Incrementally scalable optical interconnection network with a constant degree and constant diameter for parallel computing

Ahmed Louri and Costas Neocleous

A new scalable interconnection topology called the spanning-bus connected hypercube (SBCH) that is suitable for massively parallel systems is proposed. The SBCH uses the hypercube topology as a basic building block and connects such building blocks by use of multidimensional spanning buses. In doing so, the SBCH combines positive features of both the hypercube (small diameter, high connectivity, symmetry, simple routing, and fault tolerance) and the spanning-bus hypercube (SBH) (constant node degree, scalability, and ease of physical implementation), while at the same time circumventing their disadvantages. The SBCH topology permits the efficient support of many communication patterns found in different classes of computation, such as bus-based, mesh-based, and tree-based problems, as well as hypercube-based problems. A very attractive feature of the SBCH network is its ability to support a large number of processors while maintaining a constant degree and a constant diameter. Other positive features include symmetry, incremental scalability, and fault tolerance. An optical implementation methodology is proposed for the SBCH. The implementation methodology combines the advantages of free-space optics with those of wavelength-division multiplexing techniques. An analysis of the feasibility of the proposed network is also presented. © 1997 Optical Society of America

Key words: Interconnection networks, scalability, massively parallel processing, optical interconnects, wavelength division multiplexing, product networks.

1. Introduction

The interconnection network, not the processing elements (PE's) or their speed, is proving to be the decisive and determining factor in terms of the cost and performance of parallel-processing systems.[1–4] Several topologies have been proposed to fit different styles of computation. Examples include crossbars, multiple buses, multistage interconnection networks, and hypercubes, to name a few. Among these, the hypercube has received considerable attention mainly because of its good topological characteristics (small diameter, regularity, high connectivity, simple control and routing, symmetry, and fault tolerance) and its ability to permit the efficient embedding of numerous topologies, such as rings, trees, meshes, and shuffle exchanges, among others.[5]

The authors are with the Department of Electrical and Computer Engineering, The University of Arizona, Tucson, Arizona 85721.

Received 13 September 1996; revised manuscript received 22 April 1997.
0003-6935/97/266594-11$10.00/0

However, a drawback to the use of the hypercube is its lack of scalability, which limits its use in building large-size systems out of smaller-size systems. The lack of scalability of the hypercube stems from the fact that the node degree is not bounded and varies as $\log_2 N$. This property makes the hypercube cost prohibitive for large N. Most hypercube-based interconnection networks proposed in the literature[6–8] suffer from similar size-scalability problems.

Recently some networks have been introduced that are products of hypercube topology with some fixed-degree networks such as the mesh, the tree, and the de Bruijn[4,7,9] in the quest to preserve the properties of the hypercube while improving its scalability characteristics. Notable among these is the optical multimesh hypercube (OMMH).[10–12] The OMMH is a network that combines the positive features of the hypercube (small diameter, regularity, high connectivity, simple control and routing, symmetry, and fault tolerance) with those of a mesh (constant node degree and size scalability). The OMMH can be viewed as a two-level system: a local-connection level representing a set of hypercube modules and a global-connection level representing the mesh network that connects the hypercube modules.

The OMMH network has been demonstrated physically by use of a combination of free-space and fiber-optics technologies and has shown good performance characteristics for a reasonably sized network. However, for very large networks (greater than 1000 PE's), the OMMH experiences a logarithmic increase in terms of diameter and requires a large number of fibers that make implementation complicated and expensive.

In this paper, we propose a novel network that improves the topological characteristics as well as the implementation and performance aspects of the OMMH network. The new network topology proposed is called the spanning-bus connected hypercube (SBCH) and possesses a constant degree and a constant diameter while preserving all the properties of the hypercube. The SBCH, which is similar to the OMMH, employs the hypercube topology at the local-connection level. The global-connection level connecting the hypercube modules is a spanning-bus hypercube (SBH) network.[13]

The SBH is a D-dimensional lattice of width w in each dimension. Each node is connected to D buses, one in each of the orthogonal dimensions; w nodes share a bus in each dimension. The SBH offers small node degree, small diameter, low cost, and scalability. It can be scaled up by expansion of the size of the spanning buses.[13] However, expanding the size of the buses leads to a O(w) increase in traffic density,[13] which in turn leads to bus-congestion problems.[14]

The advantage of the SBCH network is that it utilizes the hypercube local-interconnection level to decrease traffic density, thereby alleviating the bus-congestion problems encountered in pure SBH networks. This feature allows the SBCH buses to support a larger number of processors than the SBH network, thus allowing larger systems to be built. As such, the SBCH is incrementally scalable, with a high degree of connectivity and a small diameter. Additionally, we also propose an optical implementation of such a network. Optical interconnects offer many desirable features, such as very large communication bandwidth, reduced cross talk, immunity to electromagnetic interference, and low power requirements.[3,4,15,16]

2. Structure of the Spanning-Bus Connected Hypercube Network

In this section, we formally define the structure of the SBCH network and discuss its properties.

A. Topology of the Spanning-Bus Connected Hypercube Interconnection Network

The size of the SBCH is characterized by a three-tuple (w, n, D), where w, n, and D are positive integers. The first parameter, w, defines the number of nodes attached to a bus. The second parameter, n, is the degree of the point-to-point n cube (hypercube). The third parameter, D, identifies the number of buses spanned by a PE in the network.

For a SBCH(w, n, D) the number of nodes $|V|$ is equal to $w^D 2^n$. A node address in the SBCH is denoted by a $(D + 1)$-tuple $(a_1, a_2, \ldots, a_D, a_n)$ by use of a mixed-radix system, where, for $i = 1$ to $i = D$, $0 \le a_i \le (D - 1)$ and $0 \le a_n \le (2^n - 1)$.

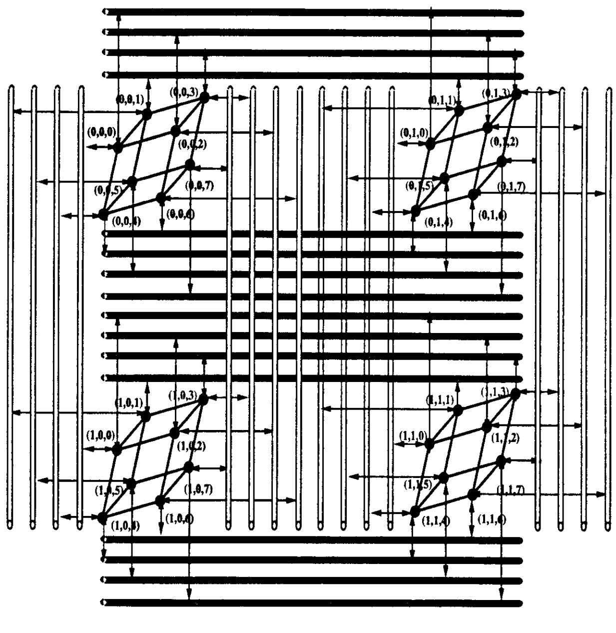
Given the set of nodes (V), the set of edges (E) is constructed as follows: For two nodes $(a_1, a_2, \ldots, a_D, a_n)$ and $(b_1, b_2, \ldots, b_D, b_n)$, where, for $i = 1$ to $i = D$, $0 \le a_i < D$ for $j = 1$ to $j = D$, we have $0 \le b_i < D$, $0 \le a_n < 2^n$, and $0 \le b_n < 2^n$:

1. The two nodes span the same bus (i) if $a_n = b_n$ and (ii) if, for $i = 1$ to $i = D$, there are only two components a_i and b_i that are identical, while all other components are different.
2. There is a link (called a hypercube link) between two nodes if and only if, for $i = 1$ to $i = D$, (i) $a_i = b_i$ and (ii) a_n and b_n differ by one bit position in their binary representation (Hamming distance of 1).

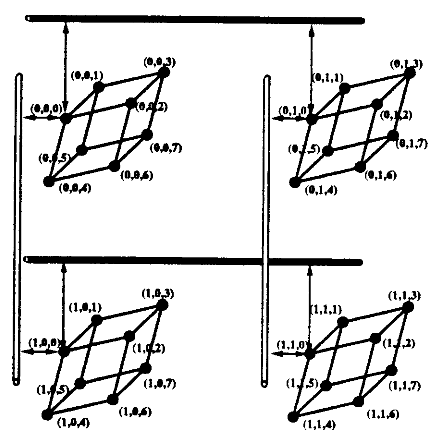
In this paper, we consider only SBCH networks with $D = 2$. Therefore, in the notation the third parameter, D, is dropped. Consequently, a SBCH$(w, n, 2)$ network is referred to as SBCH(w, n). Figure 1(a) shows a SBCH(2, 3) interconnection; the solid lines represent point-to-point hypercube links, and the dark thick lines represent buses. Small filled circles represent nodes of the SBCH network, which are, in this paper, abstractions of PE's or memory modules or switches. Note that, because $D = 2$, each node spans two buses, one bus along each dimension. Furthermore, there are three bidirectional point-to-point links, attached to a node, that correspond to the hypercube links. Careful observation of Fig. 1(a) shows that the node addresses satisfy the connection rules outlined above.

As can be seen from Fig. 1(a), the SBCH(2, 3) consists of $2^2 \times 2^3 = 32$ nodes. It can be viewed as eight concurrent two-dimensional (2-D) SBH's. Note that w horizontal buses and w vertical buses are needed to form one $w \times w$ 2-D SBH network. Figure 1(b) shows one such 2-D SBH formed by nodes with the same hypercube addresses and belonging to different hypercube modules. Similar considerations apply to the other seven 2-D SBH's shown in Fig. 1(a). The SBCH(2, 3) network can also be viewed as four concurrent three-dimensional (3-D) hypercubes in which four nodes having identical hypercube addresses form a 2×2 SBH. The SBCH(2, 3) shown in Fig. 1(a) looks like a hypercube-clustered SBCH. In general, there are 2^n 2-D SBH's and w^2 hypercube modules.

The choice of two parameters, w and n, completely determines the size of the network, the resources and implementation requirements, and the scaling complexity. The parameter w determines the size of the buses, whereas the parameter n defines the size of the hypercubes. From a scaling viewpoint, two scaling rules can be applied to a SBCH(w, n) network. The first rule, which we call the fixed-w rule, keeps the size of the buses constant and increases the size of the network by increasing n. The second rule, which we call the fixed-n rule, keeps the size of the



(a)



(b)

Fig. 1. (a) Example of the SBCH network: A SBCH(2, 3) (32 nodes) interconnection is shown. The solid thick lines represent bus connections, while the bold thin lines represent point-to-point hypercube connections. (b) Example of a 2-D SBH network within a SBCH(2, 3) network. Note that the nodes that construct the 2-D SBH belong to different hypercube modules but that they posses the same binary hypercube-address representation within their corresponding hypercube modules. Eight such 2-D SBH's coexist in the SBCH(2, 3) interconnection.

hypercube constant and increases the size of the network by increasing w. Clearly, the advantage of the SBCH(w, n) network is its flexibility to scale up by use of either or a combination of the two scaling rules.

For instance, the size of the SBCH can grow without altering the number of links per node by expansion of the size of the buses. For example, 3-D hypercubes can be added on the perimeter of the 2-D SBH's of Fig. 1. Figure 2 illustrates a SBCH(3, 3) that is constructed by expansion of the SBCH(2, 3) network by addition of the hypercube modules along an outer row and an outer column. The existing configuration of the nodes of the SBCH(2, 3) network did not change because each node still spans two buses and still has three bidirectional point-to-point links for the hypercube connections. This option allows the SBCH to be truly size scalable.

B. Properties of the Spanning-Bus Connected Hypercube Interconnection Network

1. *Diameter and Link Complexity*

The diameter of a network is defined as the maximum distance between any two processors in the network. Thus, the diameter determines the maximum number of hops that a message may have to take. Bearing in mind that $D = 2$, the diameter of a 2-D SBH is 2. The diameter of a hypercube with N nodes is $n = \log_2 N$; therefore the diameter of SBCH(w, n) is $(n + 2)$. For the SBCH(w, n) network, $N = w^2 2^n$, therefore $n = \log_2(N/w^2)$. Consequently the diameter of the SBCH(w, n) network can be written as $\log_2(N/w^2) + 2$. Using the fixed-w scaling rule shows that the diameter of the SBCH(w, n) network experiences a logarithmic increase [$O(\log_2 N)$] when the network size increases. However, using the fixed-n scaling rule would make the diameter constant for any network size. The constant value is $n + 2$.

Link complexity or node degree is defined as the number of physical links per node. For a regular network in which all nodes have the same number of links, the node degree of the network is that of a node. The node-link complexity, or degree, of a hypercube with N nodes is $n = \log_2 N$ and that of a 2-D SBH is 2. A node of a SBCH(w, n) network possesses links for both the hypercube connections and the bus connections. Consequently, the node degree of the SBCH network is $(n + 2)$ or $\log_2(N/w^2) + 2$. Again, when the fixed-w scaling rule is used the SBCH network experiences a logarithmic increase in degree [$O(\log_2 N)$]; however, when the network is expanded by use of the fixed-n scaling rule, the degree becomes constant, i.e., $(n + 2)$.

2. *Bisection Width*

The bisection width of a network is defined as the minimum number of links that have to be removed to partition the network into two equal halves.[17] The bisection width indicates the volume of communication allowed between any two halves of the network with an equal number of nodes. The bisection width of an n-dimensional hypercube is $2^{(n-1)} = N/2$, since that is the number of links that are connected between two $(n - 1)$-dimensional hypercubes to form an n-dimensional hypercube. Since there are w^2 such n-dimensional hypercubes connecting 2^n 2-D SBH's, the bisection width of a SBCH(w, n) is equal to $w^2 \times 2^{n-1} = N/2$.

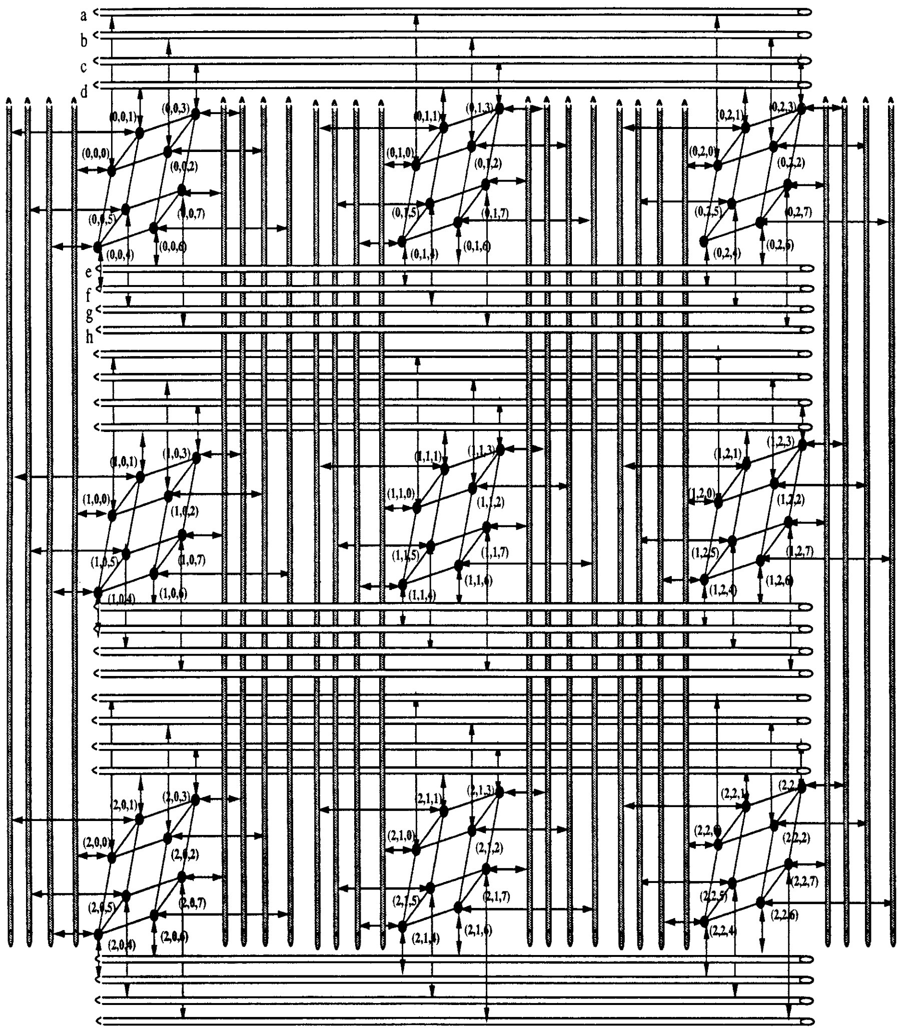


Fig. 2. SBCH(3, 3) (72 nodes) interconnection. This SBCH network can be constructed by the addition of hypercube modules along a row and a column to the network SBCH (2, 3) of Fig. 1.

3. *Granularity of Size Scaling*

Ideally it should be possible to create larger and more powerful networks by the simple addition of more nodes to the existing network. For a 2-D SBH the granularity of size scaling is only $2w + 1$ since at a minimum one bus per dimension could be added to the network to increase its size. Therefore the granularity of the size scaling in a $w \times w$ 2-D SBH of $N = w^2$ nodes is $2N^{1/2} + 1$. However, we can increase the size of a hypercube only by doubling the number of nodes; that is, the granularity of size scaling in an n-dimensional hypercube is 2^n.

In Subsection 2.A, we explained how the SBCH(w, n) network can be scaled up by use of two different scaling rules. When the fixed-w scaling rule is applied, the granularity of size scaling follows the hypercube size scaling. Therefore, the granularity of size scaling by use of the fixed-w rule is $w^2 \times 2^n = N$. When the fixed-n scaling rule is used, the granularity of size scaling follows that of the SBH. Therefore, the granularity of size scaling with the fixed-n rule is $2^n(2w + 1) = 2(N/w) + 2^n$. Note that the granularity of size scaling for the fixed-w rule is O(N), whereas for the fixed-n rule it is O(N/w).

4. *Topological Comparisons of the Spanning-Bus Connected Hypercube with Other Known Networks*

In this subsection we compare the SBCH network with existing well-known topologies. These include the Boolean hypercube (BHC),[5] the Torus network,[18] the SBH,[13] and the OMMH.[10] The comparison parameters include diameter, degree, number of links, and average traffic density. More details of the derivation of average traffic density and other parameters can be found in Ref. 19. The topological characteristics of the results of the comparison are shown in Fig. 3.

In Fig. 3 the notation SBCH(16, n) denotes that the network is expanded following the fixed$_{w=16}$ rule; that is, the size of the buses is kept constant (16 PE's per bus) and the size of the hypercube module is changed to have the same network size for comparison purposes. The notation SBCH(w, 4) denotes that the network is expanded following the fixed$_{n=4}$

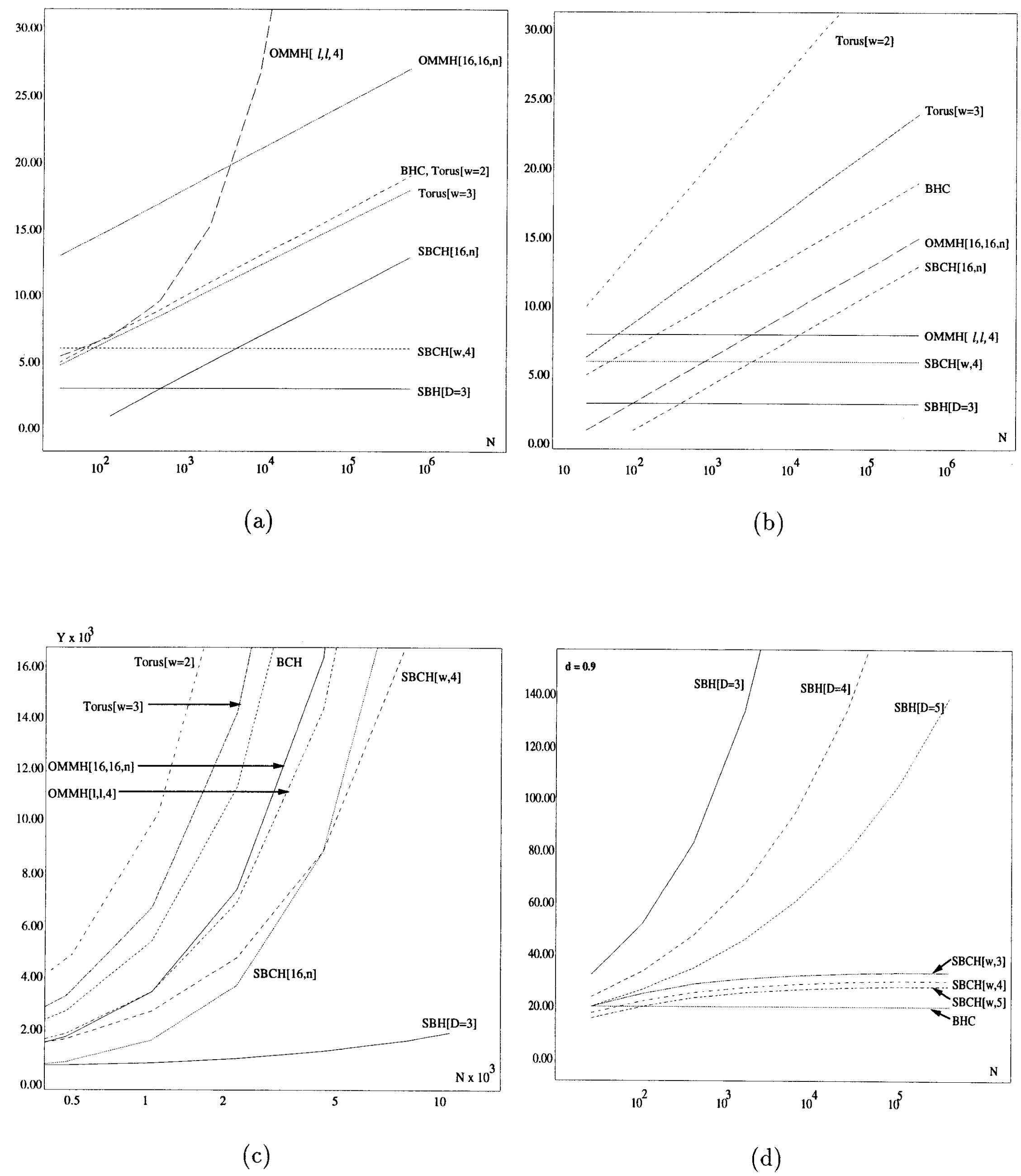


Fig. 3. Network comparisons for (a) diameter, (b) degree, (c) number of links, and (d) average traffic density.

rule, which means that the size of the hypercube module is kept fixed ($n = 4$) and the size of the buses is increased. Note that, when expanding the SBCH network, some mathematical constraints exist. The notation (16, 16, n)-OMMH denotes that the size of the mesh network in the OMMH is fixed, whereas the size of the hypercube is varied. Similarly the (l, l, 4)-OMMH notation denotes that the size of the hypercube is fixed, and the mesh size is varied. Finally, the notation SBH($D = 3$) means that the dimension of the SBH network is kept constant, and the size of the buses (w) is changed.

Figures 3(a) and 3(b) show the graph comparisons in terms of diameter and degree as the network size is increased. At the key mark of 10,000 nodes (desirable for massively parallel processing), SBCH(w, 4) and SBCH(16, n) exhibit very good performances in terms of diameter and degree, with values of 6 and 7, respectively. Note that SBCH(w, 4) is more desirable than SBCH(16, n) because it possesses constant degree and diameter, features that allow it to be scalable. Figure 3(c) shows that large SBCH networks are feasible with a small number of links. Conversely, the OMMH network experiences fairly large diameters, high numbers of links, and high topological cost. However the (l, l, 4)-OMMH possesses constant degree (8), a feature that makes it also size scalable. Nevertheless, the graphs indicate that the SBCH network improves the OMMH network drastically in terms of every topological characteristic. The SBCH network possesses very small diameter and low degree, while it offers size scalability and regularity.

Figures 3(a)–3(c) suggest that the SBH network experiences the best topological characteristics. Figure 3(d) illustrates a graphic comparison among the SBCH, the SBH, and the BHC in terms of average traffic density. The average traffic density is defined as the product of the average distance and the total number of nodes, divided by the total number of communication links.[15] The BHC has low traffic density, and it is insensitive to variations in network size. The SBCH network demonstrates the capacity for a higher traffic density than does the BHC, but for a larger network size it also exhibits no sensitivity to variations in network size. On the other hand, the SBH network shows the capacity for an increase in traffic density; therefore, for larger networks the SBH network most likely would experience severe bus-congestion problems that would lead to long message delays. Hence, even though the SBH network demonstrates better topological characteristics than does the SBCH network, the latter is more efficient because it can utilize a larger number of PE's with fewer communication delays (Ref. 19).

3. Optical Implementation of the Scanning-Bus Connected Hypercube Network

Obviously an electronic implementation of the proposed SBCH network is feasible. One methodology would be to use multiprocessor-board technology (e.g., multichip-module technology) for the hypercube subnetwork connections and backplanes for the bus connections. For limiting the number of boards required, k hypercube modules can be clustered together on a single multiprocessor board. However, for a large number of PE's and a greater bandwidth and interconnect density, conventional backplanes have major limitations.[3,4,20] These include signal skew, wave reflection, impedance mismatch, skin effects, and interference, among many others.

A possible alternative is the use of optical interconnects. Optical interconnects offer many communication advantages over electronics, including gigahertz transfer rates in an environment free from capacitive loading effects and electromagnetic interference, high interconnection density, low power requirements, and possibly a significant reduction in design complexity through the use of multiple-access techniques and the third dimension of free-space optics. The effectiveness of optical interconnects has been examined extensively.[3,4,15,16] In the following subsections we propose an all-optical implementation of the SBCH(w, n) network in which the hypercube modules are implemented by use of free-space space-invariant optics,[10] and the bus modules are implemented by use of wavelength-division multiple-access (WDMA) techniques.

A. Optical Implementation of Space-invariant Hypercubes by Use of Holographic Optical Elements

The free-space optical implementation of the hypercube network has been studied and analyzed rigorously in Refs. 4, 10, and 11. The main objective was to exploit the third dimension and the communication advantages of free-space optics to provide efficient and adequate implementation of the hypercube network. The design methodology is based on an observation that PE's in an interconnection network can be partitioned into two different sets such that any two PE's in a set do not have a direct link (except for completely connected networks). This is a well-known problem of bipartitioning a graph if the interconnection network is represented as a graph. For a binary n cube, PE's whose addresses differ by more than 1 in the Hamming distance can be in the same partition, since no link exists between two PE's if their Hamming distance is greater than 1.

Besides bipartitioning the graph, we arrange the PE's in each partition in the plane such that interconnection between two planes becomes space invariant (the connection pattern is identical for every PE in the plane). This self-imposed requirement reduces the design complexity of the optical setup. The two partitions of PE's are called plane$_l$ and plane$_r$. Optical sources and detectors are assumed to be resident on processor–memory boards located in plane$_l$ and plane$_r$. Free-space holographic optical elements (HOE's) are used to implement the connection patterns required among the PE's of the two planes.[4,10,11]

B. Implementation of the Spanning-Bus Hypercube by Use of Wavelength-Division Multiple-Access Techniques

In this subsection we present the implementation of the SBH subnetwork using WDMA techniques. For exploiting the large communication bandwidth of optics, WDMA techniques that permit multiple multiaccess channels to be realized on a single physical channel can be utilized. Optical passive star couplers can be utilized for the WDMA channels. The purpose of an $N \times N$ star coupler is to couple light from one of its N input guides to all the N output guides uniformly. Star couplers with 128×128 ports and the capacity to handle more than 100 different wavelengths are feasible with the currently available technology. An experimental integrated-services digital network (ISDN) switch architecture that uses eight 128×128 multiple star couplers to handle more than 10,000 input-port lines has been reported.[21]

The SBCH(w, n) network consists of w^2 hypercube modules and 2^n, $w \times w$ 2-D SBH's. From the discussion of Subsection 3.A, every hypercube module is bipartitioned into two planes, called plane$_l$ and plane$_r$. In the SBCH network all planes of plane$_l$ are grouped together to form a plane called PLANE$_L$, while all planes of plane$_r$ are grouped to form another plane called PLANE$_R$. The SBH buses can be implemented by interconnection of the individual planes of

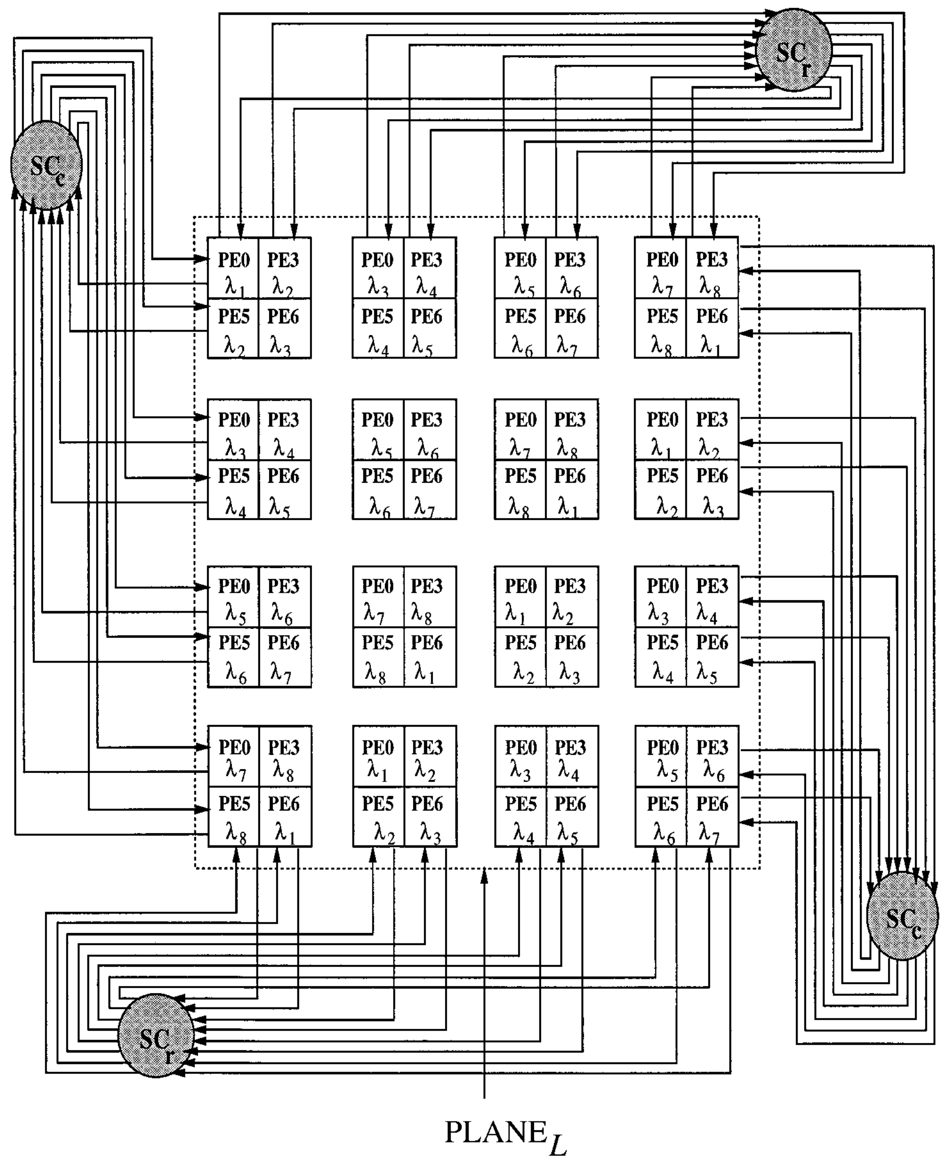


Fig. 4. Optical implementation of PLANE$_L$ of a SBCH(4, 3) network by use of optical star couplers. We need two tunable transmitters–receivers for every node. Similar connections exist for PLANE$_R$. For clarity of the figure only a few buses are shown. Note that each SC$_r$ implements two rowwise buses, while every SC$_c$ implements two columnwise buses.

plane$_l$ of PLANE$_L$ and planes of plane$_r$ of PLANE$_R$. The hypercube modules are implemented by use of free-space optics to provide the connectivity between the planes of plane$_l$ and plane$_r$. Additionally, 2^{n-1} 2-D SBH subnetworks per plane (PLANE$_L$ or PLANE$_R$) need to be implemented. Each 2-D SBH consists of $2w$ buses, therefore a total of $2w \times 2^{n-1}$ buses per PLANE are required.

A trivial implementation of the SBH subnetwork is to assign a distinct wavelength for every PE in PLANE$_L$ and PLANE$_R$ and then to perform WDMA techniques to implement the buses. However, such a straightforward method requires a prohibitively large number of different wavelengths and fibers. For example, for a SBCH(4, 3) consisting of 128 PE's, a total of 64 wavelengths would be necessary. A wavelength-assignment technique[10,16] can be employed to reduce the number of wavelengths used in the system.

Let us take a running example, a SBCH(4, 3). Figure 4 shows how wavelengths are assigned for each PE of PLANE$_L$. The following wavelengths are assigned to the first row: $\lambda_1, \lambda_2, \ldots, \lambda_8$. Then, $\lambda_2, \ldots, \lambda_7, \lambda_1$ are assigned as wavelengths in the second row. In general, wavelength assignment in a row is achieved by rotation of the wavelength assignment of the previous row by one column. This wavelength assignment results in no two PE's in the same row or column of PLANE$_L$ that have an identical wavelength. Similar considerations take place for PE's of PLANE$_R$. With this wavelength-assignment technique, the total number of wavelengths required to implement the SBCH(4, 3) network is reduced from 64 to 8. In general, for a

SBCH(w, n) the following wavelength assignment for the first row must be performed: $\lambda_1, \lambda_2, \ldots, \lambda_{w2^{(n-1)/2}}$, and then $\lambda_2, \ldots, \lambda_{w2^{(n-1)/2}}, \lambda_1$, are assigned to the PE's of the second row, and so on. Thus an implementation of a SBCH(w, n) with the above wavelength assignment requires no more than $w \times 2^{(n-1)/2}$ wavelengths.

With reference to Fig. 4, the wavelengths assigned to the PE's of the first row are divided into two groups of four wavelengths each. The groups are ($\lambda_1, \lambda_3, \lambda_5, \lambda_7$) and ($\lambda_2, \lambda_4, \lambda_6, \lambda_8$). Each of these groups corresponds to the implementation of a rowwise bus. Every PE in the group should be able to tune in to any of the wavelengths assigned to that group. For example, the node of group 1 with wavelength λ_1 must be able to tune to wavelengths λ_3, λ_5, and λ_7, which correspond to wavelengths that were assigned to the other PE's of that group. Rotating the wavelength assignments of the previous rows will form the new wavelength groups that correspond to every row.

Similarly, each column of Fig. 4 must also be divided into two groups of four wavelengths each. For example, for the second column of Fig. 4 the following groups are formed: ($\lambda_2, \lambda_4, \lambda_6, \lambda_8$) and ($\lambda_3, \lambda_5, \lambda_7, \lambda_1$), Each of these wavelength groups corresponds to the implementation of a columnwise bus. Again, rotation of the columnwise wavelength assignment will result in the formation of the wavelength groups for the other columns.

We now consider the overall optical implementation of a SBCH(w, n). For simplicity and without loss of generality we consider the implementation of an example network of size SBCH(4, 3). Figure 4 shows an example PLANE_L of the SBCH(4, 3) network. We assume that each PE has three light sources: one fixed source S_h that illuminates the HOE to generate the required hypercube links and two others S_r and S_c that are relatively tunable sources and are coupled into optical fibers to implement the two spanning buses. It should be noted that full tunability is not required, as explained above. Furthermore, each PE is equipped with three receivers. One fixed receiver, R_h, receives light from the free-space optics implementing the hypercube, and the other two receivers, R_r and R_c, receive light from fibers coming from star couplers. The key component that provides bus connectivity here is the tunable-transmitter–fixed-receiver scheme. The wavelength assignment shown in Fig. 4 corresponds to the receiver wavelength assignment of every PE. Other PE's can communicate with a particular PE by simply tuning in to the wavelength assigned to that PE. Therefore, it is important that tunable devices with sufficient tuning range, as well as tuning time, be available. Rapid progress is being made in the development of tunable devices, both in the range over which they are tunable and in their tuning times.[21,22] Current tuning ranges are of the order of 4–10 nm, and the tuning times vary from nanoseconds to milliseconds.[21]

With reference to Fig. 4, each node utilizes two star couplers, one for each spanning bus. Let each

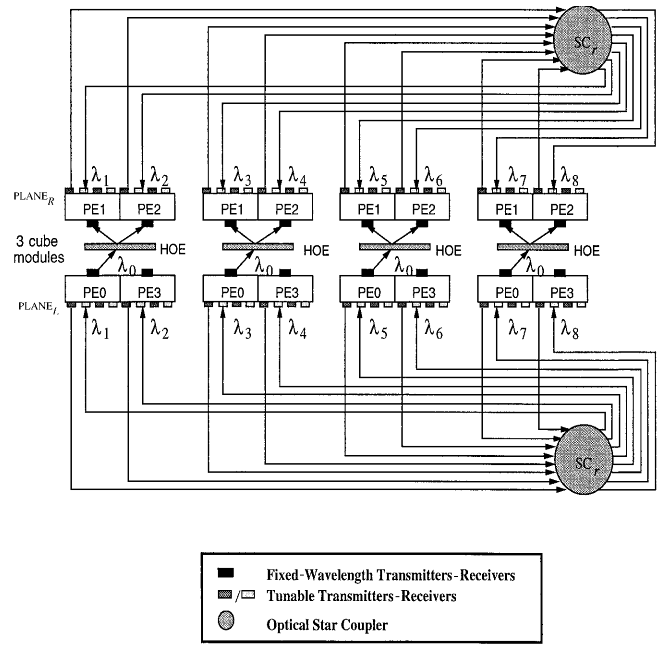


Fig. 5. Top view of a SBCH(4, 3) network. The figure shows the implementation of the first rowwise buses of PLANE_L and PLANE_R. Each star coupler implements two buses of size 4. Similar connections exist for the other rows and columns of the SBCH(4, 3).

star coupler that implements the rowwise buses be denoted by SC_r and each star coupler that implements the columnwise buses by SC_c. In a given SBCH network, a SC_r multiplexes light from S_r sources coming from nodes lying in the same row of the plane, while SC_c multiplexes light from S_c sources coming from nodes lying in the same column of the plane. Note that, instead of using a star coupler for every rowwise or columnwise bus, every star coupler implements $2^{(n-1)/2} = 2$ buses of a $w = 4$ number of nodes. Which rowwise or columnwise buses are implemented is dictated by the wavelength assignment and wavelength grouping, as explained above. For example, the SC_r of the first row of Fig. 4 implements the buses with wavelengths ($\lambda_1, \lambda_3, \lambda_5, \lambda_7$) and ($\lambda_2, \lambda_4, \lambda_6, \lambda_8$). Using a single WDMA channel reduces the number of star couplers required for implementation by a factor of $2^{(n-1)/2}$. In general, a single star coupler implements $2^{(n-1)/2}$ buses.

Figure 5 shows a top view of both planes of the SBCH(4, 3) network. In the middle of the figure the HOE's that implement the hypercube modules are shown. Only two star couplers are visible. The top SC_r implements the two rowwise buses of the first row of PLANE_R, while the bottom SC_r implements the respective buses of the first row of PLANE_L. A total of $2w2^{(n-1)/2}$ star couplers/plane are required. The SBCH(4, 3) network shown in Fig. 4 requires $2 \times 4 \times 2^{(3-1)/2} = 16$ star couplers/plane, resulting in a total of 32 star couplers for both planes. For the case in which each and every bus were to be implemented by use of separate star couplers, the total number for the complete implementation would rise to 64 star couplers.

For alleviating bus collisions (e.g., different messages destined to the same PE at the same time), the time domain along each subchannel can be utilized. Time-division multiple-access techniques can be combined with the proposed WDMA scheme. However, this issue is beyond the scope of this paper.

4. Comparisons with Other Networks that Employ Optical Star Couplers for Their Implementation

When optical star couplers are used to implement an optical network the implementation cost is dominated basically by the number of star couplers and tunable transmitters–receivers needed for the implementation. Recently two star-coupler-based optical interconnection networks have been proposed. Dowd[15] proposed the wavelength-division multiple-access channel hypercube, WMCH(k, n). A WMCH(k, n) network has k PE's along each of n dimensions for a total of k^n PE's. One tunable transmitter and one tunable receiver per PE and per dimension are required. All PE's along each dimension are connected by means of a passive optical star coupler with the WDMA technique.

The WMCH network is essentially an optical version of the conventional SBH network. Therefore they have similar network characteristics and performance. Liu *et al.*[16] proposed an optical interconnection network called a dBus-array(k, n) as a hardware improvement to the WMCH. A dBus-array(k, n) consists of an n-dimensional array with k^n PE's and k^{n-1} unidirectional buses that are also implemented by use of optical passive star couplers. The dBus-array(k, n) required one tunable transmitter and one tunable receiver per PE, compared with n tunable transmitters and n tunable receivers per PE for the WMCH. In what follows, we compare the hardware costs of these two networks with that of the SBCH.

Figure 6(a) compares the SBCH(8, n) network with the WMCH(8, n) and the dBus-array(8, n) networks in terms of the number of star couplers as the networks grow in size. The SBCH(8, n) grows by use of the fixed$_{w=8}$ rule, while the other two networks grow by use of the fixed$_{k=8}$ rule. Figure 6(b) shows a comparison of the SBCH(w, 3), WMCH(k, 3), and dBus-array(k, 3) networks as they grow in size by use of the fixed$_{n=3}$ scaling rule. Figures 6(c) and 6(d) show the comparison in terms of the number of tunable transmitters–receivers used. As the figures show, the SBCH network requires fewer optical star couplers and a smaller number of tunable transmitters–receivers than do the WMCH and the dBus-array(8, n) networks. Hence the SBCH network offers reduced implementation costs compared with the WMCH and the dBus-array(8, n) while providing performance advantages. We note that Szymanski describes a star-based optical network called Hypermesh.[23] This research was not available to us at the time of this study. Comparison of the SBCH and the Hypermesh networks is under way and is the subject of a separate paper.

5. Power-Loss Analysis of the Optical Implementation

In this section we present some system-noise calculations to investigate the bit-error rate (BER) characteristics of the proposed optical implementation of the SBCH network. Calculation of the BER of an optical system requires the estimation of the signal-to-noise ratio (SNR). Estimation of total power losses, leading into receiver-sensitivity calculations is required to obtain the SNR. In what follows the optical power loss of the implementation methodology is calculated. Then receiver sensitivity is estimated, and consequently the BER of the proposed implementation is evaluated.

The number of PE's that an optical system can support is determined by the emitting power of the transmitter, the required receiver sensitivity, and the losses occurring between the transmitter and the receiver. Let L_{sf} be the source-to-fiber coupling loss, L_{fd} be the fiber-to-detector coupling loss, and L_f be the fiber-attenuation loss. We also assume that all PE's are equidistant. Let L_e be the excess loss of the optical star coupler. To estimate the star-coupler splitting loss, it is necessary to know the input power to the coupler and the fan-out. Let P_{in} be the power coming into the coupler from an input channel and P_{out} be the power coming out from an output channel; then $L_{\text{sp}} = 10 \log (P_{\text{out}}/P_{\text{in}})$.

The total transmission loss is then

$$L_{\text{total}} = L_e + L_{\text{sp}} + dL_f + L_{\text{sf}} + L_{\text{fd}}. \tag{1}$$

P_{out} is equal to P_{in}/k, where k is the fan-out of the star coupler. For the SBCH, k is equal to $w \times 2^{(n-1)/2}$ (number of PE's in a row or column of PLANE$_L$ or PLANE$_R$). Consequently, Eq. (1) can be rewritten as

$$\begin{aligned} L_{\text{total}} &= L_e - 10 \log k + dL_f + L_{\text{sf}} + L_{\text{fd}} - 3 \\ &= L_e - 10 \log w - \frac{3}{2}(n-1) + dL_f + L_{\text{sf}} \\ &\quad + L_{\text{fd}} - 3. \end{aligned} \tag{2}$$

To estimate the total loss of the optical system, we consider values from commercially available components. We assume laser diode sources with a characteristic +7 dBm. Also, the insertion loss for a commercially available fiber coupler is taken to be −1 dB, while fiber-to-detector losses are $L_{fd} = -0.46$ dB. The fiber loss is taken to be $L_f = 0.3$ dB/km, but since d is of the order of centimeters the total fiber loss is negligible. In addition, a −3-dB loss is added for engineering errors. Rearranging Eq. (2) and using the above values determine the number of PE's supported by the star couplers, given a desirable BER. For a desired 10^{-17} BER the required receiver sensitivity of the GaAs metal-Semiconductor field-effect-transistor transimpedance[24] can be calculated to be −19.2 dBm. For laser diodes of 7.0-dBm power the total loss in the optical system should be −26.2 dB,

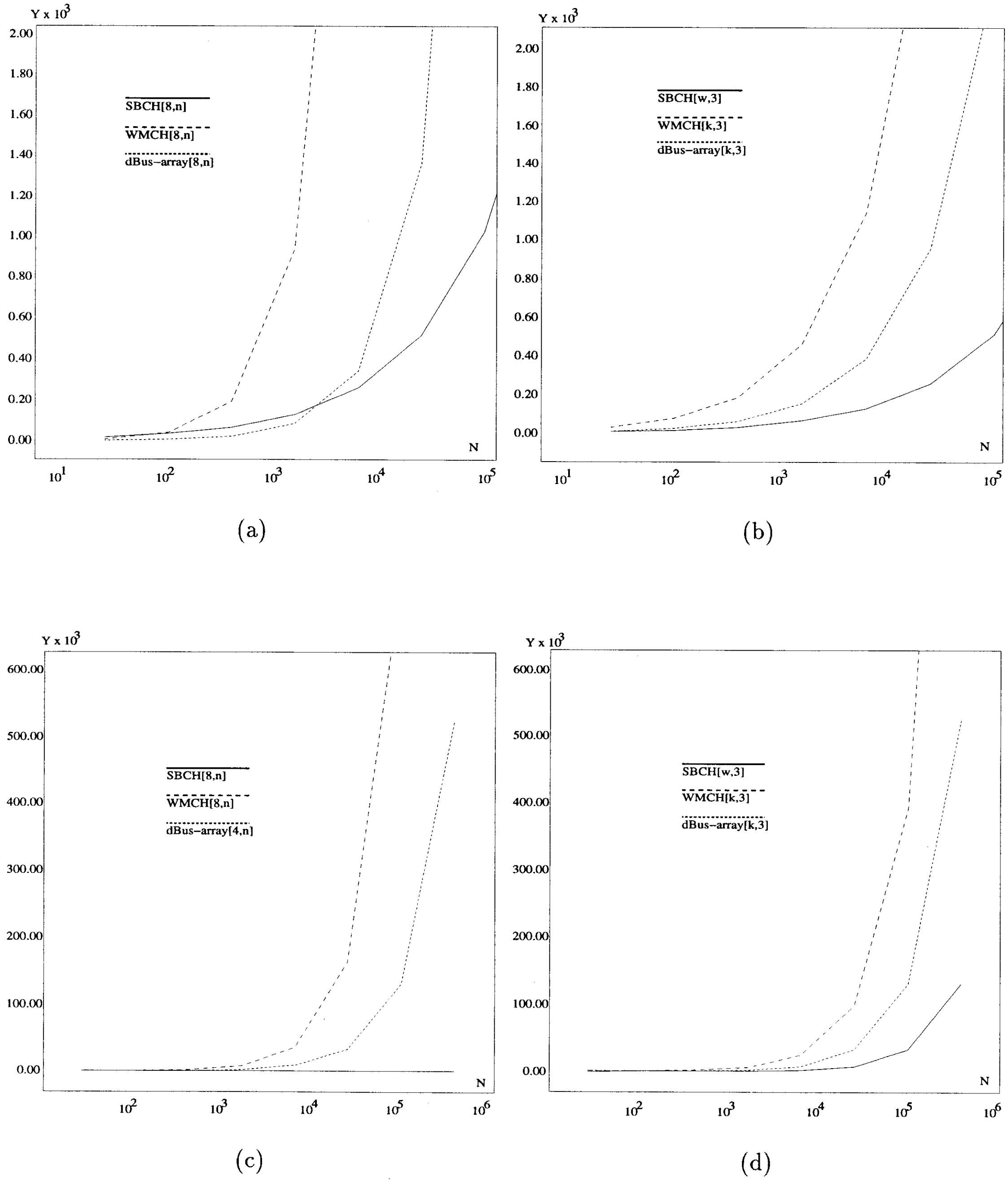


Fig. 6. Comparison of the number of optical star couplers and number of tunable transmitters or receivers among the SBCH, dBus-array(k, n), and WMCH networks. (a) and (b) show the results for optical star couplers, and (c) and (d) show the results for the number of tunable transmitters or receivers.

yielding a star-coupler fan-out of $k = 118$. This value is within the capabilities of current star-coupler technology. The optical fan-out of star couplers reported to date is 128×128.[21] For $k = 118$, large SBCH networks are feasible. For example, a SBCH(30, 5) network supporting ~28,800 PE's could be implemented.

6. Conclusions

In this paper, we proposed a novel hybrid network that significantly improves hypercube-based topologies in general and the SBH and OMMH networks in particular. The key attractive features of the proposed network include the possibility of a constant-diameter and a constant-degree network, while it is feasible to interconnect thousands of processors at a reasonable cost. Additionally, the network is incrementally scalable and fault tolerant. Theses features make the SBCH very suitable for massively parallel systems. A WDMA technique has been proposed for the optical implementation of the SBCH network. Analysis of the implementation reveals that a greater than 20,000-PE SBCH network with a BER of less than 10^{-17} is currently feasible.

This research is supported by National Science Foundation grant MIP-9310082.

References

1. K. Hwang, *Advanced Computer Architecture: Parallelism, Scalability, Programmability* (McGraw-Hill, New York, 1993).
2. H. J. Siegel, *Interconnection Networks for Large-scale Parallel Processing* (McGraw-Hill, New York, 1990).
3. H. S. Stone and J. Cocke, "Computer architecture in the 1990s," Computer **24,** 30–38 (September 1991).
4. A. Louri and H. Sung, "3D optical interconnects for high-speed interchip and interboard communications," Computer **27,** 27–37 (October 1994).
5. Y. Saad and M. H. Schultz, "Topological properties of hypercubes," IEEE Trans. Comput. **37,** 867–872 (1988).
6. Q. M. Malluhi and M. A. Bayoumi, "The hierarchical hypercube: a new interconnection topology for massively parallel systems," IEEE Trans. Parallel Distrib. Sys. **5,** 17–30 (1994).
7. C. Chen, D. P. Agrawal, and J. R. Burke, "dBCube: a new class of hierarchical multiprocessor interconnection networks with area efficient layout," IEEE Trans. Parallel Distrib. Sys. **4,** 1332–1344 (1993).
8. K. Efe, "The crossed cube architecture for parallel computation," IEEE Trans. Parallel Distrib. Sys. **3,** 513–524 (1992).
9. J. R. Goodman and C. H. Sequin, "Hypertree: a multiprocessor interconnection topology," IEEE Trans. Comput. **C-30,** 923–933 (1981).
10. A. Louri and H. Sung, "Scalable optical hypercube-based interconnection network for massively parallel computing," Appl. Opt. **33,** 7588–7598 (1994).
11. A. Louri and H. Sung, "An optical multi-mesh hypercube: a scalable optical interconnection network for massively parallel computing," J. Lightwave Technol. **12,** 704–716 (1994).
12. A. Louri and S. Furlonge, "Feasibility study of a scalable optical interconnection network for massively parallel processing systems," Appl. Opt. **35,** 1296–1308 (1996).
13. L. D. Wittie, "Communication structures for large networks of microcomputers," IEEE Trans. Comput. **C-30,** 264–273 (1981).
14. G. Lerman and L. Rudolph, *Parallel Evolution of Parallel Processors* (Plenum, New York, 1993).
15. P. W. Dowd, "Wavelength division multiple access channel hypercube processor interconnection," IEEE Trans. Comput. **41,** 1223–1241 (1992).
16. Y. Li, A. W. Lohmann, and S. B. Rao, "Free-space optical mesh-connected bus networks using wavelength-division multiple access," Appl. Opt. **32,** 6425–6437 (1993).
17. V. Kumar, A. Grama, A. Gupta, and G. Karypis, *Introduction to Parallel Computing: Design and Analysis of Algorithms* (Benjamin Cummings, New York, 1994).
18. D. A. Reed and R. M. Fujimoto, *Multicomputer Networks: Message-Based Parallel Processing* (MIT Press, Cambridge, Mass., 1987).
19. A. Louri and C. Neocleous, "A spanning bus connected hypercube: a new scalable optical interconnection network for multiprocessors and massively parallel systems," IEEE J. Lightwave Technol. **15,** 1241–1253 (1997).
20. P. Sweazey, "Limits of performance of backplane buses," in *Digital Bus Handbook* (McGraw-Hill, New York, 1990).
21. A. Cisneros and C. A. Brackett, "A large ATM switch based on memory switches and optical star couplers," IEEE J. Select. Areas Commun. **9,** 1348–1360 (1991).
22. P. F. Moulton, "Tunable solid state lasers," Proc. IEEE **80,** 348–364 (1992).
23. T. Szymanski, "Hypermeshes: optical interconnection networks for parallel computing," J. Parallel Distributed Comput. **26,** 1–23 (1995).
24. T. V. Moui, "Receiver design for high-speed optical fiber systems," IEEE J. Lightwave Technol. **LT-2,** 243–267 (1984).