbf{a} \geq \mathbf{c}$, then $\mathbf{a}\uparrow.\mathbf{b}\downarrow.\mathbf{c}\uparrow$ is bitonic. Similarly, if $\mathbf{a} \leq \mathbf{c}$, then $\mathbf{a}\downarrow.\mathbf{b}\uparrow.\mathbf{c}\downarrow$ is bitonic.

Fact 6: If $\mathbf{a} \geq \mathbf{c}$, then $\vec{\mathbf{a}}.\vec{\mathbf{c}}$ is bitonic. Similarly, if $\mathbf{a} \leq \mathbf{c}$, then $\vec{\mathbf{a}}.\vec{\mathbf{c}}$ is bitonic.

Fact 7: If $\hat{\mathbf{a}} = \mathbf{d}\uparrow.\mathbf{e}\downarrow$ and $\hat{\mathbf{c}} = \mathbf{f}\downarrow.\mathbf{g}\uparrow$ and $\hat{\mathbf{a}} \geq \hat{\mathbf{c}}$, then $\hat{\mathbf{a}}.\hat{\mathbf{c}}$ is bitonic. Similarly, if $\hat{\mathbf{a}} = \mathbf{d}\downarrow.\mathbf{e}\uparrow$ and $\hat{\mathbf{c}} = \mathbf{f}\uparrow.\mathbf{g}\downarrow$ and $\hat{\mathbf{a}} \leq \hat{\mathbf{c}}$, then $\hat{\mathbf{a}}.\hat{\mathbf{c}}$ is bitonic.

Definition 5: Let $\mathbf{a} = \{a_j \mid 1 \leq j \leq N\}$ be a sequence of length $N = kn$. A set of *modulo n subsequences* of $\mathbf{a}$ is a set of n subsequences $\alpha^n = \{\mathbf{a}_1^n, \mathbf{a}_2^n, \ldots, \mathbf{a}_n^n\}$, where $\mathbf{a}_i^n = \{a_j \mid j = i \text{ (modulo } n) \text{ for } 1 \leq j \leq N\}$ is a subsequence of length k for $1 \leq i \leq n$. Let $b_{i,1}$ be the smallest element of $\mathbf{a}_i^n$, $b_{i,2}$ be the second smallest element of $\mathbf{a}_i^n$,..., and $b_{i,k}$ be the largest element of $\mathbf{a}_i^n$ for $1 \leq i \leq n$. Let $\mathbf{b}_j = \{b_{i,j} \mid 1 \leq i \leq n\}$ be a sequence of length n for $1 \leq j \leq k$. The *k-way decomposition* of a sequence $\mathbf{a}$ is to decompose a sequence $\mathbf{a}$ into a set of sequences $\{\mathbf{b}_j \mid 1 \leq j \leq k\}$. As we prove later, if $\mathbf{a}$ is bitonic, then k-way decomposition of a bitonic sequence $\mathbf{a}$, $\{\mathbf{b}_j \mid 1 \leq j \leq k\}$, are bitonic sequences and $\mathbf{b}_1 \leq \mathbf{b}_2 \leq \cdots \leq \mathbf{b}_k$.

Definition 6: A *n-sorter* can sort an arbitrary sequence of length n to a monotonic sequence of length n. A *n-bitonic sorter* can sort any bitonic sequence of length n to a monotonic sequence of length n.

3. Fundamental Theorems for k-Way Bitonic Sort

In this section, we prove the main theorem for k-way bitonic sort which is more general than the one outlined by H. Stone (Stone[1971]) and Batcher (Batcher[1968]) for Batcher's bitonic sort (Batcher[1968]). We start from restating Batcher's bitonic sequence decomposition theorem.

Theorem: (Batcher[1968]) Let $\mathbf{a}$ be a bitonic sequence of length $2n$. Let $\{\mathbf{a}_1^n, \mathbf{a}_2^n, \ldots, \mathbf{a}_n^n\}$ be a set of modulo n subsequences of $\mathbf{a}$, where $\mathbf{a}_i^n$ $(1 \leq i \leq n)$ is a subsequence of length 2. Let b_i be the smallest element of $\mathbf{a}_i^n$ and c_i be the largest element of $\mathbf{a}_i^n$. Let $\mathbf{b} = \{b_i \mid 1 \leq i \leq n\}$ and $\mathbf{c} = \{c_i \mid 1 \leq i \leq n\}$. That is, $\{\mathbf{b}, \mathbf{c}\}$ is the 2-way decomposition of $\mathbf{a}$. Then, both $\mathbf{b}$ and $\mathbf{c}$ are bitonic, and $\mathbf{b} \leq \mathbf{c}$.

Proof: This is a special case of k-way bitonic sort. In other words, Batcher's bitonic sort is 2-way bitonic sort. The proof is immediate from the next main theorem.□

We generalize this theorem to the main theorem for k-way bitonic sort as follows:

Theorem: (k-way bitonic sort) Let $\mathbf{a}$ be a bitonic sequence of length kn. Let $\{\mathbf{a}_1^n, \mathbf{a}_2^n, \ldots, \mathbf{a}_n^n\}$ be a set of modulo n subsequences of $\mathbf{a}$, where $\mathbf{a}_i^n$ $(1 \le i \le n)$ is a subsequence of length k. Let $b_{i,k}$ be the smallest element of $\mathbf{a}_i^n$, $b_{i,2}$ be the second smallest element of $\mathbf{a}_i^n$,..., and $b_{i,l}$ be the largest element of $\mathbf{a}_i^n$. Let $\mathbf{b}_j = \{b_{i,j} \mid 1 \le i \le n\}$ be a sequence of length n for $1 \le j \le k$. That is, for $\mathbf{b}_j = \{b_{i,j} \mid 1 \le i \le n\}$, $\{\mathbf{b}_j \mid 1 \le j \le k\}$ is the k-way decomposition of $\mathbf{a}$. Then, $\mathbf{b}_j$ for $1 \le j \le k$ is bitonic, and $\mathbf{b}_1 \le \mathbf{b}_2 \le \cdots \le \mathbf{b}_k$.

Proof: There are four cases for a bitonic sequence $\mathbf{a}$: 1) $\mathbf{a} = \mathbf{c}\uparrow.\mathbf{d}\downarrow$, 2) $\mathbf{a} = \mathbf{c}\downarrow.\mathbf{d}\uparrow$, 3) $\mathbf{a} = \mathbf{c}\uparrow.\mathbf{d}\downarrow.\mathbf{e}\uparrow$ and $\mathbf{c} \ge \mathbf{e}$, or 4) $\mathbf{a} = \mathbf{c}\downarrow.\mathbf{d}\uparrow.\mathbf{e}\downarrow$ and $\mathbf{c} \le \mathbf{e}$. We use the following notation: $|\mathbf{c}| = p$, $|\mathbf{d}| = q$, $|\mathbf{e}| = r$, and $1 \le t, s, u \le k, p, q, r$. We now look at $\mathbf{b}_1$ more closely:

1) If $\mathbf{a} = \mathbf{c}\uparrow.\mathbf{d}\downarrow$, then $\mathbf{b}_1 = \mathbf{c}_{1,k}$, $\mathbf{d}_{1,k}$, or $\mathbf{c}_{1,t}.\mathbf{d}_{q-t,q}$.

2) If $\mathbf{a} = \mathbf{c}\downarrow.\mathbf{d}\uparrow$, then $\mathbf{b}_1 = \mathbf{c}_{p-k,p}$, $\mathbf{d}_{1,k}$, $\mathbf{d}_{1,s}.\mathbf{c}_{t,p}$, $\mathbf{d}_{s,u}.\mathbf{c}_{t,p}.\mathbf{d}_{1,s-1}$, $\mathbf{c}_{t,p}.\mathbf{d}_{1,s}.\mathbf{c}_{u,t-1}$, or $\mathbf{c}_{t,p}.\mathbf{d}_{1,s}$.

3) If $\mathbf{a} = \mathbf{c}\uparrow.\mathbf{d}\downarrow.\mathbf{e}\uparrow$ and $\mathbf{c} \ge \mathbf{e}$, then $\mathbf{b}_1 = \mathbf{d}_{q-k,q}$, $\mathbf{e}_{1,k}$, $\mathbf{e}_{1,s}.\mathbf{d}_{t,q}$, $\mathbf{e}_{s,u}.\mathbf{d}_{t,q}.\mathbf{d}_{1,s-1}$, $\mathbf{d}_{t,q}.\mathbf{e}_{1,s}.\mathbf{c}_{u,t-1}$, $\mathbf{d}_{t,q}.\mathbf{e}_{1,s}$, or $\mathbf{c}_{1,t}.\mathbf{d}_{s,q}.\mathbf{e}$.

4) If $\mathbf{a} = \mathbf{c}\downarrow.\mathbf{d}\uparrow.\mathbf{e}\downarrow$ and $\mathbf{c} \le \mathbf{e}$, then $\mathbf{b}_1 = \mathbf{c}_{p-k,p}$, $\mathbf{d}_{1,k}$, $\mathbf{d}_{1,s}.\mathbf{c}_{t,p}$, $\mathbf{d}_{s,u}.\mathbf{c}_{t,p}.\mathbf{d}_{1,s-1}$, $\mathbf{c}_{t,p}.\mathbf{d}_{1,s}.\mathbf{c}_{u,t-1}$, $\mathbf{c}_{t,p}.\mathbf{d}_{1,s}$, or $\mathbf{c}.\mathbf{d}_{1,s}.\mathbf{e}_{t,r}$.

In either case, $\mathbf{b}_1$ is bitonic and $\mathbf{b}_1 \le \mathbf{a} - \mathbf{b}_1$. That is, $\mathbf{b}_1 \le \mathbf{b}_j$ for $2 \le j \le k$.

Since $\mathbf{a} - \mathbf{b}_1$ is also bitonic in either case, using the same argument as above repeatedly for $\mathbf{b}_j$ $(2 \le j \le k)$ and $\mathbf{a} - \mathbf{b}_1$, we can prove that $\mathbf{b}_j$ for $1 \le j \le k$

is bitonic, and $\mathbf{b}_1 \leq \mathbf{b}_2 \leq \cdots \leq \mathbf{b}_\mathbf{k}$.□

By the following corollary, a bitonic sequence of length $N=kn$ can be sorted using k-bitonic sorters and n-bitonic sorters:

Corollary 1: A bitonic sequence of length $N=kn$ can be sorted by one stage of n k-bitonic sorters followed by one stage of k n-bitonic sorters (see Figure 3.1 for example).

Proof: Let $\{\mathbf{a}_1^\mathbf{n}, \mathbf{a}_2^\mathbf{n}, \ldots, \mathbf{a}_\mathbf{n}^\mathbf{n}\}$ be a set of modulo n subsequences of $\mathbf{a}$, where $\mathbf{a}_\mathbf{i}^\mathbf{n}$ $(1 \leq i \leq n)$ is a subsequence of length k. Since each subsequence $\mathbf{a}_\mathbf{i}^\mathbf{n}$ is also bitonic, it can be sorted by a k-bitonic sorter at the first stage. For $\mathbf{b}_\mathbf{j} = \{b_{i,j} \mid 1 \leq i \leq n\}$, n k-bitonic sorters produce the k-decomposition, $\{\mathbf{b}_\mathbf{j} \mid 1 \leq j \leq k\}$, of $\mathbf{a}$. Furthermore, $\mathbf{b}_\mathbf{j}$ for $1 \leq j \leq k$ is bitonic, and $\mathbf{b}_1 \leq \mathbf{b}_2 \leq \cdots \leq \mathbf{b}_\mathbf{k}$. Each $\mathbf{b}_\mathbf{j}$ can be sorted by a n-bitonic sorter at the second stage. That is, the concatenation of all the output subsequences is monotonic.□

Corollary 2: An arbitrary sequence of length $N=kn$ can be sorted by $\log k$ stages of smaller bitonic sorters following one stage of k n-sorters: (Stage 0) k n-sorters, (Stage 1) $k/2$ $2n$-bitonic sorters, (Stage 2) $k/4$ $4n$-bitonic sorters,..., (Stage $\log k$) a kn-bitonic sorter, where we assume k is an integral power of two (see Figure 3.2 for example).

Proof: The first stage of k n-sorters produces $k/2$ bitonic sequences of length $2n$, each of which can be sorted to a monotonic sequence of length $2n$ by a $2n$-bitonic sorter (that is, one stage of n 2-bitonic sorters followed by one stage of 2 n bitonic sorters). Two monotonic sequences of length $2n$ are equivalent to a bitonic sequence of length $4n$, which can be sorted to a monotonic sequence of length $4n$ by a $4n$-bitonic sorter (that is, one stage of n 4-bitonic sorters followed by one stage of 4 n-bitonic sorters). Repeat this process for a total of $\log k$ times until we reach a monotonic sequence of length N.□

Corollary 3: An arbitrary sequence of length $N=n^2$ can be sorted by $2logn+1$ stages of n n-sorters (see Figure 3.3 for example).

Proof: For any integer m $(n=lm)$, l m-bitonic sorters can be replaced by an n-sorter. Therefore, for any k ($k \leq n$ and k is an integral power of two), a kn-bitonic sorter can be replaced by two stages of k n-sorters. Thus, the $N=n^2$-sorter in Corollary 2 can be constructed by $2logn+1$ stages of n n-sorters.$\square$

4. Conclusions

In this paper, we proved that Batcher's bitonic sequence decomposition theorem still holds with multi-way decompositions. This leads to the applications of the sorting network with bitonic sorters of arbitrary or mixed sizes.

References

Batcher, K. E. [1968]. "Sorting networks and their applications," *1968 Spring Joint Computer Conf., AFIPS Proc.,* vol. 32. Washngton, D.C., pp. 307-314.

Bilardi, G. and F. P. Preparata [1984]. "An architecture for bitonic Sorting with optimal VLSI performance," *IEEE Trans. on Computers* C-33:7, pp. 646-651.

Bonuccelli M. A. and L. Pagli [1984]. "External sorting in VLSI," *IEEE Trans. on Computers* C-33:10, pp. 931-934.

Chung K. M., F. Luccio, C. K. Wong [1980]. "On the complexity of sorting in magnetic bubble memory systems," *IEEE Trans. on Computers* C-29:7, pp. 553-563.

Hsiao C.C. and N. Shen [1985]. "k-Fold bitonic sort on a mesh-connected parallel computer," *Information Processing News Letters* 21:10, pp. 207-212.

Ja' Ja' J. and R. M. Owens [1984]. "VLSI sorting with reduced hardware," *IEEE Trans. on Computers* C-33:7, pp. 668-671.

Jayanata B. and D. K. Hsiao [1979]. "Parallel bitonic record sort - an effective algorithm for the realization of a post processor," Technical Report OSU-CISRC-TR-79-1, Ohio State Univeristy Columbus Computer and Information Science Research Center.

Knuth, D. E. [1973]. *The Art of Computer Programming, Vol 3: Sorting and Searching.* Reading, M.A., Addison-Wesley.

Loui M. C. [1984]. "The complexity of sorting on distributed systems," *Information and Control* 60, pp. 70-85.

Meertens L. G. L. T. [1979]. "Bitonic sort on ultracomputers," Technical Report MC-IW-117-79, Mathematisch Centrum, Amsterdam, Netherlands.

Nassimi D. and S. Sahni [1979]. "Bitonic sort on a mesh-connected parallel computer," *IEEE Trans. on Computers* C-27:1, pp. 2-7.

Nath D., S. N. Maheshwari, and P. C. P. Bhatt [1983]. "Efficient VLSI networks for parallel processing based on orthogonal trees," *IEEE Trans. on Computers* C-32:6, pp. 569-581.

Orcutt, S. E. [1976]. "Implementation of permutation functions in ILLIAC IV-type computers," *IEEE Trans. on Computers* C-25:9, pp. 929-936.

Owens R. M. and J. Ja' Ja' [1985]. "Parallel sorting with serial memories," *IEEE Trans. on Computers* C-34:4, pp. 379-383.

Preparata F. P. and J. Vuillemin [1981]. "The cube-connected cycles: a versatile network for parallel computation," *Comm. of ACM* 24:5, pp. 300-309.

Stone, H. S. [1971]. "Parallel processing with the perfect shuffle," *IEEE Trans. on Computers* C-20:2, pp. 153-161.

Stone, H. S. [1978]. "Sorting on STAR," *IEEE Trans. on Software Engineering* SE-4:3, pp. 138-146.

Thompson, C.D., H. T. Kung [1977]. "Sorting on a Mesh-connected Parallel Computers," *Comm. ACM* 20:4, pp. 263-271.

Thompson, C.D. [1983]. "The VLSI comlexity of sorting," *IEEE Trans. on Computers* C-32:12, pp. 1171-1184.

Tseng P. S., K. Hwang, and Prasanna Kumar V. K. [1985]. "A VLSI-based multiprocessor architecture for implementing parallel algorithms," *Proc. of IEEE Internatinal Conference on Parallel Processing,* pp. 657-664.

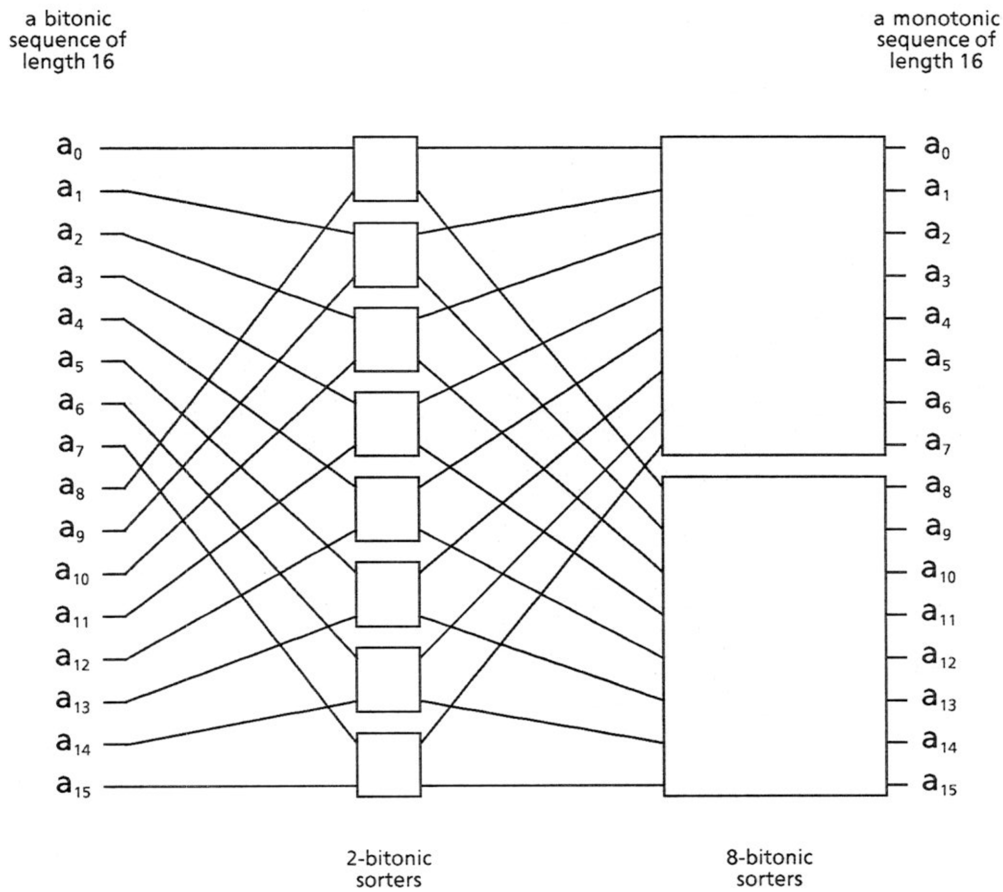

Figure 3.1a: A 16-bitonic sorter by 2-way decomposition (Batcher's construction)

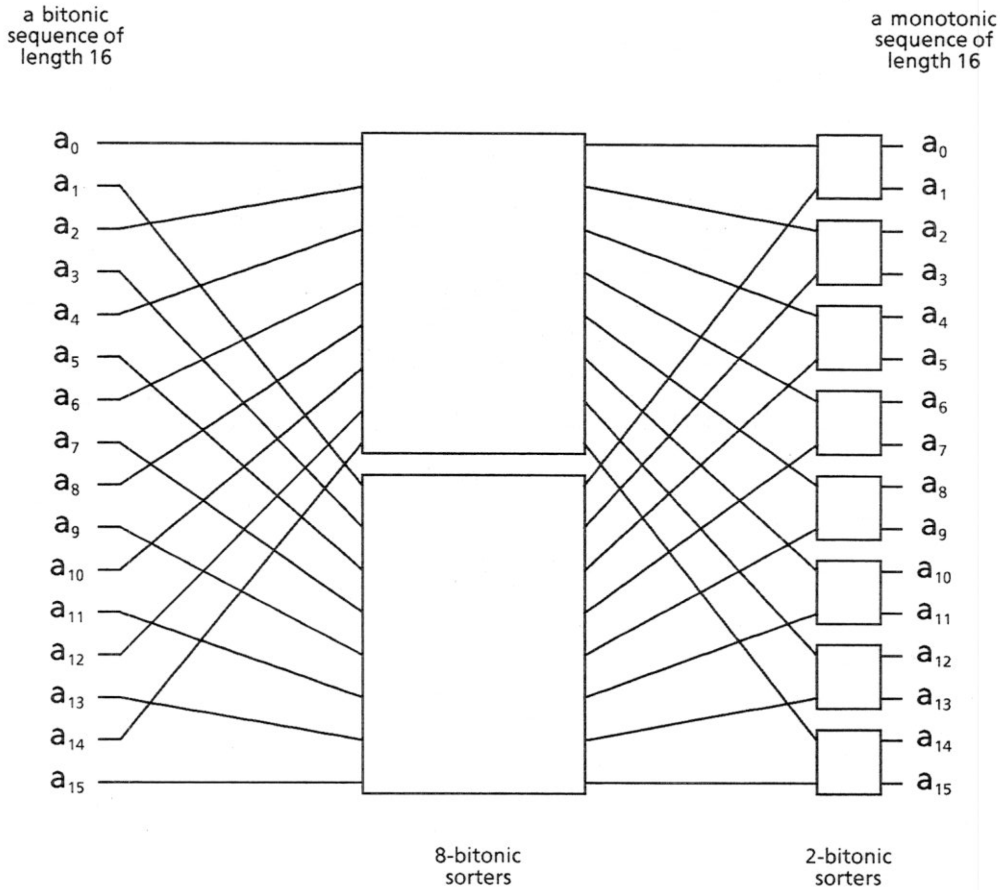

Figure 3.1b: A 16-bitonic sorter by 8-way decomposition

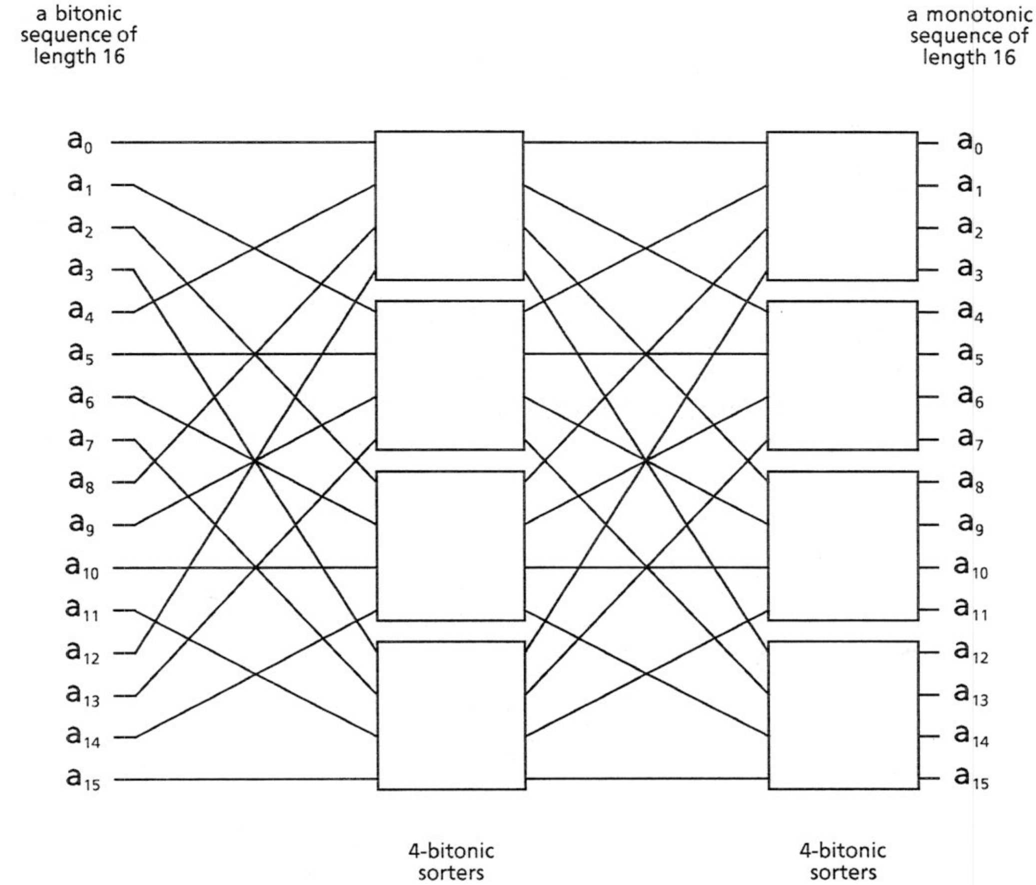

Figure 3.1c: A 16-bitonic sorter by 4-way decomposition

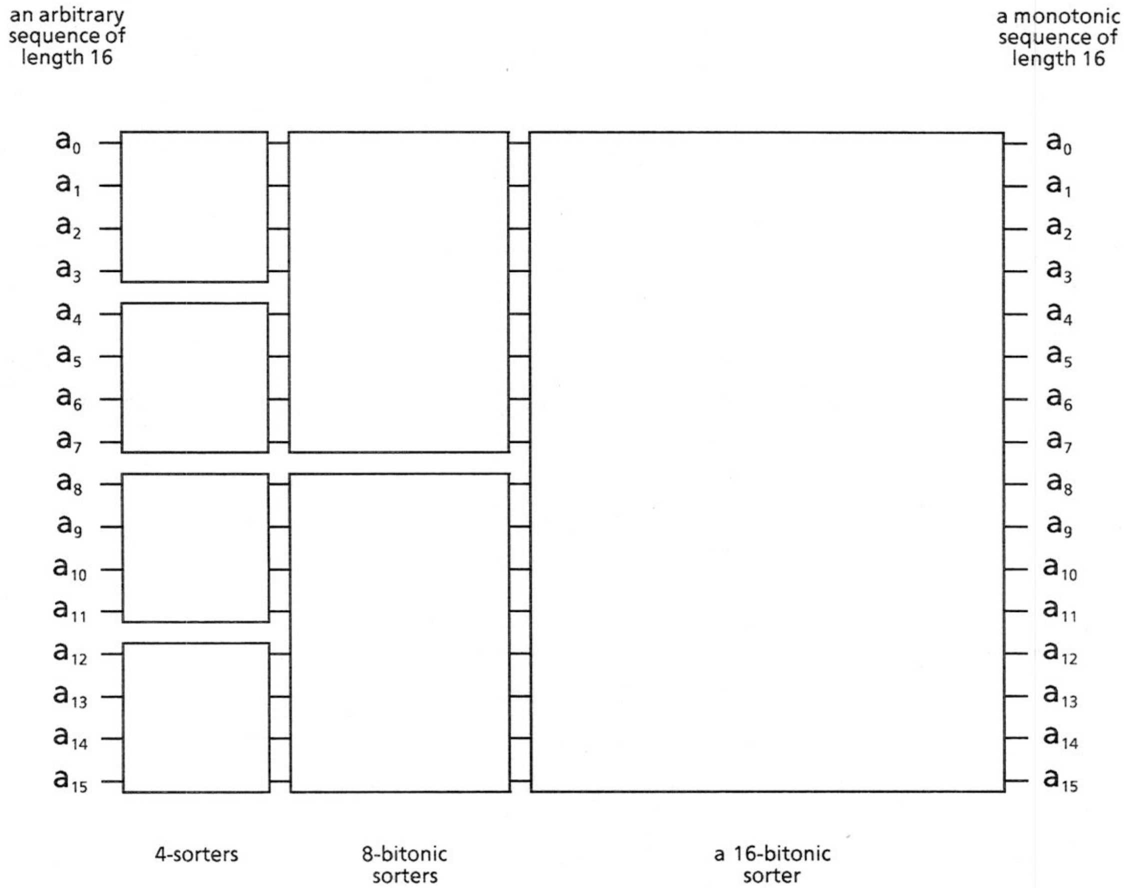

Figure 3.2: A 16-sorter constructed from 4-sorter, 8-bitonic and 16-bitonic sorters

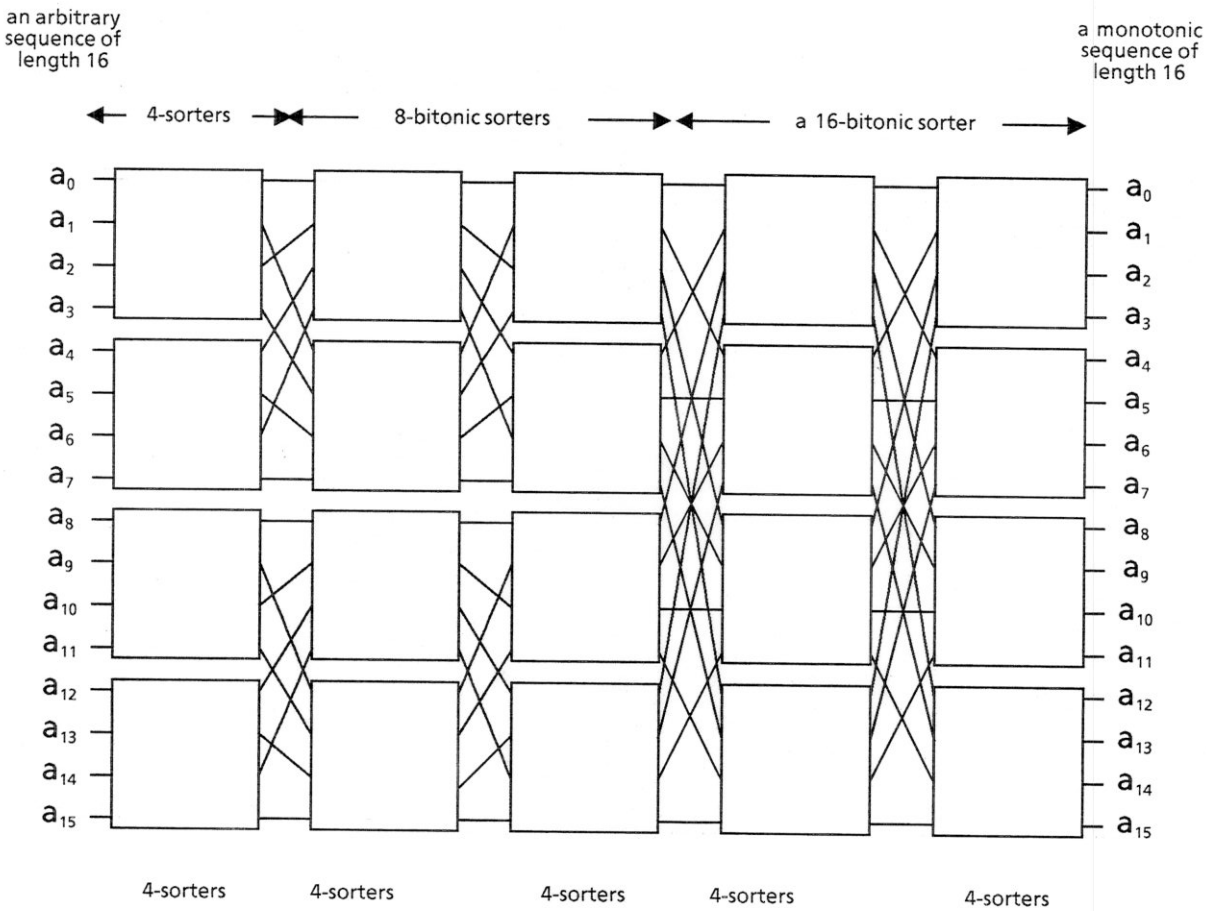

Figure 3.3: A16-sorter constructed from 4-sorters