Information Processing Letters 110 (2010) 211–215

Contents lists available at ScienceDirect

Information Processing Letters

www.elsevier.com/locate/ipl

Fully symmetric swapped networks based on bipartite cluster connectivity ☆

Wenjun Xiao [a], Behrooz Parhami [b,*], Weidong Chen [a], Mingxin He [a], Wenhong Wei [a]

[a] *School of Software Engineering, South China University of Technology, Guangzhou 510641, PR China*
[b] *Department of Electrical and Computer Engineering, University of California, Santa Barbara, CA 93106-9560, USA*

ARTICLE INFO

Article history:
Received 15 August 2009
Received in revised form 15 December 2009
Accepted 15 December 2009
Available online 21 December 2009
Communicated by A.A. Bertossi

Keywords:
Bipartite graph
Hierarchical network
Interconnection network
OTIS network
Routing algorithm
Swapped network
Topological parameters

ABSTRACT

The class of swapped or OTIS (optical transpose interconnect system) networks, built of n copies of an n-node cluster by connecting node i in cluster j to node j in cluster i for $i \neq j$, has been studied extensively. One problem with such networks is that node i of cluster i has no intercluster link. This slight asymmetry complicates a number of algorithms and hinders both theoretical investigations and practical pursuits, such as building parallel node-disjoint paths for fault tolerance. We introduce biswapped networks that are fully symmetric and have cluster connectivity very similar to swapped/OTIS networks. We derive basic topological parameters, present a simple distributed shortest-path routing algorithm, and point to a number of other interesting properties under investigation for biswapped networks.

1. Introduction

Symmetry is a very desirable property of an interconnection network. For example, node symmetry (node transitivity) allows one to develop a single generic algorithm for routing that is applicable to every node in the network. If the routing algorithm is to be fully distributed, so that each node makes its own decision as to how to forward an incoming message, the advantages of symmetry become even more pronounced. Symmetry also facilitates theoretical proofs and certain derivations of practical significance, such as computing the average internode distance or constructing a set of node-disjoint paths between a pair of nodes (to allow the routing of messages in parallel or to identify alternate paths in the event of node/link failures).

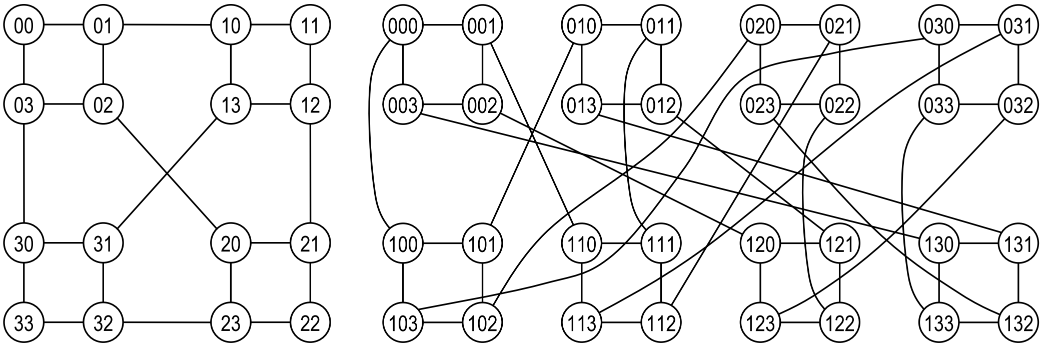
Our motivation for this study arose from a small, yet very important (in both theoretical and practical terms), asymmetry in swapped [8,6] or OTIS [5,3] networks, that have been found of interest by researchers in communications and parallel computing. A swapped/OTIS network $Sw(\Omega)$ is built from n clusters, which are identical copies of an n-node basis network Ω, by connecting node i of cluster j to node j of cluster i, for all $i \neq j$. The latter condition is what causes the asymmetry, because node i of cluster i, which has no intercluster link, maintains its original degree δ, whereas all other nodes have degree $\delta + 1$. Consequently, many analyses and algorithms for swapped/OTIS networks become complicated by the need to treat node (i, i) differently from a typical node (i, j) having unequal cluster index i and node index j. This lack of full symmetry prevents us from using results that are applicable to classes of symmetric networks, such as the ubiquitous Cayley graphs [1,2,4,7].

We thus pondered the existence question for "an alternate or modified form of swapped network that is a Cayley graph when the basis network is a Cayley graph" [6].

☆ Research of W. Xiao, W. Chen, M. He, and W. Wei was supported by the Natural Science Foundation of China (60973150).

* Corresponding author.
E-mail addresses: wjxiao@scut.edu.cn (W. Xiao), parhami@ece.ucsb.edu (B. Parhami).

0020-0190/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.ipl.2009.12.006



(a) $Sw(C_4)$ (b) $Bsw(C_4)$

Fig. 1. Example 16-node swapped/OTIS network and the corresponding 32-node biswapped network based on $\Omega = C_4$. To avoid clutter, the node indices $\langle c, g\rangle$ for $Sw(C_4)$ and $\langle i, c, g\rangle$ for $Bsw(C_4)$ are shown as cg and icg, where i, c, and g are part, cluster, and node indices, respectively.

Biswapped networks, being proposed here, constitute our positive answer to this interesting question.

2. Definitions and basic properties

Let Ω be any undirected graph with the vertex set $V(\Omega) = \{g_1, g_2, \ldots, g_n\}$ and the edge set $E(\Omega)$. The biswapped interconnection network $Bsw(\Omega) = \Sigma = (V(\Sigma), E(\Sigma))$ is a digraph with its vertex and edge sets specified as:

$$V(\Sigma) = \{\langle i, c, g\rangle \mid i \in \{0, 1\}; c, g \in V(\Omega)\},$$

$$E(\Sigma) = \{(\langle i, c, g_1\rangle, \langle i, c, g_2\rangle) \mid i \in \{0, 1\}, c \in V(\Omega), (g_1, g_2) \in E(\Omega)\} \cup \{(\langle 0, c, g\rangle, \langle 1, g, c\rangle) \mid c, g \in V(\Omega)\}.$$

Intuitively, the definition postulates $2n$ clusters, each cluster being an Ω graph: n clusters, with nodes indexed $\langle 0, \text{cluster\#}, \text{node\#}\rangle$, form part 0 of the graph, and n clusters constitute part 1, with associated node indices $\langle 1, \text{cluster\#}, \text{node\#}\rangle$. Each cluster c in either part of Σ has the same internal connectivity as Ω (intracluster edges, forming the first set in the definition of $E(\Sigma)$). In addition, node g of cluster c in part i is connected to node c in cluster g of part $1 - i$ (intercluster or swap edges forming the second set in the definition for $E(\Sigma)$). The name "biswapped network" arises from two defining properties of the network just introduced: when clusters are viewed as supernodes, the resulting graph of supernodes is the complete $2n$-node bipartite graph $K_{n,n}$, and the intercluster links connect nodes in which the cluster number and the node number within cluster are interchanged or swapped.

For example, when the basis graph is $\Omega = C_4$ (undirected cycle of order 4), the resulting $Bsw(C_4)$ is shown in Fig. 1b. Part 0 of the network is drawn at the top and part 1 at the bottom, with clusters 0–3 positioned from left to right. For comparison, $Sw(C_4)$ is shown in Fig. 1a.

We need a few more notational conventions in what follows. For any graph Γ, the number of its nodes is denoted as $|\Gamma|$. The degree of a node g in Γ is $\deg_\Gamma(g)$. The distance, that is, the length of the shortest path, between nodes g_1 and g_2 in Γ is given by $\text{dist}_\Gamma(g_1, g_2)$. The diameter of Γ, that is, the maximum distance between any two nodes in Γ, is $D(\Gamma)$. We next prove a number of results on the basic parameters of $\Sigma = Bsw(\Omega)$.

Theorem 1. *Let $\Sigma = Bsw(\Omega)$. Then:*

(1) $|\Sigma| = 2|\Omega|^2$;
(2) $\deg_\Sigma(\langle i, c, g\rangle) = \deg_\Omega(g) + 1$;
(3) $\text{dist}_\Sigma(\langle i, c_1, g_1\rangle, \langle i, c_2, g_2\rangle)$ *equals* $\text{dist}_\Omega(g_1, g_2)$ *if* $c_1 = c_2$;
otherwise, it equals $\text{dist}_\Omega(c_1, c_2) + \text{dist}_\Omega(g_1, g_2) + 2$;
(4) $\text{dist}_\Sigma(\langle i, c_1, g_1\rangle, \langle 1 - i, c_2, g_2\rangle) = \text{dist}_\Omega(c_1, g_2) + \text{dist}_\Omega(c_2, g_1) + 1$.

Proof. Statements (1) and (2) are evident from the definition of biswapped networks; there are $2n^2$ nodes in a biswapped network based on an n-node basis graph, and the node degree increases by 1 owing to the introduction of intercluster or swap links. To prove the first part of statement (3), we note that by the definition of $Bsw(\Omega)$, we have $\text{dist}_\Sigma(\langle i, c, g_1\rangle, \langle i, c, g_2\rangle) \leqslant \text{dist}_\Omega(g_1, g_2)$. To complete the first part of the proof for statement (3), we must show that no shorter path exists between nodes g_1 and g_2 of cluster c that goes through other clusters. This is established by contradiction. Suppose that the following path from $\langle 0, c, g_1\rangle$ to $\langle 0, c, g_2\rangle$, via intermediate clusters $x_1, y_1, x_2, y_2, \ldots, x_k, y_k, x_{k+1}$, is shorter than the path of length $\text{dist}_\Omega(g_1, g_2)$ within cluster c: $\langle 0, c, g_1\rangle \to \cdots \to \langle 0, c, x_1\rangle \to \langle 1, x_1, c\rangle \to \cdots \to \langle 1, x_1, y_1\rangle \to \langle 0, y_1, x_1\rangle \to \cdots \to \langle 0, y_1, x_2\rangle \to \langle 1, x_2, y_1\rangle \to \cdots \to \langle 1, x_2, y_2\rangle \to \cdots \to \langle 1, x_k, y_k\rangle \to \langle 0, y_k, x_k\rangle \to \cdots \to \langle 0, y_k, x_{k+1}\rangle \to \langle 1, x_{k+1}, y_k\rangle \to \cdots \to \langle 1, x_{k+1}, c\rangle \to \langle 0, c, x_{k+1}\rangle \to \cdots \to \langle 0, c, g_2\rangle$. The length of this path includes distances from g_1 to x_1, x_1 to $x_2, \ldots, x_{k+1}$ to g_2 within a cluster, plus a number of other segments. Given that all the clusters are isomorphic, the latter path cannot be shorter than $\text{dist}_\Omega(g_1, g_2)$. Now assume $c_1 \neq c_2$ and consider the following path between two nodes in part 0 that goes through a single intermediate cluster c in part 1. The path consists of five segments, for which the hop distance associated with each segment is provided below the corresponding arrows:

$$\langle 0, c_1, g_1\rangle \underset{\mathrm{dist}_\Omega(g_1,c)}{\rightarrow} \cdots \rightarrow \langle 0, c_1, c\rangle \underset{1}{\rightarrow} \langle 1, c, c_1\rangle$$
$$\underset{\mathrm{dist}_\Omega(c_1,c_2)}{\rightarrow} \cdots \rightarrow \langle 1, c, c_2\rangle \underset{1}{\rightarrow} \langle 0, c_2, c\rangle$$
$$\underset{\mathrm{dist}_\Omega(g,g_2)}{\rightarrow} \cdots \rightarrow \langle 0, c_2, g_2\rangle.$$

Based on the path above, we conclude that

$$\begin{aligned}&\mathrm{dist}_\Sigma\big(\langle i, c_1, g_1\rangle, \langle i, c_2, g_2\rangle\big)\\ &\quad\leqslant \min_{c\in V(\Omega)}\big\{\mathrm{dist}_\Omega(g_1, c) + \mathrm{dist}_\Omega(c, g_2)\big\}\\ &\qquad + \mathrm{dist}_\Omega(c_1, c_2) + 2\\ &\quad= \mathrm{dist}_\Omega(g_1, g_2) + \mathrm{dist}_\Omega(c_1, c_2) + 2.\end{aligned}$$

An argument similar to the one presented for proving the first part of statement (3) can be used to establish that no shorter path can go through more than one intermediate cluster. Statement (4) can be proven similarly by considering the following three-segment path, which includes a single intercluster edge, from node $\langle i, c_1, g_1\rangle$ to node $\langle 1-i, c_2, g_2\rangle$:

$$\langle i, c_1, g_1\rangle \underset{\mathrm{dist}_\Omega(g_1,c_2)}{\rightarrow} \cdots \rightarrow \langle i, c_1, c_2\rangle \underset{1}{\rightarrow} \langle 1-i, c_2, c_1\rangle$$
$$\underset{\mathrm{dist}_\Omega(c_1,g_2)}{\rightarrow} \cdots \rightarrow \langle 1-i, c_2, g_2\rangle.$$

The rest of the argument parallels that used for proving statement (3), and is thus omitted here. □

Corollary 1. *The diameter of $\Sigma = Bsw(\Omega)$ is related to the diameter of the basis network Ω by the equality $D(\Sigma) = 2D(\Omega) + 2$.*

Proof. By statement (3) in Theorem 1, $D(\Sigma) \leqslant 2D(\Omega) + 2$. Let $\mathrm{dist}_\Omega(g_1, g_2) = D(\Omega)$. Then, we have $\mathrm{dist}_\Sigma(\langle i, g_1, g_1\rangle, \langle i, g_2, g_2\rangle) = 2\,\mathrm{dist}_\Omega(g_1, g_2) + 2$, which establishes the desired result. □

3. Distributed shortest-path routing

Let Ω be any undirected graph with the vertex set $V(\Omega) = \{g_1, g_2, \dots, g_n\}$ and the edge set $E(\Omega)$. Based on Theorem 1, we can easily obtain a shortest-path (optimal) routing algorithm for a biswapped network, given the availability of an optimal routing algorithm for the basis graph Ω. Assume that the latter routing algorithm is a distributed one, using the local function $next_\Omega(g_1, g_2)$ to obtain the first intermediate node in the routing path from g_1 to g_2. Then, the algorithm in Fig. 2 can be used to derive the first intermediate node on a shortest routing path from node $\langle i, c_1, g_1\rangle$ to node $\langle j, c_2, g_2\rangle$ in $\Sigma = Bsw(\Omega)$. Optimality of this algorithm is proven in Theorem 2.

Theorem 2. *The routing function $next_\Sigma$, defined in Fig. 2, is optimal, that is, it guarantees shortest-path routing, provided $next_\Omega$ is an optimal routing function for Ω.*

Proof. A routing path in $Bsw(\Omega)$ may be one of three types. First, for a routing path that begins and ends in the same cluster ($i = j, c_1 = c_2, g_1 \neq g_2$), the forwarding

```
function nextΣ(⟨i, c1, g1⟩, ⟨j, c2, g2⟩)
if i = j
then                                      // i = j; routing in the same part
  if c1 = c2 and g1 = g2
  then return ⟨i, c1, g1⟩                 // destination has been reached
  else
    if g1 = g2
    then return ⟨1 − i, g1, c1⟩           // g1 = g2 and c1 ≠ c2
    else return ⟨i, c1, nextΩ(g1, g2)⟩    // g1 ≠ g2
    endif
  endif
else                                      // i ≠ j; routing between parts
  if c1 = g2 and g1 = c2
  then return ⟨j, c2, g2⟩                 // destination is one hop away
  else
    if g1 = c2
    then return ⟨j, g1, c1⟩               // ⟨j, g1, c1⟩ = ⟨1 − i, g1, c1⟩
    else return ⟨i, c1, nextΩ(g1, c2)⟩
    endif
  endif
endif
```

Fig. 2. Optimal routing function for a biswapped network $Bsw(\Omega)$ based on the optimal routing function $next_\Omega(g_1, g_2)$ for its basis network Ω.

node chosen by $next_\Sigma$ is $\langle i, c_1, next_\Omega(g_1, g_2)\rangle$. Thus, the forwarding path, which remains in the same cluster until $g_1 = g_2$ holds, is optimal by our assumption regarding the optimality of $next_\Omega$. Second, for a path that begins and ends in the same part, but not in the same cluster ($i = j, c_1 \neq c_2$), the routing algorithm first causes the message to be moved within the source cluster until the node $\langle i, c_1, g_2\rangle$ has been reached (note that if $g_1 = g_2$, this segment of the path is empty). Then, the intercluster link to $\langle 1-i, g_2, c_1\rangle$ is used, followed by routing within the same cluster to $\langle 1-i, g_2, c_2\rangle$ using $next_\Omega$, and finally via an intercluster link to $\langle i, c_2, g_2\rangle$. This path is an instance of the five-segment shortest path hypothesized in the proof of Theorem 1, with g_2 here taking the place of c in that path, thus making the fifth segment unnecessary (of length 0). The third case pertains to a path that begins in part i and ends in part $j = 1 - i$. In this case, the algorithm first aims to reach the node $\langle i, c_1, c_2\rangle$ in the source cluster. Routing is then completed via $\langle 1-i, c_2, c_1\rangle$, followed by a path dictated by $next_\Omega$ within the destination cluster to the final destination $\langle 1-i, c_2, g_2\rangle$. The preceding path is precisely the three-segment shortest path hypothesized near the end of the proof of Theorem 1. □

4. Comparison with swapped/OTIS networks

Because swapped/OTIS networks are known to have advantages over other well-known networks in terms of topological properties, performance, scalability, and fault tolerance (see, e.g., [6] and the references therein), demonstrating that biswapped networks are preferable to swapped networks can be taken as indirect evidence of advantages over those other networks.

A biswapped network of node degree $\delta + 1$ (where δ is the node degree of its n-node basis graph Ω) has $2n^2$ nodes, compared with n^2 nodes for a degree-$(\delta + 1)$ swapped network formed from the same basis graph. The only penalty for doubling the number of nodes is a unit increase in network diameter, from $2D + 1$ for swapped/OTIS to $2D+2$ for biswapped, where D is the diameter of Ω. This doubling of network size for a unit increase in diameter is a worthwhile trade-off, in that for most component networks that already have the maximum number of nodes for their node degree and diameter (this is true of square meshes/tori, hypercubes, star graphs, and so on), increasing the size of Ω to approximately $1.4n$ so as to have close to $2n^2$ nodes in a swapped/OTIS network would lead to a 1-unit increase in D, and thus a 2-unit increase in the network diameter.

Another way to look at the trade-off just discussed is that if $D = \log_2 n$, the diameter of a swapped network is $\log_2(n^2)+1$ and the diameter of a biswapped network is $\log_2(2n^2)+1$; i.e., with logarithmic-diameter basis networks, the two networks are similar in terms of diameter, given their sizes. For any basis network that has super-logarithmic diameter, however, biswapped networks would win on account of their diameter. Note that biswapped networks are similar to swapped/OTIS networks in two other important respects: (1) they have power-of-2 size when n is a power of 2; and (2) their incremental scalability, that is, the relative expansion in size when n is increased to $n+1$, is approximately $2/n$.

As evidence of algorithmic simplicity, we point to the simple and elegant distributed routing algorithm depicted in Fig. 2, which requires only a few numerical comparisons between the components of the current node's address and those of the destination address to decide on the outgoing channel belonging to a shortest path. These comparisons can be performed in hardware and imply a very small routing latency (pipelined, if necessary), thus enabling fast wormhole switching. By contrast, known optimal routing algorithms for swapped/OTIS networks are quite complicated and require computing and comparing distances in the basis network. This is because the shortest path from node (i, j) to node (k,l) can be one of the two paths, that include one and two intercluster links: which one is shorter depends on the intracluster distances between certain intermediate nodes. In effect, we would need separate optimal routing algorithms for swapped/OTIS networks built of different basis networks (referred to in the literature as OTIS-mesh, OTIS-hypercube, OTIS-star, and so on). Other evidence of the superiority of biswapped networks to swapped/OTIS networks will be reported in the near future.

To summarize, in view of indirect evidence of superiority for biswapped networks, arising from properties that they inherit from swapped/OTIS networks, biswapped networks constitute an important addition to the repertoire of parallel computer designers. These networks are at the same time competitive with, and complementary to, existing interconnection networks. They are competitive because they offer a cost-effective way of scaling network size, while maintaining desirable architectural features. They are complementary owing to the fact that they allow the use of virtually any existing network as the basis network, thus combining the advantages of particular cluster interconnections with the benefits resulting from the biswapped connectivity.

Table 1
Comparison of topological, cost-effectiveness, and robustness parameters of a network as it is scaled up from n nodes to $2n^2$ nodes.

Network	Degree	Diameter	Cost ratio	Connectivity
Cycle (ring)	$\delta = 2$	$4D^2$	$4D$	$\kappa = 2$
Square 2D torus	$\delta = 4$	$\approx 1.4D^2$	$\approx 1.4D$	$\kappa = 4$
Hypercube	$2\delta + 1$	$2D + 1$	≈ 4	$2\kappa + 1$
Biswapped	$\delta + 1$	$2D + 2$	≈ 2	$\delta + 1$

5. Conclusions

A biswapped network, which has a two-level structure, takes any graph as modules and connects them in a complete bipartite manner. Hence, the architecture of biswapped networks offers a simple general scheme for constructing larger networks from any component or basis network. Since the topology of a biswapped network is closely related to the topology of its basis network, it inherits some favorable properties from the latter. We have derived some general properties of a biswapped network based on parameters and structure of its basis network. Examples of properties that are inherited by biswapped networks, and which will be reported separately in the near future, include hamiltonicity and being a Cayley graph.

Biswapped networks offer advantages over the widely studied swapped networks, as previously discussed in Section 4. Table 1 shows how the topological, cost-effectiveness, and robustness parameters of a network change as it is scaled up from n nodes to $2n^2$ nodes, that is, the same size expansion offered by the biswapped architecture. The network's original node degree δ, diameter D, and connectivity κ change to values shown in Table 1. For example, when an n-node hypercube with node degree and diameter $\delta = D = \log_2 n$ is scaled up to include $2n^2$ nodes, its node degree and diameter become $\log_2(2n^2) = 2\delta + 1 = 2D + 1$. In particular, we note that a biswapped network is nearly twice as cost-effective as a hypercube of equal size using the commonly suggested degree-diameter product as an indicator. More importantly, we note that the biswapped structure offers maximal fault tolerance (connectivity) for its node degree, independent of the robustness of the basis network. This is an important advantage brought about by the biswapped architecture that we will exploit in our future work.

Results in this paper bring some closure to the topic of swapped/OTIS networks, which as previously defined lack full symmetry. Biswapped networks are completely symmetric and offer twice as many nodes as the corresponding swapped networks with the same node degree and with a unit increase in network diameter. This is an advantageous tradeoff for nearly all basis networks of practical interest.

References

[1] S.B. Akers, B. Krishnamurthy, A group theoretic model for symmetric interconnection networks, IEEE Trans. Comput. 38 (1989) 555–566.

[2] F. Annexstein, M. Baumslag, A.L. Rosenberg, Group action graphs and parallel architectures, SIAM J. Comput. 19 (1990) 544–569.

[3] K. Day, A. Al-Ayyoub, Topological properties of OTIS-networks, IEEE Trans. Parallel Distrib. Systems 13 (4) (April 2002) 359–366.

[4] M. Heydemann, Cayley graphs and interconnection networks, in: Graph Symmetry: Algebraic Methods and Applications, 1997, pp. 167–224.

[5] G. Marsden, P. Marchand, P. Harvey, S. Esener, Optical transpose interconnection system architectures, Opt. Lett. 18 (13) (1993) 1083–1085.

[6] B. Parhami, Swapped interconnection networks: Topological, performance, and robustness attributes, J. Parallel Distrib. Comput. 65 (2005) 1443–1452.

[7] W.J. Xiao, B. Parhami, Some mathematical properties of Cayley digraphs with applications to interconnection network design, Int. J. Comput. Math. 82 (2005) 521–528.

[8] C.-H. Yeh, B. Parhami, Swapped networks: Unifying the architectures and algorithms of a wide class of hierarchical parallel processors, in: Proc. International Conf. Parallel and Distributed Systems, June 1996, pp. 230–237.