Comparing the Performance Parameters of Two Network Structures for Scalable Massively Parallel Processors

Ding-Ming Kwai and Behrooz Parhami

Department of Electrical and Computer Engineering
University of California
Santa Barbara, CA 93106-9560, USA

Abstract

Packed exponential connection (PEC) networks and periodically regular chordal (PRC) rings, proposed as enhanced versions of linear arrays and rings respectively, introduce skip links into the respective base networks. Both networks can be viewed as derivatives of symmetric chordal rings with restricted or reduced connection assignments. Thus, an interesting question arises: How can one prune or thin the connections of a given network to a prescribed degree such that the key network properties are still acceptable? We show the hierarchical arrangements and grouping strategies on which PEC networks and PRC rings are based and examine their effects on expandability, ease of implementation, network diameter, and routing schemes.

1 Introduction

Networks with nearest-neighbor connections, such as linear arrays and rings, have an area complexity of O(N) with N nodes. Although their implementation is very compact, lack of long-distance connections adversely affects the communication latency and network reliability. One common approach to improving the performance is to increase the connectivity by introducing a fixed number of *skip* links or *express* channels into the networks. Based on this multiple fixed-skip construction, different augmented structures have been proposed to secure desired properties of small diameter and robustness. Among these, topologies derived from unidirectional and bidirectional rings have proven quite effective in view of routing simplicity.

Homogeneous networks in which all nodes have the same degree and uniform connection patterns offer several advantages over arbitrary networks with respect to task scheduling and data routing. In fact, an important reason for including end-around links that connect boundary nodes in some architectures, such as augmented data manipulator networks [13] and continuously wrapping hexagonal meshes [2], is to introduce connection uniformity. On the other hand, homogeneity may lead to certain undesirable properties. For example, in some symmetric networks, such as tori and chordal rings, one either has to cope with relatively large diameters or be willing to pay the cost of complex nodes whose degree increases with network size. A natural question, then, is: How can one do away with homogeneity and still maintain desirable properties of a given class of networks?

Several attempts have been made to answer the above question [4, 10, 13, 14, 17]. We compare two such architectures: Packed exponential connection (PEC) networks and periodically regular chordal (PRC) rings. The former has been proposed as an enhanced bidirectional linear array [10] while the latter is based on a unidirectional ring [13]. These two architectures are similar in node degree, use of skip connections, and derivability from symmetric chordal ring through the removal of some links. Hence an in-depth comparison of these "locally heterogeneous" architectures may lead to inferences about the best strategies for "pruning" edges from a homogeneous network without significantly altering its key topological parameters (e.g., diameter).

2 Network definitions

We assume that $n-1$ types of auxiliary connections are introduced into a unidirectional or bidirectional ring. Let $R(N, s_1, s_2, \ldots, s_{n-1})$ denote a unidirectional symmetric chordal ring with N nodes indexed from 0 to $N-1$ and $n-1$ skip links per node such that node v is connected to $n-1$ remote nodes $v+s_1, v+s_2, \ldots, v+s_{n-1}$, in addition to its near neighbor $v+1$. $R(N, \pm s_1, \pm s_2, \ldots, \pm s_{n-1})$ is the bidirectional version where node v is connected to nodes $v \pm 1, v \pm s_1, v \pm s_2, \ldots, v \pm s_{n-1}$. For notational convenience, we define the local or nearest-neighbor connection as skip distance $s_0 = 1$ and also introduce $s_n = N$. Note that the node degree for bidirectional chordal rings is even except in the special case where $N/2$ constitutes one of the skips.

The symmetric chordal ring is homogeneous in that all nodes are topologically identical. Fig. 1 shows an 8-node bidirectional symmetric chordal ring with node degree 5 and skips $\pm s_0 = \pm 1$, $\pm s_1 = \pm 2$, and $s_2 = 4$. For these networks (and their generalized versions with $s_0 \neq 1$), Wong and Coppersmith [15] have shown that the diameter is bounded by a function of $nN^{1/n}$. Hence, the diameter can be reduced by increasing the node degree n. See also the survey by Bermond, Comellas, and Du [1].

0-8186-7235-8/96 $5.00 © 1996 IEEE
Proceedings of MASCOTS '96

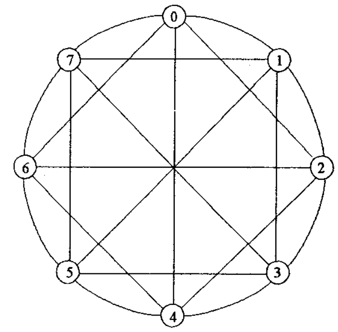

Fig. 1. An 8-node bidirectional symmetric chordal ring $R(8, \pm 2, 4)$.

2.1 PEC Networks

PEC networks use skip distances $s_h = 2^h$, $0 \le h \le n-1$, incorporating all possible skips (i.e., $n = \log_2 N$). The skip distance 2^h is assigned to a node v if $v = i2^h + 2^{h-1}$ for some non-negative integer i. Each node v has at most four links connecting to nodes $v+1$, $v-1$, $v+2^h$ and $v-2^h$, if these nodes exist. All odd-numbered nodes (having IDs of the form $2i+1$) are assigned skip distance of 2. In general, a node is assigned skip distance of 2^h if its ID ends with exactly $h-1$ zeros.

2.2 PRC Rings

In PRC rings, the $n-1$ skip distances of a symmetric chordal ring are distributed among a sequence of nodes, each being assigned only one skip distance. This is similar to the derivation of cube-connected cycles (CCC) architecture [14] from a hypercube. The N nodes in a PRC ring are split into N/g groups of g consecutive nodes, where $g = n-1$ divides N. Each group then possesses a complete sequence of skip distances from s_g to s_1. The skip distance s_h is assigned to a node v if $v = ig + (g-h)$, $0 \le i \le N/g - 1$ and $1 \le h \le g$. Node v has two outgoing links connecting to nodes $v+1$ and $v+s_h$ and two incoming links originating from nodes $v-1$ and $v-s_h$.

By varying these parameters, one can obtain architectures with widely different characteristics (e.g., from polynomial to logarithmic diameter). Assuming that s_{h+1} is divisible by s_h, the network diameter of a PRC ring is $\sum_{h=0}^{g} s_{h+1}/s_h - 3$ [13]. Recall that $s_0 = 1$ and $s_{g+1} = s_n = N$. Given a particular value for g, the diameter is minimized for $s_{h+1}/s_h = N^{1/(g+1)}$. This leads to the optimal diameter $d = (g+1)N^{1/(g+1)} - 3$. Selecting $g = (\log_2 N)/2 - 1$ minimizes the diameter to $2\log_2 N - 3$.

3 Topological Properties

3.1 Node Degree

The node degree of both PEC network and PRC ring is a small constant. The features of the two structures, however, diverge due to important differences in the original networks on which they are based. A linear array is irregular in that its node degree is not uniform. In a PEC network, the node degree is 1, 2, 3, or 4. Consequently, if nodes have identical designs, as is normally the case in practice, some communication ports will remain unconnected. This apparent waste of resources may lead to better expandability (i.e., constructing a large network from smaller building blocks).

A ring, on the other hand, is both node- and edge-symmetric. In a PRC ring, circular symmetry is relaxed from a node to a group of nodes. Node degree (in-degree plus out-degree) is uniformly equal to 4. But, symmetry is in part sacrificed for realizability and scalability.

We assume, for simplicity, that the PRC ring is unidirectional. We will see that our results do not change appreciably for bidirectional PRC rings. For an actual implementation, it may be advantageous to make PRC rings possess bidirectional links. The most important reason is to allow us to exploit locality of communication which generally involves frequent passing of messages between near neighbors in both directions.

3.2 Connection Assignment

The differences between the two network structures are further accentuated by their connection assignment schemes. PEC networks adopt a simple approach to obtaining an asymmetric network; *viz*, modifying a symmetric chordal ring by removing some inter-level and all intra-level end-around links, which leads to the distinction between interface nodes and non-interface nodes in terms of their degrees. Hierarchical levels are introduced into the PEC network, where each node $v = i2^h + 2^{h-1}$ is located at a unique level h and each level consists of $N/2^h$ nodes. Thus, two connections of equal skip distance never overlap in a PEC network.

On the other hand, groups are introduced into PRC rings and optimal skip distances are selected to be close to powers of $N^{1/n}$. The periodically regular connections cluster every $g = n-1$ consecutive nodes into one group. Each node $v = ig + (g-h)$ is identified as being at the $(g-h)$th position within group i. The group size g is also the period of the connection assignment so that if node u is connected to node v, then node $u+g$ is connected to node $v+g$ for all u and v. Thus, the connection pattern is repeated after every g nodes around the ring [5].

Fig. 2 shows an example of how 8-node PEC network and PRC ring are constructed by removing some connections from the symmetric chordal ring of Fig. 1. In the hierarchical layout of Fig. 3 nodes of the PEC network are arranged in $\log_2 N + 1$ levels. Direct connections are provided from each higher level to the "ground" level but not between the higher levels. Thus, all inter-level communications must go through nodes at the ground level. The PRC ring may be seen as providing both inter-level (7 to 0 and 0 to/from 4) and intra-level (6 to 2 and 7 to 1) end-around connections. Such differences in arrangements, together with the absence or presence of the ring property, result in different routing strategies for these networks.

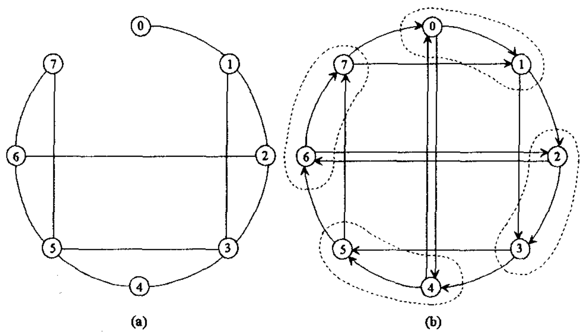

Fig. 2. Circular representation of (a) PEC network and (b) PRC ring.

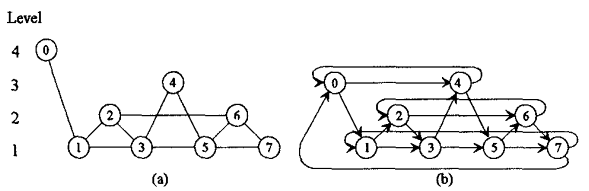

Fig. 3. Hierarchical construction of (a) PEC network and (b) PRC ring.

4. Routing Schemes

4.1 PEC Network

An irregular network structure incurs extra routing complexity in view of the need to examine the boundary conditions. H-routing [8] is a deterministic routing algorithm which favors the greedy use of links with large skip distances subject to the condition of no overlap between the endpoints of the links used in the route. If level h is selected to pass a message but is inaccessible to the source and destination nodes, the routing steps must go up and down to level h by taking two extra hops along the local links.

Let H(n) be the number of hops used by H-routing for passing a message from node 1 to node $2^n - 1$. The path length from node 0 to node $2^n - 1$ is thus H(n) + 1. The greedy use of skip distance $s_h = 2^h$ will exhaust all the $2^{n-h} - 1$ links in the middle of the PEC network. The process then repeats for the remaining non-overlapped $2^{h-1} - 1$ nodes at either end by selecting the next available level. This divide-and-conquer approach leads to the recurrence

$$H(n) = \min_h (2H(h-1) + 2^{n-h}) + 1$$

The initial condition H(1)=0 leads to $H(n) = \Theta(\sqrt{n}\, 2^{\sqrt{2n}})$ [9]. Unfortunately, this algorithm is centralized and we see no way of modifying it to run in a distributed manner such that routing decisions are made distributively on a hop-by-hop basis.

4.2 PRC Ring

An important property of PRC rings, inherited from simple rings, is the modulo N reduction that allows the use of relative addressing scheme, or the "distance" from the current node to the destination, as opposed to using the actual destination address. A routing algorithm for unidirectional symmetric chordal rings is given in [7]. The same routing rule is used at each node: Reduce the distance to destination by the largest possible amount without overshooting the destination. Unlike the complex centralized algorithm needed for PEC networks, the price of asymmetry is much lower in the case of PRC rings, for which the following two-phase routing algorithm has been proposed [13].

In phase 0, which is the initial condition, the message is sent to a node at the first position within its group (with the largest skip distance s_g). This ensures that the message routing can take the largest available skip. At most $g - 1$ extra hops are implied by this phase. The next phase determines whether reducing the distance to the destination by the skip distance leads to a non-negative residue. If so, the message is sent along the corresponding skip link. Otherwise, the message is moved along the local link to the node which has the next longest skip distance s_{g-1}. The routing algorithm stays in this phase until the message reaches its destination. At most $g - 1$ extra hops are required for such circumferential transitions from one skip distance to the next.

This simple routing algorithm does not incur significant overhead. On the contrary, the reduced connectivity simplifies the decision making process. In a symmetric chordal ring, we need to examine several skip distances at each node in order to select the outgoing link. In PRC rings, we only need to examine one skip distance. The worst-case routing distance d_R satisfies

$$d_R \le n(\lceil N^{1/n} \rceil + 1) - 4$$

This upper bound is $2n - 4$ hops more than that of a symmetric chordal ring with node degree $2n$.

We can derive an optimal group size that minimizes the worst-case routing distance. This optimal group size increases with the number of nodes N. Fig. 5 shows the worst-case routing distances in N-node PEC network and PRC ring. One can see that lack of inter-level connections in the PEC network does lead to a higher communication overhead.

5. Other Properties

5.1 Bisection Width

The bisection width of a network is the minimum number of links cut when dividing the network into two equal halves. The bisection width b for a PEC network is

$$b = \log_2 N$$

which depends only on the network size N. The network is incrementally extensible by concatenating two N-node PEC networks.

The number of additional connections to be established is the new bisection width $b + 1$.

The bisection width b of an N-node PRC ring with group size $g = n - 1$ is equal to $2 + 2\sum_{h=1}^{n-1} s_h/(n-1)$. When $N^{1/n}$ is an integer and s_h is optimally chosen to be $N^{h/n}$, the exact bisection width becomes

$$b = 2 + \frac{2}{n-1}\left(\frac{N - N^{1/n}}{N^{1/n} - 1}\right)$$

which depends on both the network size N and the group size g. One disadvantage of PRC rings is that the optimal networks are not incrementally extensible. By fixing the bisection width and selecting a constant set of skip distances, the use of a smaller PRC ring as a building block incurs the penalty of greater routing distances and increased diameter. This effect is shown in Table 1, where the skip distances 4, 16, 64 and 256 are used to construct PRC rings from 256 to 4096 nodes with $g = 4$.

Table 1. PRC Rings Using Fixed Skip Distances

Number of nodes N	256	512	1024	2048	4096
Diameter d	14(12)	15(14)	17	21(18)	29(20)
Average distance $\bar{d}$	8.1(7)	8.8(8.1)	10	12.1(10.9)	16.1(12)
Worst-case routing distance d_R	16(15)	17(16)	19	23(22)	31(26)
Average routing distance $\bar{d}_R$	9.8(8.6)	10.3(9.7)	11.3	13.4(12.8)	17.4(16.4)

Note that even though the diameter with respect to the optimal value (shown in parentheses if different) is relatively large for $N = 4096$ nodes, both the worst-case routing distance d_R and the average routing distance $\bar{d}_R$ exhibit smaller differentials. From a practical standpoint, the average routing distance is a more accurate predictor of performance than the worst-case routing distance, which is in turn more relevant than the diameter d. Hence, these results are quite encouraging.

5.2 Number of Links

In an N-node PEC network, we have 1 node with degree 1 (node 0), 2 nodes with degree 2 (nodes $N/2$ and $N - 1$), $2\log_2 N - 3$ nodes with degree 3, and $N - \log_2 N$ nodes with degree 4. The total number of bidirectional links is therefore $2N - \log_2 N - 2$. The number of unidirectional links in an N-node PRC ring is $2N$. Thus, link count is roughly the same for both networks. However, one must note that, depending on the implementation technology, time overhead may be involved in switching a bidirectional link from one direction to the other. Hence, higher-performance (and thus higher-cost) bidirectional links may be needed to provide the same throughput as a corresponding unidirectional link.

Because PEC networks have a larger number of short links and a smaller number of long links compared to PRC rings, link cost (e.g., as measured by VLSI area) is lower for PEC networks.

5.3 Cartesian Product Networks

A natural way to extend a network topology is to combine basic graphs by their Cartesian product [16]. Packed exponential connections, as originally proposed, are built on top of a 2D mesh [10]. The resulting topology is the Cartesian product of two 1D PEC networks as shown in the 8×8 example of Fig. 4(a). The Cartesian product is equally applicable as a composition tool to PRC rings, where the periodically regular connections are built on top of a 2D torus. Fig. 4(b) shows an 8×8 PRC ring (edge directions are not shown to avoid clutter and to facilitate comparison). Although the node degree is doubled in both cases, the extent to which the network diameter is affected is different for the two networks.

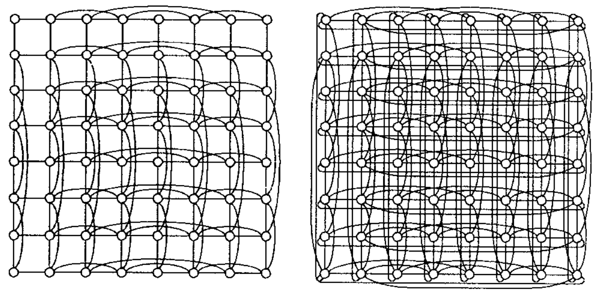

Fig. 4. Cartesian product of a pair of (a) PEC networks (b) PRC rings.

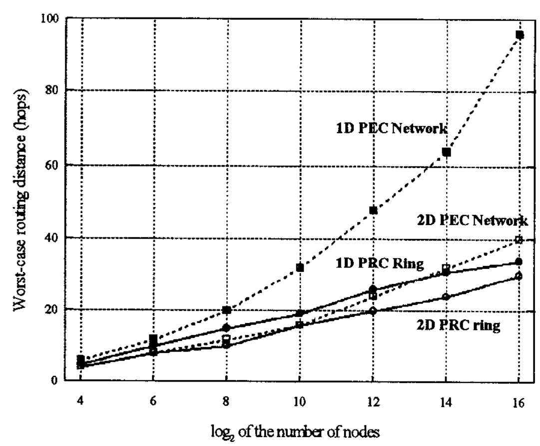

Fig. 5. Improvement of worst-case routing distance in 2D product networks.

We compare the worst-case routing distances obtained by an $N{\times}N$ 2D structure and an N^2-node 1D structure for both PEC networks and PRC rings (Fig. 5). For an N^2-node 1D PEC network, the worst-case routing distance $\Theta(\sqrt{\log_2 N}\, 2^{2\sqrt{\log_2 N}})$ is considerably greater than $\Theta(\sqrt{\log_2 N}\, 2^{\sqrt{2\log_2 N}})$ of an $N{\times}N$ 2D PEC network. The worst-case routing distance for an $N{\times}N$ 2D PRC ring with group size $g = n - 1$ is bounded by $2n(\lceil N^{1/n} \rceil + 1) - 8$. This is comparable to that of an N^2-node 1D PRC ring with group size $g = 2n - 1$, which has diameter $2n(\lceil N^{1/n} \rceil + 1) - 4$. Thus, the improvement is not significant. However, since the skip distances and connection patterns are different for the 2D version of PRC rings, their use may be justified on the basis of lower implementation cost as opposed to smaller diameter. This issue merits further investigation.

6 Conclusions

In this paper, we have compared two distributed and reduced connection assignment strategies. PEC networks arrange nodes in hierarchical levels and PRC rings cluster nodes in groups. These two "thinning" strategies lead to different routing schemes and architectural properties. The cost/performance parameters of the two networks are summarized in Table 2.

Table 2. Cost/Performance Parameters for Different Degree-4 Networks

	Bidirectional		Unidirectional	
	PEC	2D Torus	PRC Ring	2D MSN
Worst-case routing distance d_R	$\sqrt{\log_2 N}\, 2^{\sqrt{2\log_2 N}}$	$N^{1/2}$	$nN^{1/n}+n-4$	$N^{1/2}+1$
Average routing distance $\overline{d}_R$	Not available	$\sim d_R/2$	$> d_R/2$	$\sim d_R/2$
Bisection width b	$\log_2 N$	$2N^{1/2}$	$\frac{2(N-N^{1/n})}{(n-1)(N^{1/n}-1)}+2$	$2N^{1/2}$
Total number of links	$2N-\log_2 N-2$	$2N$	$2N$	$2N$

For a fixed number of nodes, the number of edges captures the cost of links and I/O ports, and hence is a reasonable first approximation to the cost of different network topologies. However, a single overall measure for the performance of networks is hard to define and little agreement over a standard measure exists. Despite their shortcomings, the worst-case and average routing distances and bisection width are listed in Table 2 as performance indicators. For completeness, 2D torus and its unidirectional variant, known as Manhattan street network (MSN), with same node degrees are also shown [3]. It is noteworthy that the PEC network and MSN [11] require more complicated routing schemes, while the PRC ring has the same routing simplicity as the torus.

Although in such heterogeneous networks certain links may become more congested than others, PRC rings are expected to be both more resilient and less prone to congestion when there is a significant level of non-local or random communication. This can be explained by observing that PRC rings preserve periodic regularity while PEC networks have fewer links for larger skip distances (the number of skip links provided is halved with each doubling of the length). The price that one pays for the above advantages of PRC rings is limited expandability and a more complex connection pattern which translates into greater area or space for interconnections.

References

[1] J.-C. Bermond, F. Comellas, and D. F. Du, "Distributed loop computer networks: a survey," *J. Parallel Distrib. Comput.*, vol. 24, pp. 2-10, Jan. 1995.

[2] M.-S. Chen, K. G. Shin, and D. D. Kandlur, "Addressing, routing, and broadcasting in hexagonal mesh multiprocessors," *IEEE Trans. Comput.*, vol. 39, pp. 10-18, Jan. 1990.

[3] T. Y. Chung, N. Sharma, and D. P. Agrawal, "Cost-performance tradeoffs in Manhattan street networks versus 2D torus," *IEEE Trans. Comput.*, vol. 43, pp. 240-243, Feb. 1994.

[4] S.P. Dandamudi and D. L. Eager, "Hierarchical interconnection networks for multicomputer systems," *IEEE Trans. Comput.*, vol. 39, pp. 786-797, Jun. 1990.

[5] K. W. Doty, "New designs for dense processor interconnection networks," *IEEE Trans. Comput.*, vol. C-33, pp. 447-450, May 1984.

[6] K. Ghose and K. R. Desai, "Hierarchical cubic networks," *IEEE Trans. Parallel Distrib. Syst*, vol. 6, pp. 427-435, Apr. 1995.

[7] L. M. Huisman and S. Kundu, "Highly reliable symmetric networks," *IEEE Trans. Parallel Distrib. Syst.*, vol. 5, pp. 94-98, Jan. 1994.

[8] C. C. Lin and V. K. Prasanna, "A routing algorithm for PEC networks," *Proc. 4th Symp. Frontiers of Massively Parallel Computation*, Oct. 1992, pp. 170-177.

[9] ——, "Bounds on the diameter of one dimensional PEC networks," *J. Parallel and Distributed Computing, to appear.*

[10] W. W. Kirkman and D. Quammen, "Packed exponential connections – a hierarchy of 2-D meshes," *Proc. Int'l Parallel Processing Symp.*, pp. 464-470, 1991.

[11] N. F. Maxemchuk, "Routing in the Manhattan street network," *IEEE Trans. Commun.*, vol. COM-35, pp. 503-512, May 1985.

[12] R. J. McMillen and H. J. Siegel, "Routing schemes for the augmented data manipulator network in an MIMD system," *IEEE Trans. Comput.*, vol. C-31, pp. 1202-1214, Dec. 1982.

[13] B. Parhami, "Periodically regular chordal ring networks for massively parallel architectures," *Proc. 5th Symp. Frontiers Massively Parallel Comput.*, Feb. 1995, pp. 315-322.

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

[15] C. K. Wong and D. Coppersmith, "A combinatorial problem related to multimodule memory organization," *J. ACM*, vol. 21, pp. 392-402, 1974.

[16] A. Youssef, "Design and analysis of product networks," *Proc. 5th Symp. Frontiers Massively Parallel Comput.*, Feb. 1995, pp. 521-528.

[17] S. G. Ziavras, "RH: a versatile family of reduced hypercube interconnection networks," *IEEE Trans. Parallel Distrib. Syst.*, vol. 5, pp. 1210-1220, Nov. 1994.