PAPER *****

A Class of Benes-based Optical Multistage Interconnection Networks for Crosstalk-free Realization of Permutations

Xiaohong Jiang[1],Pin–Han Ho[2],Hong Shen[3] and Susumu Horiguchi[1], Regular member

Summary

Vertical stacking is a novel technique for creating nonblocking (crosstalk-free) optical multistage interconnection networks (MINs). In this paper, we propose a new class of optical MINs, the vertically stacked Benes (VSB) networks, for crosstalk-free realization of permutations in a single pass. An $N \times N$ VSB network requires at most O(NlogN) switching elements, which is the same as the Benes network, and much lower overall hardware cost than that of the existing optical MINs built on the combination of horizontal expansion and vertical stacking of banyan networks, to provide the same crosstalk-free permutation capability. Furthermore, the structure of VSB networks provides a more flexible way for constructing optical MINs because they give more choices in terms of the number of stages used in an optical MIN. We also present efficient algorithms to realize crosstalk-free permutations in an $N \times N$ VSB network in time O(NlogN), which matches the same bound as required by the reported schemes.

Key words: *Benes network, optical crosstalk, optical switch, banyan network, rearrangeably nonblocking.*

1 Introduction

Optical multistage interconnection networks (MINs) provide an attractive way to meet the increasing demands on high channel bandwidth and low communication latency for high-performance applications. Driven by the requirements of promising Optical Burst Switching (OBS) and Optical Packet Switching (OPS) technologies, one of the biggest challenges is to design high-speed optical MINs that can switch fast (e.g., tens of nanoseconds or less). The optical micro-electromechanical systems (MEMS) are attractive for building large-scale optical switches, but they switch only at the speed of milliseconds due to their inherent mechanical limit of switching speed [23]. Directional-coupler (DC) [6] is an electro-optical device implemented by manufacturing two waveguides close to each other, and a DC is similar to a 2×2 electrical switching element (SE) with both crossing and parallel states. A DC has the capability to simultaneously switch optical flows with the speed of sub-nanoseconds and with multiple wavelengths, which makes it one of the promising candidates for building future OBS and OPS enabled high-speed optical switches. Crosstalk in DCs is a major shortcoming in DC-based optical switches, which occurs between two signals carried in the two waveguides of a coupler [3,6]. When two optical signals pass through a DC, a portion of optical power in one waveguide is coupled into another waveguide, and this undesirable coupling at the DC is called the first-order crosstalk. Due to the stringent bit-error rate requirement of optical transmission facilities, elimination of crosstalk has become an important issue for making optical networks function properly [3,12,13,14]. By guaranteeing that only one input of a DC is busy while keeping the other idle at any moment, the first-order crosstalk in the DC can be avoided. This provides a cost-effective approach of solving the crosstalk problem in DC-based optical switching networks. In this paper, we focus on the design of optical MINs that are free of first-order crosstalk in each SE (we refer to this quality as 'crosstalk-free' hereafter). It is notable that the consideration of crosstalk-free constraint requires that all optical signals passing through a MIN never share a common SE in transmission (i.e., they should follow SE-disjoint paths in transmission).

Banyan [5] and its topologically equivalent (e.g. *baseline, omega*) networks [8,15] belong to a class of MINs with important applications due to their simple switch setting ability (self-routing) and small number of SEs along a path between an input-output pair. When the class of MINs is constructed as a DC-based optical switching network, a good characteristic can be yielded because the crosstalk and signal loss of an input optical flow are proportional to the number of DCs taken by the optical flow. Thus, the optical MINs based on banyan networks have been of great interest by researchers and subject to extensive study in the past years [13,17,18,20]. Since there is only a single path connecting an input-output pair in a banyan network, the banyan network is a blocking network. A nonblocking MIN can be constructed by either appending some extra stages to the back of a regular banyan network (*horizontal expansion*) or vertically stacking multiple copies (planes) of the banyan network (*vertical stacking*) [9,13].

[1] The authors were with the Graduate School of Information Sciences, JAIST, Japan. They are now with the Graduate School of Information Sciences, Tohoku University, Japan
[2] The author is with the Department of Electrical and Computer Engineering, University of Waterloo, Ontario, Canada.
[3] The author is with the Graduate School of Information Sciences, JAIST, Japan.

There are three types of nonblocking networks, namely strictly nonblocking, wide-sense nonblocking and rearrangeably nonblocking [2,16]. In a strictly nonblocking network, any input can be routed to any unused output regardless of the way the other input signals are routed. However, this high degree of flexibility is gained at the expense of taking higher hardware cost in term of the number of SEs required. Under the crosstalk constraints, strictly nonblocking optical networks built by combining both the horizontal expansion and the vertical stacking of banyan network have been explored recently [17,18]. In a wide-sense nonblocking network, a rule of establishing a connection must be defined so as to keep the switch nonblocking. Some results for this class of switch can be found in [19]. Rearrangeably nonblocking is an interesting alternative other than strictly nonblocking and wide-sense nonblocking for optical switching networks. This kind of network can route any idle input to any unused output, but one or more existing connections may have to be rerouted to guarantee the switch nonblocking. Since the cost and signal degradation of a rearrangeable optical network are always lower than its strictly nonblocking and wide-sense nonblocking counterparts, the rearrangeably nonblocking optical networks are promising for the future applications [13] and will be focused on in this paper.

The condition of being rearrangeable and crosstalk-free has been determined for the switching network composing of both vertical stacking and horizontal expansion (VH) of banyan networks [13]. Since the amount of signal loss and crosstalk in an optical MIN is proportional to the number of DCs traversed by the input optical flow, reducing the number of stages between input-output pairs in the network is one of the important design objectives. Nonetheless, the redundancy of SEs and the total number of SEs will be dramatically increased if the number of stages in a VH network is reduced [13]. For example, the minimum number of stages of a rearrangeable $N \times N$ VH network, $\log_2 N$, can be achieved by vertically stacking $2^{\lfloor (n+1)/2 \rfloor}$ (here $n = \log_2 N$) copies of an $N \times N$ banyan network, resulting in the number of SEs at the order of ($N^{3/2}\log N$), which is considerably higher than the $O(N\log N)$ switching complexity of a rearrangeable VH network when the number of stages is at its maximum possible value $2\log_2 N$-1.

In this paper, we propose a new class of optical multistage interconnection networks, called vertically stacked Benes (VSB) network, for crosstalk-free realization of a permutation in a single pass. A VSB(N,K) network has N inputs and N outputs and consists of K vertically stacked copies (planes) of ($2N/K$)× ($2N/K$) Benes network. Compared to the reported $N \times N$ rearrangeable optical MINs, the proposed VSB(N,K) network yields the following advantages:

- A VSB(N,K) network yields a maximum switching complexity of $O(N\log N)$, which meets the lower bound of the switching complexity in an $N \times N$ rearrangeable VH network. The former, however, takes a much lower overall hardware cost than that by the latter for the same permutation capability.
- A VSB(N,K) network adopts a more flexible structure than that of an $N \times N$ rearrangeable VH network in the sense that the number of stages can vary from 1 to $2\log_2 N$-1 in the former rather than from $\log_2 N$ to $2\log_2 N$-1 as in the latter.
- Crosstalk-free permutation in Benes network can be considered as a special case of that in a VSB(N,2) network when we allow *time division multiplexing* (TDM) in signal transmission and let each stacked plane of VSB(N,2) network correspond to one pass.

Based on the proposed VSB(N,K) structure, a scheme for crosstalk-free realization of permutations is introduced. The basic idea of this scheme is to first decompose a permutation into multiple partial permutations by applying the Euler-split technique for coloring bipartite graphs [4], followed by realizing each partial permutation crosstalk-free within one plane of the VSB(N,K) network. The scheme has a time complexity of $O(N\log N)$, which is the same as that for crosstalk-free realization of a permutation in a Benes network or a VH network [7].

2 MINs Built on Vertical Stacking and Horizontal Expansion of Banyan Network

Based on an $N \times N$ banyan network composed of $n = \log_2 N$ stages, an $N \times N$ rearrangeable MIN can be built by applying either a Vertical Stacking (VS) technique or a Horizontal Expansion (HE) technique [9]. A VS network is constructed by vertically stacking K copies of the banyan network and connecting each input (output) through a 1×K splitter (K×1 combiner). On the other hand, an HE network, denoted by H(n,m), constructed by appending the reversal of the first m ($m \le n$-1) stages of the banyan network to the back of the network. More generally, a rearrangeable MIN can be built by combining the VS and HE techniques in which an H(n,m) network is first created from a regular banyan network and the H(n,m) network is then vertically replicated K times [13]. The resulting network is denoted as VH(n,m,K). Fig. 1 illustrates VH (4,2,2).

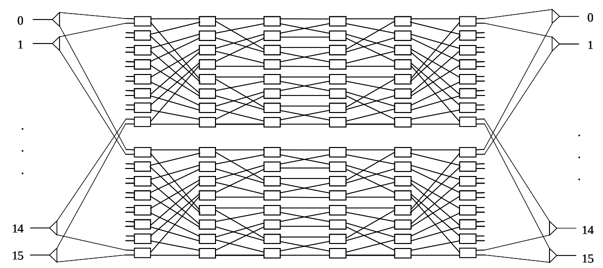

Fig.1 The construction of a VH (4,2,2) network from banyan networks.

Ideally, we are interested in designing an optical switching network without any crosstalk. The following result relating to the rearrangeable condition for a VH(n,m,K) network was given in [13].

Theorem 1: Under constraint of crosstalk-free, a VH(n,m,K) network is rearrangeable if and only if

$$K = 2^{\lfloor (n-m+1)/2 \rfloor} \quad (1)$$

Hereafter, we use RVH(N,m) to refer to the rearrangeable $N\times N$ network VH(n,m,K) with $K=2^{\lfloor (n-m+1)/2 \rfloor}$. The above result indicates that any permutation can be realized rearrangeably nonblocking and free of crosstalk in a RVH(N, m) network. In [13], an algorithm was also proposed to realize a permutation crosstalk-free in a RVH(N,m) network. The basic idea of this algorithm is to first decompose the permutation evenly into $2^{\lfloor (n-m+1)/2 \rfloor}$ partial permutations of size $N/2^{\lfloor (n-m+1)/2 \rfloor}$, and then realize each partial permutation crosstalk-free in a single plane of the RVH(N,m) network. Note that a single plane of a RVH(N,m) network is an H(n,m) network that can accommodate only one partial permutation. To meet the crosstalk-free requirement, only one signal is allowed to pass through a switch at a time. Then it is easy to verify that the total number of SEs, $M_{SE,}$ and the SE redundancy, $R_{SE,}$ (i.e., the number of unused SEs) in a RVH(N,m) network are given by:

$$M_{SE} = 2^{\lfloor (n-m+1)/2 \rfloor} \cdot \frac{N}{2} \cdot (n+m) \quad (2)$$

$$R_{SE} = 2^{\lfloor (n-m+1)/2 \rfloor} \cdot \left(\frac{N}{2} - \frac{N}{2^{\lfloor (n-m+1)/2 \rfloor}} \right) \cdot (n+m) \quad (3)$$

$$n = \log_2 N,\ 0 \le m \le n-1$$

Since an optical MIN with a small number of stages is desirable for the sake of minimizing crosstalk and signal loss, we may reduce the number of stages $\log_2 N + m$ in a RVH(N,m) network by reducing the number of extra stages m in the network. However, results (2)-(3) indicate that if we do so, both the M_{SE} and R_{SE} will increase dramatically as illustrated in Fig.2 for a 4096×4096 network. This indicates that constructing a large optical MIN with few stages by combining vertical stacking and horizontal expansion of banyan networks is impractical.

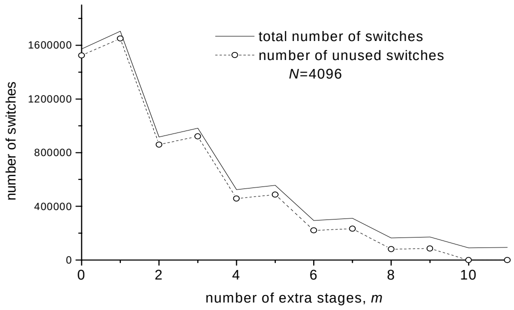

Fig. 2. Total number and the number of unused SEs in a 4096 × 4096 RVH (n,m) network with different extra stages.

3 MINs Built on Vertically Stacked Benes Networks

The large number of SEs in a RVH(N,m) network comes from the high SEs redundancy in the network because each of its $N \times N$ H(n, m) networks accommodates only one partial permutation of size $N/2^{\lfloor (n-m+1)/2 \rfloor}$. Note that we can realize any permutation crosstalk-free in an optical Benes network in two passes with zero SEs redundancy [11,21]. Motivated by the above observations, we propose a novel switch architecture in this study for implementing an $N \times N$ rearrangeable network by applying the vertical stacking technique to Benes networks instead of using banyan networks as in RVH(N,m). The proposed switch architecture yields a new class of optical MINs, namely vertically stacked Benes networks (denoted by VSB(N,K)), which is constructed by vertically stacking K ($K\le N$ is the power of 2) copies of a $(2N/K)\times(2N/K)$ Benes network. The connection between N inputs of a VSB(N,K) and its K Benes planes is implemented by using N $1\times K$ splitters and N $K\times 1$ combiners, in which each input of the VSB(N,K) network is connected to a splitter, and one inlet out of the two in a Benes network is connected to a combiner. Let the splitters in the VSB(N,K) network be numbered from top (number 0) to bottom (number N-1), and the number of the combiners in each $(2N/K)\times(2N/K)$ Benes of the network be numbered from top (number 0) to bottom (number N/K-1). The connection between the splitters and combiners will be implemented in such a way that the K outputs of the i-th splitter are linked to the K $\lfloor i/K \rfloor$-th combiners in K different Benes networks and symmetrically the connection between N outputs of a VSB(N,K) and its K Benes planes can be implemented by using N $K\times 1$ combiners and N $1\times K$ splitters. Fig. 3 illustrates a sample of VSB(N,K), VSB(16,4).

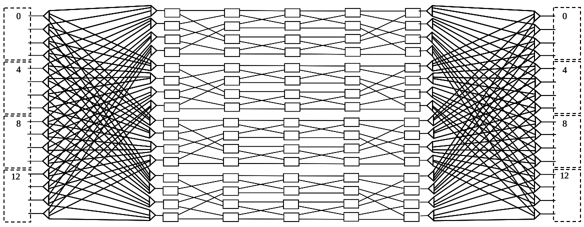

Fig.3 The construction of a VSB (16,4) network from 8 × 8 Benes networks.

We will show in the next section that VSB(N,K) architecture has the capability of supporting any permutation of an N-element set {0,1,…, N-1} rearrangeably-nonblocking and crosstalk-free. A performance comparison of VSB network with existing RVH networks is presented in Section 5. One interesting special case of VSB(N,K) networks in the VSB(N,2) network that is composed of two $N \times N$ Benes networks and has the capability of realizing any permutation rearrangeably nonblocking and crosstalk-free as discussed in [13]. However, if we increase the number of planes K to N in a VSB(N,K) network, we can achieve a VSB(N,N) network consisting of only splitters and combiners so that a path does not need to be switched to go through the network. In such a situation, each input in a VSB(N,N) network can be assigned a dedicated path through the network to each output, and thus a connection request can be routed from any input to any unused output regardless of the way by which the other existing are routed. So the network VSB(N,N) becomes another attractive special case of VSB(N,K) structure because it is actually the cascade of two strictly nonblocking Spanke architectures[22]. The structure of VSB(4,4) network is shown in Fig.4.

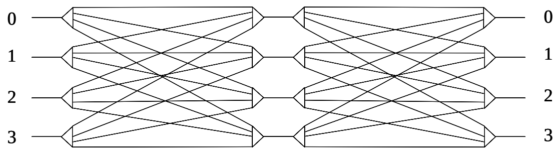

Fig. 4. The structure of strictly nonblocking VSB (4,4) network.

4 Crosstalk-free Permutation in VSB (N,K) Network

In this section, we first prove that we can realize any permutation of an N-element set {0,1,…, N-1} rearrangeably-nonblocking and crosstalk-free in a VSB(N,K) network. Based on such a property, an efficient permutation scheme is introduced. The proposed permutation scheme consists of two integrated parts: first, the decomposition of a permutation into crosstalk-free feasible partial permutations (CFPPs) by applying the well-known *Euler-split* technique for coloring a bipartite graph; second, the realization of each CFPP rearrangeably nonblocking and crosstalk-free in each stacked plane of a VSB(N,K) network.

4.1 Crosstalk-free Feasible Partial Permutation (CFPP)

To realize a permutation in a VSB(N,K) network, the task of a full permutation is evenly distributed to K stacked Benes networks by decomposing the permutation into K partial permutations, each of which is realized upon a ($2N/K$)× ($2N/K$) Benes in the VSB(N,K) network. For convenience, we introduce the following definition.

Definition 1 A partial permutation $P = \begin{pmatrix} x_0, x_1, \dots, x_{N/K-1} \\ y_0, y_1, \dots, y_{N/K-1} \end{pmatrix}$ for an $N \times N$ VSB(N,K) network, where input x_i is mapped to output y_i, with $x_i, y_i \in \{0,1,\dots,N-1\}$ and $x_0<x_1<\dots<x_{N/K-1}$, is referred to as a crosstalk-free feasible partial permutation (CFPP) to the VSB(N,K) network if:

$$\left\{\left\lfloor \frac{x_0}{K} \right\rfloor, \left\lfloor \frac{x_1}{K} \right\rfloor, \dots, \left\lfloor \frac{x_{N/K-1}}{K} \right\rfloor\right\} = \left\{0,1,\dots,\frac{N}{K}-1\right\}$$

$$\left\{\left\lfloor \frac{y_0}{K} \right\rfloor, \left\lfloor \frac{y_1}{K} \right\rfloor, \dots, \left\lfloor \frac{y_{N/K-1}}{K} \right\rfloor\right\} = \left\{0,1,\dots,\frac{N}{K}-1\right\} \quad (4)$$

Note that if we divide both the inputs and outputs of a VSB(N,K) network into K-element sets $I_j=\{K\cdot j, K\cdot j+1,\dots, K\cdot j+K-1\}$ and $O_j=\{K\cdot j, K\cdot j+1,\dots, K\cdot j+K-1\}$ for $0\le j\le N/K-1$, with I_j and O_j corresponding to inputs and outputs, respectively, partial permutation P is a CFPP to the network if and only if each input and each output of the partial permutation falls within a distinct set I_j and O_j for $0\le j\le N/K-1$, respectively. It will be shown in Section 4.3 that a CFPP to a VSB(N,K) network is actually a partial permutation that is crosstalk-free realizable in each of its planes.

Example 1 For a VSB(16,4) network, the partial permutation $\begin{pmatrix} 3 & 4 & 9 & 13 \\ 12 & 4 & 1 & 10 \end{pmatrix}$ is a CFPP, since N =16, K =4 and we have

$$\left\{\left\lfloor \frac{3}{4} \right\rfloor, \left\lfloor \frac{4}{4} \right\rfloor, \left\lfloor \frac{9}{4} \right\rfloor, \left\lfloor \frac{13}{4} \right\rfloor\right\} = \{0,1,2,3\} \text{ and}$$

$$\left\{\left\lfloor \frac{12}{4} \right\rfloor, \left\lfloor \frac{4}{4} \right\rfloor, \left\lfloor \frac{1}{4} \right\rfloor, \left\lfloor \frac{10}{4} \right\rfloor\right\} = \{3,1,0,2\} = \{0,1,2,3\}.$$

4.2 Decomposition of Permutations into CFPPs

We will show in this section that any permutation of an N-element set {0,1,…, N-1} can be evenly decomposed into

K CFPPs for a VSB(N,K) network, and then present an algorithm for such decomposition.

4.2.1 CFPP Decomposability of Permutations

Lemma 1: Any permutation of an N-element set {0,1,…, N-1} can be decomposed into K CFPPs for a VSB(N,K) network.

Proof. Let the permutation for an $N \times N$ VSB(N,K) be the form $\begin{pmatrix} x_0, x_1, \ldots, x_{N-1} \\ y_0, y_1, \ldots, y_{N-1} \end{pmatrix}$, where $x_l = l$ for $0 \le l \le N$-1 and $\{y_0, y_1, \ldots, y_{N-1}\} = \{0,1,\ldots,N-1\}$. We decompose the permutation evenly into N/K partial permutations each of which has K elements

$$\begin{pmatrix} x_0 \ldots x_{K-1} \\ y_0 \ldots y_{K-1} \end{pmatrix}, \begin{pmatrix} x_K \ldots x_{K+K-1} \\ y_K \ldots y_{K+K-1} \end{pmatrix}, \ldots, \begin{pmatrix} x_{(N/K-1)K} \ldots x_{N-1} \\ y_{(N/K-1)K} \ldots y_{N-1} \end{pmatrix} \quad (5)$$

By applying P.Hall's distinct system representatives theorem [1] recursively, the proof of this Lemma becomes an extension to any K of the proof reported in [21] for the particular case K=2.

QED.

4.2.2 CFPP Decomposition Algorithm

Lemma 1 guarantees the correctness of the CFPP decomposition of a permutation. As indicated in [13], a CFPP decomposition algorithm can be easily obtained by the repetition of the simple bi-partite graph coloring procedure with two colors [4]. A high-level description of the complete CFPP decomposition algorithm can be summarized as:

CFPP Decomposition Algorithm for VSB(N,K) network:

Initiate: i = 0 and take the permutation as the 0-level partial permutation.

Step 1: If $i = \log_2 K$, exit.

Step 2: For each i-level partial permutation, do step 3 and step 4.

Step 3: Construct the bipartite graph $G = (V_1, V_2 ; E)$ for the i-level partial permutation. The vertex sets of G are defined as:

$$V_1 = \left\{ I_0, I_1, \ldots, I_{\frac{N}{2^{i+1}}-1} \right\},\ V_2 = \left\{ O_0, O_1, \ldots, O_{\frac{N}{2^{i+1}}-1} \right\}$$

Here $I_j = \{2^{i+1} \cdot j,\ 2^{i+1} \cdot j+1, \ldots, 2^{i+1} \cdot j + 2^{i+1} - 1\}$ and $O_j = \{2^{i+1} \cdot j,\ 2^{i+1} \cdot j+1, \ldots,\ 2^{i+1} \cdot j + 2^{i+1} - 1\}$ for $0 \le j \le N/2^{i+1} - 1$, with I_j and O_j corresponding to inputs and outputs, respectively. The edge set E is defined as: for any input-output pair (x_j, y_j) in the i-level partial permutation, if $x_i \in I_{j1}$ and $y_i \in O_{j2}$, then there is an edge between vertex I_{j1} and vertex O_{j2} in E.

Step 4: For each connected component of G, start from a vertex of this component in V_1, traverse through an unvisited edge to the neighbor vertex in V_2, back and forth until returning to the starting vertex. During the traversing, a visited edge will be put into E_1 if the traverse direction on this edge is from V_1 to V_2; and is put into E_2 if the direction is in the opposite.

Step 5: Take all one-pair mappings corresponding to the edges in E_1, to form one (i+1)-level partial permutation corresponding to the i-level partial permutation; let the remaining one-pair mappings, corresponding to the edges in E_2, form another (i+1)-level partial permutation corresponding to the i-level partial permutation.

Step 6: $i \Leftarrow i + 1$. Go to Step 1.

It is clear that after running Steps 2-5 for an i-level partial permutation, two (i+1)-level partial permutations will be obtained correspondingly with each input/output of a (i+1)-level partial permutation falling within a distinct set $\{2^{i+1} \cdot j, 2^{i+1} \cdot j+1, \ldots, 2^{i+1} \cdot j + 2^{i+1} - 1\}$ for $0 \le j \le N/2^{i+1} - 1$. Thus, after running the decomposition algorithm over a full permutation for a VSB(N,K) network, the permutation will be decomposed into K partial permutations where each input (output) of a partial permutation is within a distinct set $\{K \cdot j, K \cdot j+1, \ldots, K \cdot j + K - 1\}$ for $0 \le j \le N/K - 1$. So these K partial permutations obtained are just K CFPPs for the VSB(N,K) network. It is clear that the time to construct the bipartite graph is proportional to the number of input-output pairs in the permutation which increases linearly with N, so the time to traverse all edges is $O(N)$.Since Steps 2-5 take $O(N)$ steps and these steps will repeat $O(\log_2 K)$ times, thus, the time complexity of the decomposition algorithm is $O(N \log K)$.

4.3 Realizing CFPP in Benes Network

A Benes network is constructed by appending the reversal of an $N \times N$ banyan network to the back of the banyan network with the central stages overlapped [2]. By the property of symmetry of Benes networks, we can also define a Benes network in a recursive way as shown in Fig.5.

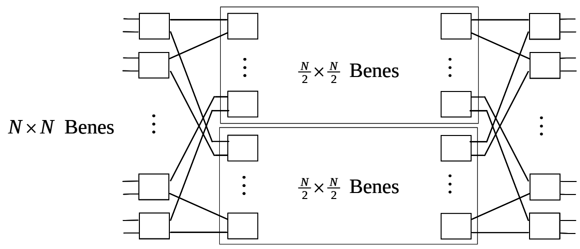

Fig. 5 Recursive definition of an $N \times N$ Benes network.

Fig. 5 shows that in a Benes network, an outlet of an input switch is connected to an inlet of the upper sub-Benes network and another outlet of the input switch is connected to an inlet of the lower sub-Benes network, respectively. Output switches have a symmetric connection pattern. By the definition of CFPPs for a VSB(N,K) network, it is easy to see that if we realize a CFPP in one $(2N/K) \times (2N/K)$ Benes of the VSB(N,K) network, there is only one active link passing through each input/output switch of the Benes and thus the CFPP is crosstalk-free in the first and last stages in the Benes. By applying the *Euler-split* technique for coloring a bipartite graph [4] to decompose the CFPP into two sub-CFPPs so as to realize each sub-CFPP in one sub-Benes of the ($2N/K$)× ($2N/K$) Benes, we can guarantee that there is only one active link passing through each input/output switch of each sub-Benes. Applying this process recursively, we can finally realize a CFPP rearrangeably nonblocking and crosstalk-free in one stacked Benes of the VSB(N,K) network, as discussed in [10,11,13,21]. A high-level description of the complete CFPP routing algorithm in one ($2N/K$)× ($2N/K$) Benes of a VSB(N,K) network can then be obtained by slightly modifying the routing algorithm reported in [21]:

Routing Algorithm (Routing of a CFPP in one $L\times L$ Benes of a VSB(N,K) network, here $L=2N/K$)

Step 1: If L is 2, make the connection of the CFPP in 2×2 switch according to the CFPP; exit.

Step 2: Construct the bipartite graph $G = (V_1,V_2\ ;\ E)$ corresponding to the CFPP in the $L\times L$ Benes. The vertex sets of G are defined as:

$$V_1 = \left\{I_0, I_1, \ldots, I_{\frac{L}{2}-1}\right\},\ V_2 = \left\{O_0, O_1, \ldots, O_{\frac{L}{2}-1}\right\}$$

Here $I_j=\{2\cdot j,\ 2\cdot j+1\}$ and $O_j=\{2\cdot j,\ 2\cdot j+1\}$ for $0\le j \le L/2-1$, with I_j and O_j correspond to the inputs and outputs of the $L\times L$ Benes, respectively. The edge set E is defined as: for any input-out pair (x_j, y_j) in the permutation, if $x_i \in I_{j1}$ and $y_i \in O_{j2}$, then there is an edge between vertex I_{j1} and vertex O_{j2} in E.

Step 3: Use the same idea as that of Step 4 in Decomposition Algorithm in Section 4.2.2 to split the bipartite graph G into two bipartite graphs $G_1 = (V_1, V_2\,; E_1)$ and $G_2 = (V_1,V_2\,; E_2)$.

Step 4: Take the partial permutation corresponding to the edges in E_1 to form one upper-sub-CFPP, and make the connection of this upper-sub-CFPP through the upper ($L/2$)×($L/2$) sub Benes network; take the partial permutation corresponding to the edges in E_2 to form the lower-sub-CFPP, and make the connection of this lower-sub-CFPP through the lower ($L/2$)×($L/2$) sub Benes network.

Step 5: Recursively call the Routing Algorithm for the upper-sub-CFPP in the upper sub Benes network.

Step 6: Recursively call the Routing Algorithm for the lower-sub-CFPP in the lower sub Benes network. QED

4.4 Crosstalk-free Realization of Permutation in VSB (N, K) Networks

The results in Sections 4.2 and 4.3 indicate that we can realize a permutation of an N-element set {0,1,…, N-1} rearrangeably-nonblocking and crosstalk-free in a VSB(N,K) network by first decomposing the permutation into K CFPPs using the CFPP decomposition algorithm and then realizing each of these CFPPs in one ($2N/K$)× ($2N/K$) Benes of the VSB(N,K) network. In this section, we start the introduction on the crosstalk-free realization of permutation in VSB(N, K) networks by demonstrating an example.

Example 2: Crosstalk-free realization of the following permutation in VSB(16,4) network.

$$\begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 \\ 8 & 2 & 0 & 12 & 4 & 13 & 3 & 11 & 9 & 1 & 6 & 7 & 5 & 10 & 15 & 14 \end{pmatrix} \quad (10)$$

Since K=4 and $\log_2 K$=2, we need two levels decomposition to decompose the permutation into K = 4 CFPPs for the VSB(16,4) network.

First-level decomposition: Take the permutation (10) as the 0-level partial permutation. The bipartite graph describing all edge traversals are shown in Fig. 6.

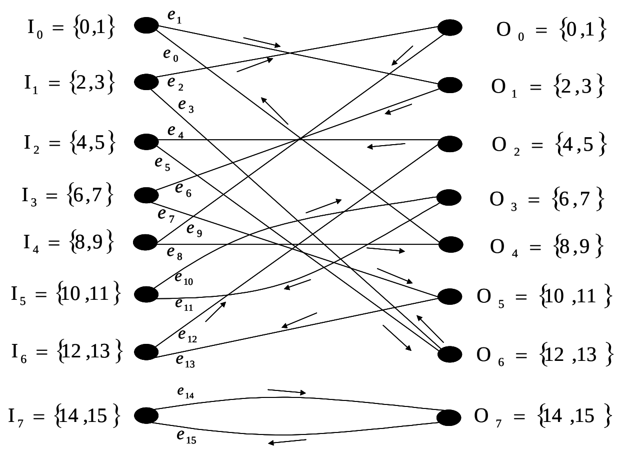

Fig.6. The bipartite graph with all edge traversals of the first-level decomposition.

where $e_0 = \begin{pmatrix} 0 \\ 8 \end{pmatrix}$, $e_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, $e_2 = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$, $e_3 = \begin{pmatrix} 3 \\ 12 \end{pmatrix}$, $e_4 = \begin{pmatrix} 4 \\ 4 \end{pmatrix}$, $e_5 = \begin{pmatrix} 5 \\ 13 \end{pmatrix}$, $e_6 = \begin{pmatrix} 6 \\ 3 \end{pmatrix}$, $e_7 = \begin{pmatrix} 7 \\ 11 \end{pmatrix}$, $e_8 = \begin{pmatrix} 8 \\ 9 \end{pmatrix}$, $e_9 = \begin{pmatrix} 9 \\ 1 \end{pmatrix}$, $e_{10} = \begin{pmatrix} 10 \\ 6 \end{pmatrix}$, $e_{11} = \begin{pmatrix} 11 \\ 7 \end{pmatrix}$, $e_{12} = \begin{pmatrix} 12 \\ 5 \end{pmatrix}$, $e_{13} = \begin{pmatrix} 13 \\ 10 \end{pmatrix}$, $e_{14} = \begin{pmatrix} 14 \\ 15 \end{pmatrix}$, $e_{15} = \begin{pmatrix} 15 \\ 14 \end{pmatrix}$.

Then the input-output pairs e_1, e_2, e_5, e_7, e_8, e_{10}, e_{12} and e_{14} corresponding to the edges in E_1 form a first-level partial permutation:

$$\begin{pmatrix} 1 & 2 & 5 & 7 & 8 & 10 & 12 & 14 \\ 2 & 0 & 13 & 11 & 9 & 6 & 5 & 15 \end{pmatrix} \quad (11)$$

and the input-output pairs e_0, e_3, e_4, e_6, e_9, e_{11}, e_{13} and e_{15} corresponding to the edges in E_2 form the other first-level partial permutation:

$$\begin{pmatrix} 0 & 3 & 4 & 6 & 9 & 11 & 13 & 15 \\ 8 & 12 & 4 & 3 & 1 & 7 & 10 & 14 \end{pmatrix} \qquad (12)$$

which completes the fist-level decomposition.

Second-level decomposition: For the first-level partial permutation (11), the bipartite graph and the corresponding edge traversals are shown in Fig. 7,

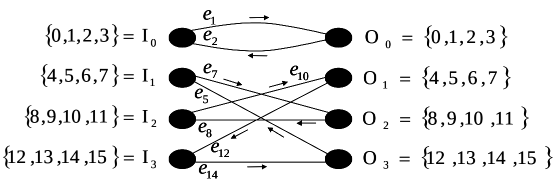

Fig. 7: The first bipartite graph and edge traverses of the second-level decomposition.

Then the input-output pairs e_1, e_7, e_{10} and e_{14} corresponding to the edges in E_1 form the first CFPP:

$$\begin{pmatrix} 1 & 7 & 10 & 14 \\ 2 & 11 & 6 & 15 \end{pmatrix} \qquad (13)$$

and the input-output pairs e_2, e_5, e_8 and e_{12} corresponding to the edges in E_2 form the second CFPP:

$$\begin{pmatrix} 2 & 5 & 8 & 12 \\ 0 & 13 & 9 & 5 \end{pmatrix} \qquad (14)$$

For the partial permutation (12), the bipartite graph the corresponding edge traversals are shown in Fig. 8.

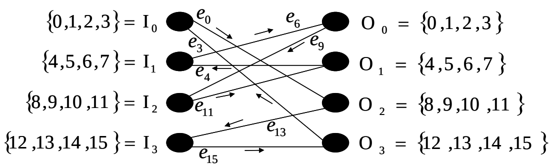

Fig. 8: The second bipartite graph and edge traverses of the second-level decomposition.

The input-output pairs e_0, e_6, e_{11} and e_{15} corresponding to the edges in E_1 form the third CFPP:

$$\begin{pmatrix} 0 & 6 & 11 & 15 \\ 8 & 3 & 7 & 14 \end{pmatrix} \qquad (15)$$

and the input-output pairs e_3, e_4, e_9 and e_{13} corresponding to the edges in E_2 form the fourth CFPP:

$$\begin{pmatrix} 3 & 4 & 9 & 13 \\ 12 & 4 & 1 & 10 \end{pmatrix} \qquad (16)$$

Till now the decomposition is completed and the four CFPPs are the partial permutations (13), (14), (15) and (16).

By realizing each of the four CFPPs (13), (14), (15) and (16) in a single 8×8 Benes of the VSB(16,4) network based on the Routing Algorithm in Section 4.3, the full permutation (10) can be realized crosstalk-free in the network in a single pass. Fig. 9 shows the final crosstalk-free routing of the permutation (10) in the VSB(16,4) network.

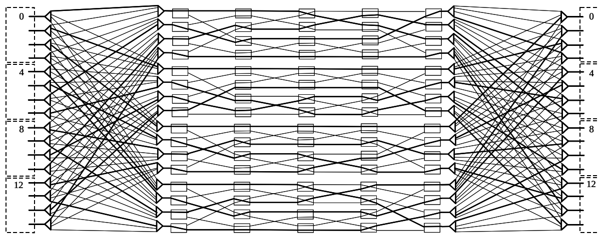

Fig. 9 Crosstalk-free realization of permutation (10) in a VSB (16,4) network.

5. Performance Comparison and Discussion

It is clear that the number of stages in a VSB(N,K) network is $2\log_2(2N/K)$-1, which is determined by the number of stages of its Benes subsets in the network, and the total number of SEs is $[2\log_2(2N/K)-1]\cdot N$. To show some quantitative comparisons, we list in Table I the number of SEs and the cost of splitters/combiners required by RVH(N,m) and VSB(N,K) respectively for a 4096×4096 MIN of various stages.

Table I. Number of SEs and cost of splitters and combiners for networks RVH(4096, m) and VSB(4096,K)

	Number of SEs		Cost of splitters and combiners	
Stages	RVH(N, m)	VSB(N, K)	RVH(N, m)	VSB(N, K)
23	94208	94208	16384	32768
21	172032	86016	32768	65536
19	311296	77824	65536	131072
17	557056	69632	131072	262144
15	983040	61440	262144	524288
13	1703936	53248	524288	1048576
11	-	45056	-	2097152

The cost of a splitter/combiner is calculated by assigning a conventional cost K to each 1×K splitter (K×1 combiner) [13]. If we also assign a conventional cost 4 to each SE, the total hardware cost of a RVH(N,m) network is then given by:

$$C_{RVH} = 2^{\lfloor (\log_2 N - m + 1)/2 \rfloor} \cdot 2N \cdot (\log_2 N + m + 1)$$

and the total hardware cost of a VSB(N,K) network is:

$$C_{VSB} = 4N \cdot \left(K + 2\log_2\left(\tfrac{N}{K}\right) + 1\right)$$

We illustrate in Fig. 10 the overall hardware cost required by RVH(4096,m) network and VSB(4096,K) network of various stages. To have a more complete comparison, we illustrate in Fig. 11 the costs for two networks of smaller sizes (N=32 and N=64).

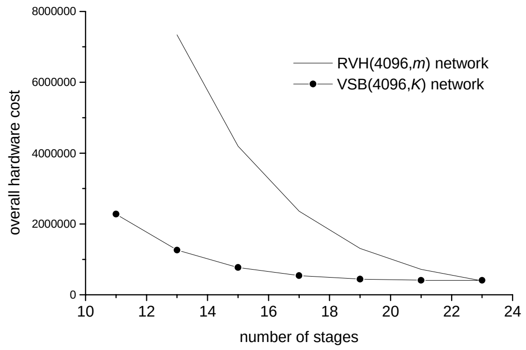

Fig. 10 Overall hardware cost of RVH(4096, m) network and VSB(4096,K) network of various stages.

The results in Table I show that to reduce the number of stages of a RVH(N,m) network, both the number of SEs and the cost of splitters/combiners in the network will be increased accordingly. Although the cost of splitters/combiners of a VSB(N,K) network will also grow with the decrease of the number of stages, the number of SEs will decrease. The results in Fig. 10 and Fig. 11 further indicate that the number of stages for both RVH(N,m) and VSB(N,K) can be reduced by increasing the overall hardware cost of the network. However, with the decrease in the number of stages, the overall hardware cost of RVH(N, m) becomes considerably larger than that of VSB(N,K). The example held in the section shows that to construct a 4096×4096 optical MIN of 13 stages, the hardware cost required by a RVH(4096,m) structure is 7340032, which is almost 6 times of that of the hardware cost required by a VSB(4096,K) structure. Furthermore, the maximum possible number of stages of both RVH(N,m) and VSB (N,K) is 2 $\log_2 N$ -1, but the minimum number of stages that RVH(N, m) and VSB(N,K) can provide is $\log_2 N$ and 1, respectively. This clearly shows that VSB(N,K) provides more alternatives for designing networks of different number of stages.

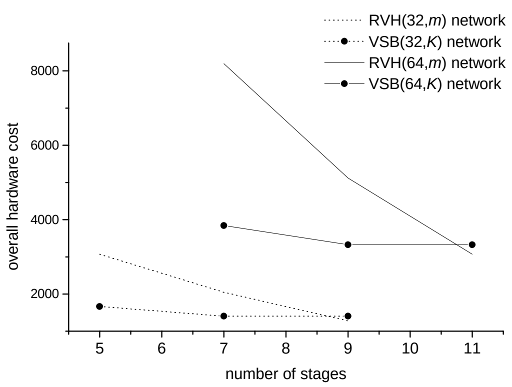

Fig. 11 Overall hardware cost of networks of sizes N=16 and N=32.

Recall that the main task in both the CFPP decomposition for a VSB network and the realization of a CFPP in a plane of the network is to construct and traverse through the bipartite graphs, respectively. Let α be the unit time for constructing and traversing an edge of a bipartite graph, then the time complexity for CFPP decomposition for a VSB(N,K) network is approximately $\alpha \cdot N \cdot \log_2 K$, and the time complexity of realizing CFPPs in the network is $\alpha \cdot N \cdot [\log_2(2\,N/K)-1]$. Thus, the overall time complexity for crosstalk-free realization of a permutation in VSB(N,K) network is approximately $\alpha \cdot N \cdot \log_2 K + \alpha \cdot N \cdot [\log_2 (2\,N/K)-1] \cong \alpha \cdot N \cdot \log_2 N$, which is in the same order of time complexity for the crosstalk-free realization of a permutation in a Benes network [21] or a RVH network [7].

Due to the fact that a VSB(N,K) network consists of K identical Benes networks and the fact that each CFPP can be realized in any of these Benes networks, a trade-off can be initiated between the number of planes and number of passes in supporting the TDM transmission mode. For example, to construct a VSB(4096,K) MIN having 17 stages, we need 16 planes to ensure any permutation to be realized crosstalk-free in one pass, but this number of planes can be reduced to 8 if we allow two passes to realize the permutation (each pass will accommodate half of CFPPs) and 4 if we allow four passes to realize the permutation (each pass will accommodate one fourth of CFPPs). Note that to realize a permutation rearrangeably nonblocking and crosstalk-free in a VSB(N,K) network, we need a central scheduler to first decompose the permutation into K CFPPs using the CFPP decomposition algorithm and then realize each of these CFPPs in one $(2N/K)\times(2N/K)$ Benes of the VSB(N,K) network based on the Routing algorithm in Section 4.3. Thus, the VSB networks proposed in this paper can be used as circuit-switched optical cross-connects (OXCs) for optical transport networks or circuit-switched switching systems for LANS.

This paper focus mainly on the crosstalk-free VSB networks, because an optical channel usually carries an extremely high volume of traffic and crosstalk in optical transmission facilities can significantly increase the bit-error rate of optical signals. Thus, the crosstalk-free VSB networks proposed here are of more interest for the error-sensitive applications (e.g., communication). Since the data communication is also loss-sensitive and the loss of an optical signal is proportional to the number of couplers that the optical signal passes through, so we can adopt a VSB network with a smaller number of stages for the loss-sensitive application at the expense of taking more planes (and thus a higher overall hardware cost as illustrated in Fig. 10 and Fig. 11). Note that the crosstalk-free constraint used in this paper may not be necessary for some applications that are not very sensitive to crosstalk (e.g., sensing or image processing), and it is foreseeable that the hardware cost of networks can be reduced further for these applications if we allow certain degrees of crosstalk in transmission. In particular, for the applications without the crosstalk constraint, all optical signals only need to follow link-disjoint paths rather than SE-disjoint paths in transmission and the vertically stacked Benes architecture proposed in this paper can be modified slightly to provide the full permutation capability for these applications with only half number of planes required by its crosstalk-free counterpart. Thus, the VSB network structure is flexible in the sense that it enables a graceful trade-off to be made among number of stages, number of passes and number of planes depending on the requirements of different applications in terms of error rate, signal loss, transmission delay and hardware cost.

6. Conclusions

In this paper, we proposed a new class of optical MINs, namely vertically stacked Benes (VSB) networks. One of the most significant contribution of the proposed VSB architecture is that it has a considerably lower overall hardware cost than any reported optical MINs counterparts based on a combination of vertical stacking and horizontal expansion (VH) of banyan networks, while providing the same capability of realizing crosstalk-free permutation rearrangeably-nonblocking in a single pass. In addition, the VSB network structure yields much higher flexibility in terms of the number of stages (and also the number of planes), which can well compromise among the number of stages, the number of passes, and the number of planes concerning the requirements of signal loss, transmission delay and hardware cost. We have presented an efficient scheme for realizing a permutation crosstalk-free in the proposed VSB network structure by first decomposing it into multiple crosstalk-free realizable partial permutations and then realizing each of them in a stacked Benes of the VSB network. The scheme has the same time complexity as that for realizing a permutation crosstalk-free in a Benes network or a RVH network.

Acknowledgement

The authors would like to thank the editor and anonymous reviewers for their valuable and constructive comments. This work is support in part by Telecommunications Advancement Organization of Japan (TAO), Japan Society for the Promotion of Science (JSPS) Grant-in-Aid for Scientific Research (B) under Grant No.16700056 and Grant No. 17300010.

References

[1] I.Anderson, *Combinatorics of Finite Sets*, Oxford Science Publications, 1987.

[2] V.Benes, *Mathematical Theory of Connecting Networks and Telephone Traffic*. New York: Academic,1965.

[3] V.R.Chinni etal., "Crosstalk in a lossy directional coupler switch," *J.Lightwave Technol.*, vol.13, no.7, pp.1530-1535,July 1995.

[4] R.Cole and J.Hopcroft, "On edge coloring bipartite graphs", SIAM J.Comput.,vol.11, pp.540-546,1982.

[5] G.R.Goke and G.J.Lipovski, "Banyan networks for partitioning multiprocessor systems," *Proc.1st Annu. Symp. Comp. Arch*,, pp.21-28,1973.

[6] H.S.Hinton, *An introduction to Photonic Switching Fabrics*, New York: Plenum, 1993.

[7] Xiaohong Jiang, Hong Shen, Md. M. Khandker and Susumu Horiguchi, "Crosstalk-free Permutation in Photonic Rearrangeable Networks Built on a Combination of Horizontal Expansion and Vertical Stacking of Banyan Networks," *IEICE Transactions on Information and Systems*, vol.E86-D/9,pp.1525-1533, 2003.

[8] C.Kruskal and M.Snir, "The performance of multistage interconnection networks for multiprocessors," *IEEE Trans. Commun*., vol.COM-32, pp.1091-1098, Dec.1983.

[9] C.Lea and D.Shyy, "Tradeoff of horizontal decomposition versus vertical stacking in rearrangeably nonblocking networks," *IEEE Trans. Commun*., vol.39, pp.899-904, June 1991.

[10] C.-T. Lea, "Muti-log$_2$N networks and their applications in high speed electronic and photonic switching systems," *IEEE Trans. Commun*., vol.38, pp.1740-1749, Oct. 1990.

[11] F. Thomson Leighton, *Introduction to Parallel Algorithms and Architectures: Arrays, Trees, Hypercubes*, Morgan Kaufmann, 1992.

[12] D.Li, "Elimination of crosstalk in directional coupler switches," *Optical Quantum Electron.*, vol.25, no.4, pp.255-260, Apr.1993.

[13] G.Maier and A.Pattavina, "Design of photonic rearrangeable networks with zero first-order switching-element-crosstalk, " *IEEE Trans. Commun*., vol.49,no.7, pp.1268-1279, July.2001.

[14] K.Padmanabhan and A.Netravali, "Dilated networks for photonic switching," *IEEE Trans. Commun*., vol.COM-35, pp.1357-1365, Dec.1987.

[15] J.Patel, "Performance of processor-memory interconnections for multiprocessors," *IEEE Trans. Comput*., vol.C-30, pp.771-780, Oct.1981.

[16] A.Pattavina, *Switching Theory - Architecture and Performance in Broadband ATM Networks*, 1st ed. New York: Wiley,1998.

[17] M.M.Vaez and C.-T. Lea, "Strictly nonblocking directional-coupler-based switching networks under crosstalk constraint," *IEEE Trans. Commun*., vol.48,no.2, pp.316-323, Feb. 2000.

[18] M.M.Vaez and C.-T. Lea, "Space-wavelength tradeoff in the design of nonblocking directional coupler based network under crosstalk constraint," *J. Lightwave Technol*.,vol.16,pp.1373-1379,Aug.1998.

[19] M.M.Vaez and C.-T. Lea, "Wide-sense nonblocking Banyan-type switching systems based on directional couplers, " *IEEE J. Select. Areas Commun*., vol.16, pp.1327-1332, Sept.1998.

[20] T.-S. Wong and C-T. Lea, "Crosstalk reduction through wavelength assignment in WDM photonic switching networks, " *IEEE Trans. Commun*., vol.49, no.7,pp.1280-1287, July.2001.

[21] Y.Yang, J.Wang and Y.Pan, "Permutation Capability of Optical Multistage Interconnection Networks," *Journal of Parallel and Distributed Computing*, vol.60, no.1, pp.72-91, Jan.2000.

[22] R.Ramaswami and K.N.Sivarajan, *Optical Networks: A Practical Perspective*. San Mateo, CA: Morgan Kaufmann,2002.

[23] Fujitsu Develops High-Speed Multichannel Optical Switch Based on MEMS Mirrors, 2003. http://pr.fujitsu.com/en/news/2003/09/29.html

Xiaohong Jiang (M'03) received his B.S., M.S. and Ph.D degrees in 1989, 1992, and 1999 respectively, all from Xidian University, Xi'an, China. He is currently an associate professor in the Department of Computer Science, Graduate School of Information Science, TOHOKU University, Japan. Before joining TOHOKU University, Dr.Jiang was an assistant professor in the Graduate School of Information Science, Japan Advanced Institute of Science and Technology (JAIST), from Oct.2001 to Jan.2005. Dr. Jiang was a JSPS (Japan Society for the Promotion of Science) postdoctoral research fellow at JAIST from Oct.1999-Oct. 2001. He was a research associate in the Department of Electronics and Electrical Engineering, the University of Edinburgh from Mar.1999-Oct.1999. Dr. Jiang's research interests include optical switching networks, WDM networks, interconnection networks, IC yield modeling, timing analysis of digital circuits, clock distribution and fault-tolerant technologies for VLSI/WSI. He has published over 70 referred technical papers in these areas. He is a member of IEEE.

Pin-Han Ho received his B.Sc. and M.Sc. Degree from the Electrical and Computer Engineering department in the National Taiwan University in 1993 and 1995. He started his Ph.D. study in the year 2000 at Queen's University, Kingston, Canada, focusing on optical communications systems, survivable networking, and QoS routing problems. He finished his Ph.D. in 2002, and joined the Electrical and Computer Engineering department in the University of Waterloo, Waterloo, Canada, as an assistant professor at the same year. Professor Pin-Han Ho is the first author of more than 50 refereed technical papers and the co-author of a book on optical networking and survivability. He is the recipient of the Best Paper Award and the Outstanding Paper Award from SPECTS'02 and HPSR'02, respectively.

Susumu Horiguchi received the B.Eng the M.Eng and PhD degrees from Tohoku University in 1976, 1978 and 1981 respectively. He is currently a professor and Chair of Department of Computer Science, the Graduate School of Information Science, Chair of Department of Information Engineering, Faculty of Engineering, Tohoku University. He was a visiting scientist at the IBM Thomas J. Watson Research Center from 1986 to 1987. He was also a professor in the Graduate School of Information Science, JAIST (Japan Advanced Institute of Science and Technology). He has been involved in organizing international workshops, symposia and conferences sponsored by the IEEE, IEICE, IASTED and IPS. He has published over 150 papers technical papers on optical networks, interconnection networks, parallel algorithms, high performance computer architectures and VLSI/WSI architectures. Prof. Horiguchi is a senior member of IEEE and member of IPS and IASTED.