$O(logN)$ Depth Routing Structure Based on truncated Concentrators

Lee, Jong-Keuk*
*Dept. of Computer Engineering, Dongeui University

잘림구조 집중기에 기초한 $O(logN)$ 깊이의 라우팅 구조

이 종극*
*동의대학교 컴퓨터공학과

ABSTRACT

One major limitation on the efficiency of parallel computer designs has been the prohibitively high cost of parallel communication between processors and memories. Linear order concentrators can be used to build theoretically optimal interconnection schemes. Current designs call for building superconcentrators from concentrators, then using these to recursively partition the connection streams $O(log_2N)$ times to achieve point-to-point routing. Since the superconcentrators each have $O(N)$ hardware complexity but $O(log_2N)$ depth, the resulting networks are optimal in hardware, but they are of $O(log_2N)$ depth. This depth is not better than the $O(log_2N)$ depth Bitonic sorting network, which can be implemented on the $O(N)$ shuffle-exchange network with message passing. This paper introduces a new method of constructing networks using linear order concentrators and expanders, which can be used to build interconnection networks with $O(log_2N)$ depth as well as $O(Nlog_2N)$ hardware cost.(All logarithms are in base 2 throughout paper).

1. Introduction

In order to route N streams of information efficiently in a parallel computer, it is necessary to construct a network with N disjoint paths from source to destination. Since parallel computing can employ a wide range of parallel algorithms and data structures, the most powerful interconnection scheme is one that can accommodate arbitrary source-destination pairings for all N information streams. One way to build such an interconnection is to use a superconcentrator to divided the input stream into two output parts, then recursively divide each part of the output with two additional superconcentrators, and so on until each steam has been connected to its specific destination[2,3]. Leighton[5] showed that a simple, fast routing algorithm can be used to compute the routing of the links through a concentrator-based splitter network. Therefore the remaining task is to determine the exact structure of a practical sized concentrator.

To develop a $O(logN)$ depth, $O(NlogN)$ hardware complexity network, one proposed solution is the AKS sorting network[1]. This network uses expanders as pathways for comparison-exchanges in a sequence of repeated approximate halving operations. The algorithm terminates in *O(logN)* time, though with a very large constant. Another approach has been proposed by Upfal[9], which terminates in *O(logN)*, but it relies on packet routing, and therefore doesn't allow fixed connections to be established from inputs to outputs. In this paper we propose a different method of using concentrators to achieve *O(logN)* depth point-to-point routing on an interconnection network with *O(NlogN)* hardware complexity. It doesn't use packet routing, so fixed connections are established from inputs to outputs.

2. Structure of Superconcentrator

First we define a concentrator and show how to build a superconcentrator with it. An (N, θ, k) concentrator is a two-stage connection network, with N inputs, θN outputs, at most kN links from the inputs to the outputs, having the property that, for every set of inputs X such that $|X| \leq N/2$, all inputs in the set X can be one-to-one connected to the outputs[6]. Since $\theta < 1$, this property guarantees that a stream of at most N/2 active inputs can be connected to the output stream along disjoint paths, while $(1-\theta)N$ of the unused inputs are disconnected from the θN outputs[8].

In order to construct a superconcentrator from this structure following Pippenger[4] we build a network with N inputs and an outputs, with a direct connection from each input to a corresponding output. In order to superconcentrate a set of inputs I to a set of outputs O where $|I|=|O|$, connect any inputs in I to any output in O that happens to be linked by the direct connection. If $|I| > N/2$, then at most N/2 of these inputs will fail to link using the direct connection. These are then passed through an (N, θ, k) concentrator, while on the output side a mirror image structure feeds the outputs. Between these two structures, a recursion of the entire superconcentrator structure is implemented, but with θN inputs and θN outputs. The total hardware cost S(N) of this structure, in terms of the number of links, is given by:

$$S(N) = N + 2kN + S(\theta N) \qquad (1)$$

or, after solving the recursion (ignoring the minor impact of restrictions on the number of inputs to the concentrators):

$$\frac{S(N)}{N} = \frac{2k+1}{1-\theta} \qquad (2)$$

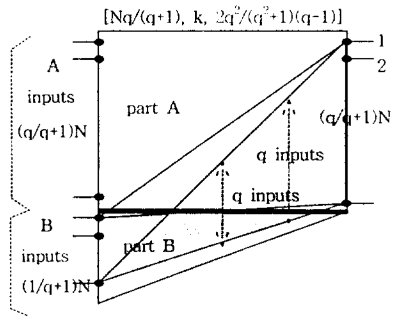

Fig. 1. The concentrator composed of two different sized parts.

The concentrator built from this expander is the union of two parts, called Part A and Part B, as shown in figure 1. Part A is an $(\lceil \frac{Nq}{q+1} \rceil, k, d)$ expander, Part B has $\lfloor \frac{N}{q+1} \rfloor$ inputs, with each input connected to q disjoint sets of the $\lceil \frac{Nq}{q+1} \rceil$ outputs of the expander, as shown in figure 1. N' is chosen so that:

$$N \geq \lceil \frac{Nq}{q+1} \rceil + \lfloor \frac{N}{q+1} \rfloor \qquad (4)$$

As N becomes large, N'-N becomes small, so we will assume $N' \approx N$ (Gabber and Galil[7] show how to use this fact to improve the size of the superconcentrator slightly). For a specific range of values of the concentration coefficient q, as given in equation 5 below, this structure is an (N, $\frac{q}{q+1}$, k) concentrator. The concentration property is guaranteed by the following version of Hall's Matching Theorem[10].

3. Design of New Concentrator from Expander

Any superconcentrator constructed from concentrators must have O(logN) depth. Thus, any interconnection network constructed from these superconcentrators, which uses a process of recursively subdividing the input streams, until the destinations are reached, must use O(logN) successive superconcentrators. This means that the interconnection network will have depth O(logN). We show in this section that it is possible to construct linear order concentrator structures with very large ratios of inputs to outputs, that is, which can concentrate an input stream into a much smaller output stream, in one stage. In section Ⅳ. we will use this type of structure to build interconnection networks of O(logN) depth and O(N logN) hardware complexity for point-to-point (permutation) routing. We will also show how to construct linear order expander type network with arbitrarily large expansion coefficient d.

As described previously, this structure guarantees that any input set I can be connected to any output set O. Now after D recursive stages of superconcentration within this superconcentrator, there can be at most ω connections from any I to any O which have not already been achieved, where ω is given by the formula:

$$\omega = \left(\frac{N}{2} \cdot \frac{q}{q+1}\right)^{D-1} \qquad (6)$$

where q is defined by the type of expander used to construct the concentrator. By choosing D to be sufficiently large, ω can be made arbitrarily small. Nevertheless, as long as D is bounded, the resulting structure will have bounded depth.

Now if we truncate the superconcentration process after D steps, then for each input set I, and each output set O, at most ω elements of I will fail to connect to elements of O. Instead, these ω elements have connections to at least ω outputs not in O. Therefore the total expansion resulting from this process is given by:

$$\Gamma_I \geq \frac{N}{|I|} \times (|I| - \omega) + \omega \qquad (7)$$

This formula applies as long as $|I| \geq \omega$, otherwise $\Gamma_I = |I|$, and no expansion occurs. This can be remedied by connecting each of the $N(q/q+1)D$ links after the stage D recursion to $(q+1/q)D$ outputs, as illustrated in figure 2, resulting in an expansion given by:

$$\Gamma_I = \max\left[\frac{N}{|I|} \times (|I| - \omega), \left(\frac{q+1}{q}\right)^{D} \times \omega\right] \qquad (8)$$

if $|I| \geq \omega$, or if $|I| < \omega$:

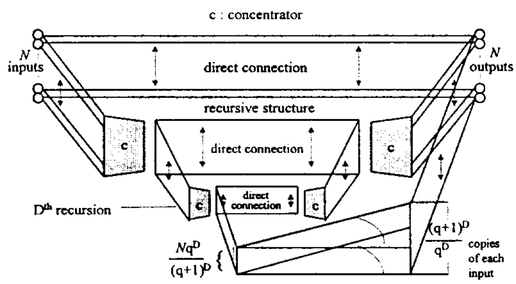

Fig. 2. A truncated superconcentrator with (q+1/q)D direct connections to the outputs after D recursion.

$$\Gamma_I = \left(\frac{q+1}{q}\right)^{D} \times |I| \qquad (9)$$

Now the above structure functions as an expander, but it is different in that it is a multistage network, unlike the bipartite expander graph. Nevertheless, it can be used to construct a bipartite expander graph, by replacing each of the paths from each input node to each output node by a direct link. Since the above network is of bounded depth and each node is of bounded outdegree, the outdegree of the input nodes of the resulting expander graph will also be bounded. If the original outdegree of the nodes in the concentration stages of the multistage network is k, then the outdegree of the nodes in the mirror image reverse concentrate stages is $k(q+1)/q$, and with D stages each input of the resulting bipartite expander is connected to K outputs where:

$$K=(k+1)^{2D}\times\left(\frac{q+1}{q}\right)^{D} \qquad (10)$$

While this structure functions as an expander with outdegree K and expansion given by equations 8 and 9, its expansion formula is different than for the usual expanders, and it has the added property that the formula can be applied for all values of $|I|$, that is for the range $0\leq|I|\leq N$. Because it is derived from a truncated superconcentrator, we will refer to it as a truncated superexpander.

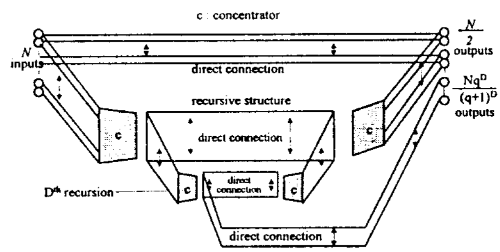

Fig. 3. A halving concentrator built from a truncated superconcentrator with single direct connections to the outputs after D recursions.

A second structure which may be useful in designing interconnection networks is a type of concentrator derived from the above truncated superconcentrator. Note that by truncating the superconcentration process after D recursion, we guarantee that for any choice of input set I and output set O all but ω links will be completed. Now consider a network which has the above structure, except that only $N/2$ of the outputs are to be used in the set O. That is, the network directs the input set I to only half of the outputs. We then add to this set an additional $N(q/q+1)D$ outputs, which we connect, one-to-one, with the $N(q/q+1)D$ outputs of the D_{th} concentration recursion as illustrated in figure 3. This structure concentrates N inputs to $N/2+N(q/q+1)D$ outputs. Clearly by making D sufficiently large the number of outputs can be made to approach $N/2$ arbitrarily.

Again, as before, the multistage structure can be replaced with a single stage bipartite network, resulting in an outdegree similar to K in equation 11 for the above truncated superexpander. Since each input is guaranteed to either superconcentrate, or appear in the $N(q/q+1)D$ extra outputs containing at most ω connections, it must function as a concentrator. Because it is derived from the truncated superconcentrator, we will refer to it as a truncated concentrator.

4. Construction of Interconnection Network

Now we will show how to build an interconnection network with $O(logN)$ stages, using the truncated concentrator. We employ this structure as a halving network, with two exact copies each designed to handle half of the connections, those connecting to the outputs i in the range $1\leq i\leq N/2$, and those connecting to the outputs j in the range $N/2<j\leq N$. Because of the $N(q/q+1)D$ extra outputs appearing after each recursion, after R recursions of the halving process, the streams have increased in size from $N/2R$ in each of the $2R$ substreams to a value S given by:

$$S=N\frac{((q+1)^{D}+2q^{D})^{R}}{2(q+1)^{D}} \qquad (11)$$

This presents no special problem for the truncated concentrator, since it can easily be modified to remove at each stage all but $(N/2R)(q/q+1)D$ extra outputs at the R_{th} recursion, at no extra cost. Since the truncated concentrator is derived from the superconcentrator, as the inputs are reduced the size of the output set can be equally reduced, plus the $(N/2R)(q/q+1)D$ extra outputs. Thus at each recursion following the first stage, the halving networks are removing a little more than half the streams. Figure 3 illustrates the entire process. The above process terminates after $logN$ stages.

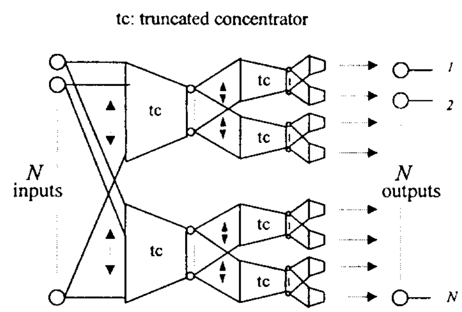

Fig. 3. An interconnection network built from truncated concentrators with successive halving.

5. Conclusions

We have shown that the truncated concentrators in particular can be used to construct a $O(logN)$ depth interconnection network, with a total hardware complexity of $O(KNlogN)$. This network creates fixed connections from the N inputs to the N outputs, it does not use packet routing, allowing repeated use of the same connection paths as long as connection requests do not change. Because K is a large constant, the complexity is not within the range of practicality, but it provides a basis for further study. There is no clear reason to believe that K cannot be reduced to a practical level.

References

1. M.Ajtai, J.Komlos, and E.Szemeredi "Sorting in c log n Parallel Steps", *Combinatorica*, Vol. 3, no. 1, 1983, pp. 1-19.

2. C.E.Leiserson, "Fat-Trees: Universal Networks for Hardware-Efficient Supercomputing", *IEEE Transactions on Computers*, Vol. C-34, No. 10, October 1985, pp. 892-901.

3 D.M.Koppelman, and A.Y.Oruc "A Self-Routing Permutation Network", *Proceedings, 1989 International Conference on Parallel Processing*, August 1989.

4. N.Pippenger "Superconcentrators", *SIAM Journal of Computing*, Vol. 6, no. 2, June 1977, pp. 298-304.

5. F.T.Leighton and B.M.Maggs "Fast Algorithms for Routing Around Faults in Multibutterflies and Randomly-Wired Splitter Networks", *IEEE Transactions on Computers*, Vol. C-41, No. 5, May 1992, pp. 578-587.

6. G.A.Margulis, "Explicit Construction of Concentrators", *Problems of Information Transmission*, 1973.

7. O.Gabber and S.Galil "Explicit Constructions of Linear Size Superconcentrators", *Proceedings of the Twentieth Annual Symposium on the Foundations of Computer Science*, 1979, pp. 364-370.

8. N.Alon, Z.Galil and V.D.Milman, "Better Expanders and Superconcentrators", *Journal of Algorithms*, Vol. 8, 1987, pp.337-347.

9. E.Upfal, "An O(logN) Deterministic Packet Routing Scheme", *Twenty-first Annual ACM Symposium on the Theory of Computing*, May 1989, pp. 241-250.

10. P.Hall, "On Representatives of Subsets" *Journal of the London Mathematical Society*, Vol. 10, 1935, pp. 26-30.