BEHAVIOR ANALYSIS OF OPTICAL BANYAN NETWORK ON VERTICAL STACKING

Thesis submitted in partial fulfillment of the requirements for the award of degree of

Master of Engineering
in
Computer Science & Engineering

By:
K. Nagarjun
(80732009)

Under the supervision of:
Rinkle Aggarwal
Lecturer (SS), CSED

COMPUTER SCIENCE AND ENGINEERING DEPARTMENT
THAPAR UNIVERSITY
PATIALA – 147004

JUNE - 2009

Certificate

I hereby certify that the work which is being presented in the thesis entitled, **“Behavior analysis of Optical Banyan Network on Vertical Stacking ”**, in partial fulfillment of the requirements for the award of degree of Master of Engineering in Computer Science and Engineering submitted in Computer Science and Engineering Department of Thapar University, Patiala, is an authentic record of my own work carried out under the supervision of Rinkle Aggarwal and refers other researcher’s works which are duly listed in the reference section.

The matter presented in this thesis has not been submitted for the award of any other degree of this or any other university.

(**K. NAGARJUN)**

This is to certify that the above statement made by the candidate is correct and true to the best of my knowledge.

(**RINKLE AGGARWAL)**
Lecturer (SS),
Computer Science and Engineering Department,
Thapar University, Patiala.

Countersigned by

(SEEMA BAWA)
Professor & Head,
Computer Science & Engineering Department,
Thapar University,
Patiala.

(R.K.SHARMA)
Dean (Academic Affairs),
Thapar University,
Patiala.

Acknowledgement

It is a great pleasure for me to acknowledge the guidance, assistance and help I have received from Rinkle Aggarwal, Lecturer (SS), Computer Science and Engineering Department. I am thankful for her continual support, encouragement, and invaluable suggestions. She not only provided me help whenever needed, but also the resources required to complete this thesis report on time.

I am also thankful to Dr. Seema Bawa, Head, Computer Science and Engineering Department for her kind help and cooperation.

I would also like to thank all the staff members of Computer Science and Engineering Department for providing me all the facilities required for the completion of my thesis work.

I would like to say thanks for support of my Classmates. I want to express my appreciation to every person who contributed with either inspirational or actual work to this thesis.

I am deeply grateful to my parents and brothers for the inspiration and ever encouraging moral support, which enabled me to pursue my studies.

Place: TU, Patiala

Date: **(K. NAGARJUN)**

Abstract

Multistage Interconnection Networks (MINs) are very popular in switching and communication applications and have been used in telecommunication and parallel computing systems for many years. Extensive research has been done on Electronic Multistage Interconnection Networks (EMINs) but these days advances in electro-optic technologies have made optical communication a promising networking choice to meet the increasing demands of high-performance computing communication applications for high channel bandwidth and low communication latency. The electronic Multistage Interconnection Networks (EMINs) and the Optical Multistage Interconnection Networks (OMINs) have many similarities, but there are some fundamental differences between them. The new challenges facing optical MINs such as the optical-loss during switching and the crosstalk (it is caused by coupling two signals within a switching element). Banyan networks are attractive switching networks because they are fast in switch setting (self-routing) and also have small number of switches between an input-output pair.

Vertical stacking of optical banyan networks is a novel scheme for building nonblocking and crosstalk-free optical switching network, but significantly increases the hardware cost. This thesis, analyses the blocking probability in the optical networks and the deterministic conditions for strictly nonblocking and rearrangeably nonblocking Vertically Stacked Optical Banyan (VSOB) Networks with and without link failures. It also includes the new approach called pruning of banyan structure to save the hardware cost. In this approach two kinds of banyan planes are used: pruned and regular banyan. A plane is pruned by eliminating some of its switching elements as those are never used in the system.

Table of Contents

List of Figures

List of Tables

CHAPTER 1

Introduction to MINs

1.1 Multistage Interconnection Networks (MINs)

Electronic multistage interconnection networks (MINs) have been studied extensively as an important interconnecting scheme for communication and parallel computing systems [1]. Multistage interconnection networks (MINs) consist of more than one stages of small interconnection elements called switching elements and links interconnecting them [1,2]. The number of stages and the connection patterns between stages determine the routing capability of the networks. MINs were initially proposed for telephone networks and later they are used in multiprocessing systems to provide cost-effective, high-bandwidth communication between processors and/or memory modules. In these cases, a central controller establishes the path from input to output. In cases where the number of inputs equals the number of outputs, each input synchronously transmits a message to one output, and each output receives a message from exactly one input. Such unicast communication patterns can be represented as a permutation of the input addresses [3]. On the other hand, in asynchronous multiprocessors, centralized control and permutation routing are infeasible. In this case, a routing algorithm is required to establish the path across the stages of a MIN [2]. Depending on the interconnection scheme employed between two adjacent stages and the number of stages, various MINs have been proposed [2,4]. Multistage interconnection networks are:

- Attempt to reduce cost
- Attempt to decrease the path length

1.2 A Generalized MIN Model

There are many ways to interconnect adjacent stages [2]. Figure 1.1 shows a generalized multistage interconnection network with M inputs and N outputs. It has n stages S_0 to S_{n-1}

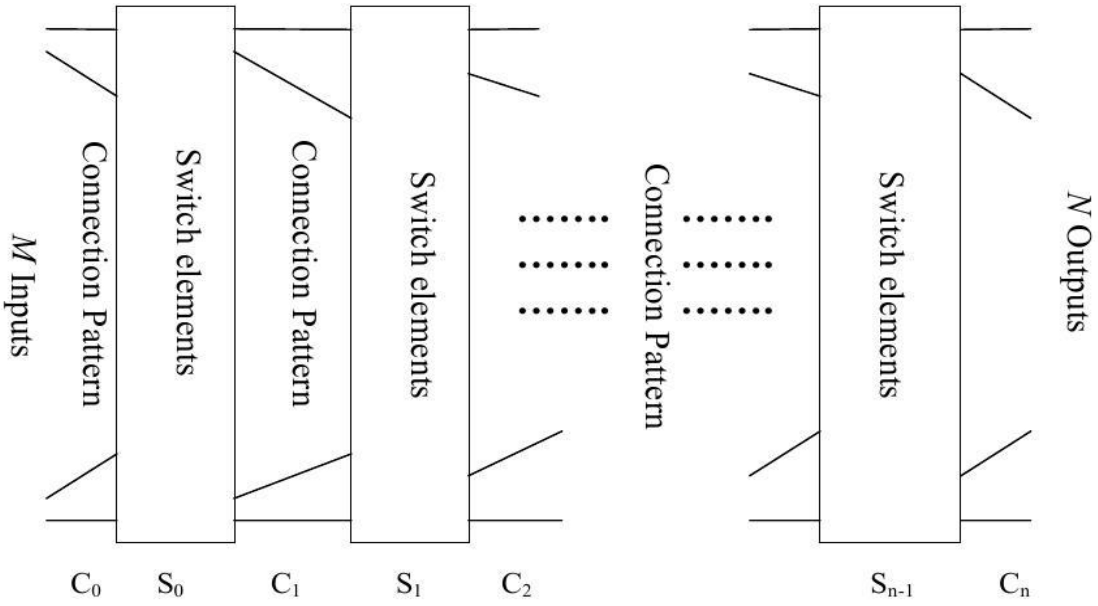


Figure 1.1: A generalized MIN with M inputs, N outputs, and n stages [2]

As shown in Figure 1.2, each stage, say S_i has w_i switches of size $a_{i,j} \times b_{i,j}$, where $1 \leq j \leq w_i$. Thus, stage S_i has p_i inputs and q_i outputs, where

$$p_i = \sum_{j=1}^{w_i} a_{i,j} \text{ and } q_i = \sum_{j=1}^{w_i} b_{i,j}$$

The connection between two adjacent stages, S_{i-1} and S_i, denoted C_i, defines the connection pattern for $p_i = q_{i-1}$ links, where $p_0 = N$ and $q_{g-1} = M$. A MIN thus can be represented as

$C_0(M)S_0(w_0)C_1(p1)S1 \ldots Sn - 1(w_{n-1})C_n(N)$

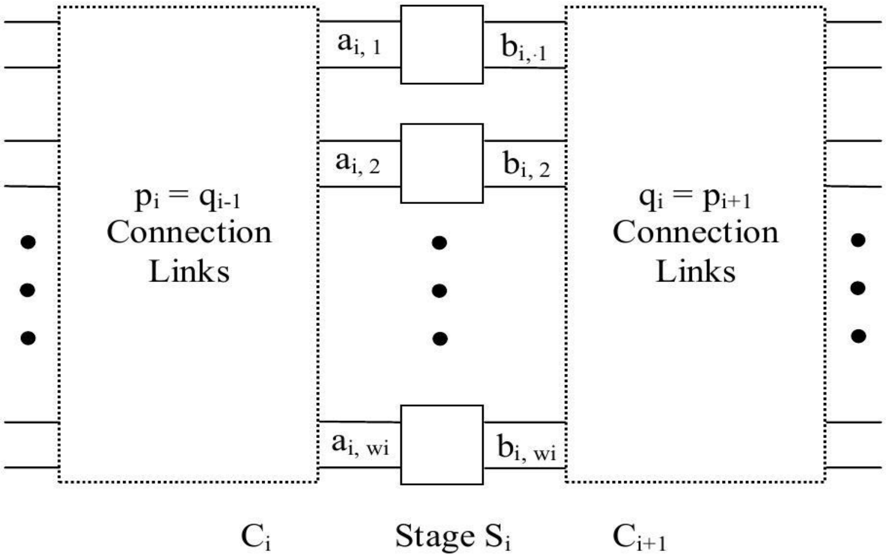


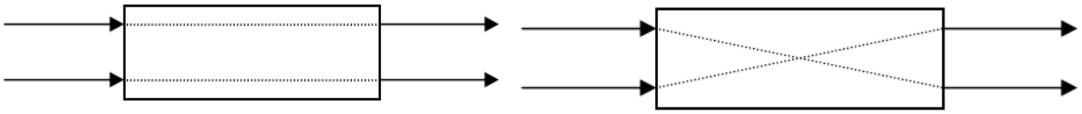
Figure 1.2: A closer view of one of the stages [2]

A connection pattern $C_i(p_i)$ defines those p_i links should be connected between the $q_{i-1} = p_i$ outputs from stage G_{i-1} and the p_i inputs to stage G_i. Different connection patterns give different characteristics and topological properties of MINs. The links are labeled from 0 to p_{i-1} at C_i. From a practical point of view, it is interesting that all the switches are identical, thus amortizing the design cost. Banyan networks are a class of MINs with the property that there is a unique path between any pair of source and destination [4]. An N-node ($N= k^n$) Delta network is a sub class of banyan networks, which is constructed from identical k × k switches in n stages, where each stage contains N/k switches. Many of the known MINs, such as Omega, flip, cube, butterfly, and baseline [2], belong to the class of Delta networks [5] and have been shown to be topologically and functionally equivalent. A good survey of those MINs can be found in [1,2].

1.3 Switching Elements

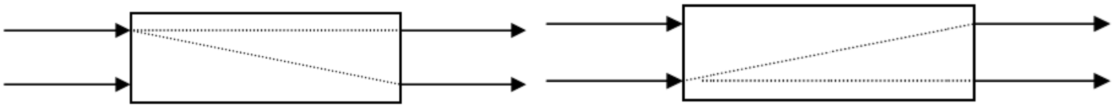
The switching element is the basic element of a multistage interconnection network. It may be viewed as a very small network. These switches are the devices having multiple inputs and multiple outputs [6]. A switch box have any one of the following four states i.e. straight, exchange, upper broadcast and lower broadcast shown in Figure 1.3.

Switching Element



Straight

Exchange



Lower Broadcast

Upper Broadcast

Figure 1.3 Different states of a switching element

1.4 Classification of Multistage Interconnection Networks

Multistage interconnection networks can be classified according to different categories. The main classification categories are:

- According to number of paths
- According to switch
- According to control
- According to availability of path

1.4.1 Classification According to Number of Paths

Unique path networks: These networks provide unique path between every source and destination. The failure of any switching element along the path disconnects some source destination pairs. These are not reliable for a large multiprocessor system, as they cannot tolerate even a single fault. These networks provide poor performance as a source destination connection may be blocked by a previously established connection [7].

Multipath networks: These provide more than one path between source and destination. In case, there is a failure of one switching element in the path, the request is routed through some alternative path. Multi path multistage interconnection networks can be static or dynamic. For static networks, backtracking is used if fault is encountered. In dynamic networks, if a fault is encountered in a particular stage, a switching element in preceding stage will re-route data through an alternative available path [7].

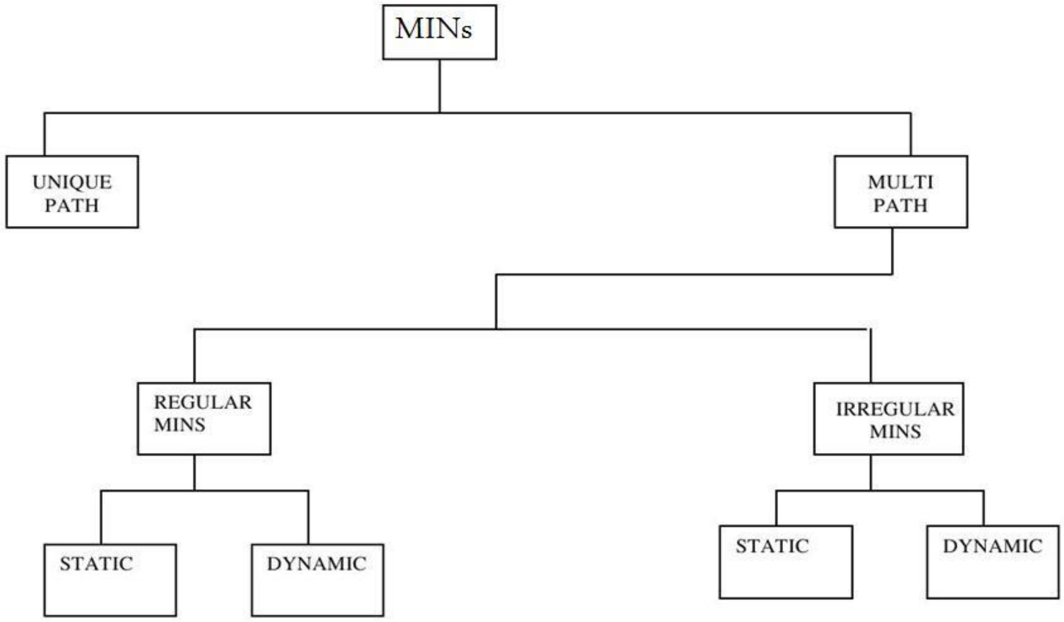


Figure 1.4: Classification of MINs

1.4.2 Classification According to Switches

Regular networks: Regular multistage interconnection networks have an equal number of switching elements per stage. As a result they may impose equal time delay to all requests passing through them [8].

Irregular networks: Irregular multistage interconnection networks have unequal number of switching elements at each stage and thus they are inherently multi path in nature. For a given source destination pair, multiple paths are available [8].

1.4.3 Classification According to Control

Flip controlled networks: Flip controlled multistage interconnection networks have a common control signal for switching in various switching elements at a given stage [9].

These networks are less complicated due to lesser number of control signals but have lesser bandwidth.

Distributed control networks: Distributed control multistage interconnection networks have a separate control signal for every switching element. These have higher bandwidth due to selection of source destination pair at a given time and are quite complex [9].

1.4.4 Classification According to Availability of Path

Blocking networks: Blocking is the conflict that arises between two paths established between two source and destination pairs. When sources generate connection requests to an identical destination, it is possible that some paths need to pass simultaneously through an identical output of the SE. A path can never share the output with another path at any point in time, and hence blocking arises in the SE. Omega network is a type of blocking network [2].

Non blocking networks: A network is called nonblocking if it is possible to route from any source to any destination, in presence of other established source-destination routes, provided no two sources have same destination. In other words, networks that can handle all possible connections without blocking is called nonblocking network [2].

Table 1.1 Comparison of bus, crossbar and Multi-stage interconnection networks

Property	Bus	Crossbar	Multistage
Speed	Low	High	High
Cost	Low	High	Moderate
Reliability	Low	High	High
Configur-ability	High	Low	Moderate
Complexity	Low	High	Moderate

CHAPTER 2

Literature Survey

From the research it is studied that Optical communication is better than the electronic communication in terms of bandwidth and latency. There are various similarities between the two networks but optical multistage interconnection networks have their own challenges. These problems can be removed by using the various techniques. For all these various literatures which is studied from them are as follows

1) "Optical Multistage Interconnection Networks: New Challenges and Approaches"

This paper presents an idea about Optical interconnections for communication networks and multiprocessor systems. The basic element of optical switching networks is a directional coupler with two inputs and two outputs or switching elements. Depending on the control voltage applied to it, an input optical signal is coupled to either of the two outputs, setting the SE to either the straight or cross state. A class of topologies that can be used to construct optical networks is multistage interconnection networks, which interconnect their inputs and outputs via several stages of SEs (Switching elements), are also mentioned in this paper. In this various ways to deal with the unique problem of avoiding crosstalk in the SEs (Switching elements) are given.

2) "A New Scheme to Realize Crosstalk-free Permutations in Optical MINs with Vertical Stacking"

In this paper a new concept vertical stacking, a new approach for constructing nonblocking multistage interconnection networks (MINs) is presented. This paper gives an idea of the crosstalk-free permutation in rearrangeable, self-routing Banyan-type optical MINs built on vertical stacking and propose a new scheme for realizing permutations in this class of optical MINs crosstalk-free.

3) "Nonblocking Optical MINs under Crosstalk-free Constraint"

In this nonblocking condition for crosstalk-free optical MINs under packing strategy is given. It is shown that packing strategy is very efficient in reducing the hardware cost of a nonblocking optical MIN under the constraint of crosstalk-free. The necessary condition for a banyan type network to be nonblocking is the number of planes $p \geq \sqrt{N}$if $\log_2 N$ is even, and planes $p \geq \sqrt{2N}$if $\log_2 N$is odd.

4) "Blocking Behaviors of Crosstalk-free Optical Banyan Network on Vertical Stacking"

The blocking probabilities of Vertically Stacked Optical Banyan (VSOB) networks studied from probabilistic view and upper and lower bounds on VSOB networks presented. These bounds depict accurately the overall blocking behaviors of VSOB networks and agree with the conditions of strictly nonblocking and rearrangeably nonblocking VSOB networks respectively.

5) "Performance of Fast Routing Algorithms in Large Optical Switches Built on the Vertical Stacking of Banyan Structures"

In this paper, two fast routing algorithms proposed for the VSOB network based on the idea of inputs grouping. The two algorithms, namely plane fixed routing (PFR) algorithm and partially random routing (PRR) algorithm, have the time complexities of O($\log N$) and O($\sqrt{N}$) respectively, and FR algorithm can actually turn a VSOB network into a self-routing one. Results are shown that both routing algorithms can achieve comparable and considerably low blocking probabilities for large VSOB networks.

6) "Blocking Probability of Self-Routing Vertically Stacked Optical Banyan Switching Networks with Given Crosstalk Constraint"

The recently proposed Plane Fixed Routing (PFR) algorithm , which is distributed control routing algorithm, turns the VSOB networks into self-routing, and therefore provides optimum delay in the order of $\log_2 N$.The PFR algorithm results comparably low blocking probability(<9%) with zero first-order crosstalk. This paper presents the simulation results of the blocking probability of the self-routing VSOB networks with given crosstalk constraint.

7) “Effects of Link Failures on the Overall Blocking Behavior of Banyan-based Optical Switches”

This paper gives an idea about the study of overall blocking behavior of a VSOB switch by establishing an upper bound on its blocking probability in presence of link failures. The upper bound accurately depicts the overall blocking behavior of a VSOB switch for a reasonable small link failure rate, as verified by extensive simulation results. This upper bound is significant because it provides switch designers a quantitative tool to determine the effects of link failure on the blocking behavior and to estimate the maximum blocking probability of a VSOB switch.

8) “On Number of planes of Rearrangeably Nonblocking Optical Banyan Networks with Link Failures”

This paper presents the approximate number of planes required to make a VSOB networks rearrangeably nonblocking allowing link-failures. It also show an interesting behavior of the blocking probability of a faulty VSOB networks that the blocking probability may not always increase monotonously with the increase of link-failures; blocking probability decreases for certain range of link-failures and then increases again.

9) “Lower Bound on Number of Planes for Vertically Stacked Optical Banyan Networks with Link Failures”

This paper deals with the blocking behavior of VSOB networks when some links are broken or failed. Also gives the approximate value for lower bound on number of planes required to make VSOB networks nonblocking allowing link-failures. An interesting finding is that the blocking probability of VSOB networks not always increase with the increase in link-failures, the fluctuations happen in the blocking behavior.

10) “Upper Bound on Blocking probability for Vertically Stacked Optical Banyan Networks with Link Failures and Given Crosstalk Constraint”

As blocking behavior analysis is an effective approach to studying network performance, this paper presents the simulation results for upper bound on blocking probability considering both link-failures and given degree of crosstalk constraint. These results can guide the network designers in finding a tradeoff among the blocking probability, the degree of crosstalk and link failures of VSOB networks.

11) “Blocking Behavior of Crosstalk-free Pruned Optical Banyan Networks”

This paper presents a new concept called pruning of banyan network for faster connection establishment that can be used in optical cross connect. A plane is pruned by eliminating some of its switching elements those are never used in the system. And also it gave an idea about the blocking behavior of generalized architecture of pruned optical banyan networks. This paper presented the new structure called extended pruned vertically stacked optical banyan (EP-VSOB) network. In this structure two kinds of banyan planes have been used: pruned and regular banyan.

12) “Blocking Behavior Analysis of Extended Pruned Vertically Stacked Optical Banyan networks with Link Failures”

This paper presents two routing algorithms. These routing algorithms show that this switching network can reduce the blocking probability to very low value even with zero-crosstalk constraint while keeping the hardware cost almost the same as that of EP-VSOB networks, and also studied the blocking behavior of EP-VSOB networks when some links are failed or broken in the network.

CHAPTER 3

Problem Statement

It is expected that users of telecommunication services such as Internet, Web browsing, and tele-education will increase dramatically. This has greatly increased the demand for high bandwidth and high capacity communication systems. Advances in electro-optic technologies have made optical communication a good networking choice for the increasing demands of high channel bandwidth and low communication latency of high-performance computing/communication applications. Fiber optic communications offer a combination of high bandwidth, low error probability, and gigabit transmission capacity. Multistage Interconnection Networks (MINs) are very popular in switching and communication applications and have been used in telecommunications and parallel computing systems. In electronic MINs electricity is used, where as in Optical MINs light is used to transmit the messages. The electronic MINs and the optical MINs have many similarities, but there are some fundamental differences between them such as the optical-loss during switching and the crosstalk problem in the optical switches.

Available optical MINs were built mainly on banyan or its equivalent (e.g. *baseline*, *omega*) networks because they are fast in switch setting (self-routing) and also have a small number of switches between an input-output pair. Banyan networks have a unique path between an input-output pair, and this makes them blocking networks. Nonblocking networks can be constructed by either appending some extra stages to the back of a regular banyan network.

Problems in Optical Networks are:

1. Path Dependent Loss
2. Optical crosstalk

Two ways are used to solve crosstalk problem in Optical Networks:

1. Space Domain Approach
2. Time Domain Approach

One of the most serious problems is optical crosstalk in optical MIN. This crosstalk occurs when two signal channels interact with each other. When a crosstalk happens, a small fraction of the input signal power may be detected at another output although the main signal is injected at the right output. For this reason, when a signal passes many switching elements, the input signal will be distorted at the output due to the loss and crosstalk introduced on the path. This was not a big issue in electrical MINs, but because the more stringent bit error rate in optical network, it has become a big problem.

Vertical stacking of multiple copies of an optical banyan network is a novel scheme for building nonblocking (crosstalk-free) optical switching networks. The resulting network namely Vertically Stacked optical Banyan (VSOB) network preserves all the properties of banyan network, but increases the hardware cost significantly.

CHAPTER 4

Optical Multistage Interconnection Networks

Optical switches are used in optical networks for a variety of applications. The different applications require different switching times and number of switch ports. In terms of the switching function achievable, switches are of two types: blocking or nonblocking. A switch is said to be nonblocking if an unused input port can be connected to any unused output port. If some interconnection cannot be realized, the switch is said to be blocking, to be nonblocking otherwise. There are three types of nonblocking networks, namely strictly nonblocking, wide sense nonblocking, and rearrangeably nonblocking. This will focus on one of the rearrangeably nonblocking networks, Banyan network.

4.1 Optical Multistage Interconnection Networks

Optical Multistage Interconnection Networks (OMINs) [10] differ from Electrical Multistage Interconnection Networks; in which optical signal is converted to/from electrical signal at the network input/output, optical Multistage Interconnection Networks work in optical domain. This advantage makes the signal transmission in optical network faster. As networks face increasing bandwidth demand and diminishing fiber availability [11], network providers are moving towards a crucial milestone in network evolution: the optical network. Optical networks, based on the emergence of the optical layer in transport networks, provide higher capacity and reduced costs for new applications such as the Internet, video and multimedia interaction, and advanced digital services.

Optical networks are high-capacity telecommunications networks based on optical technologies and components that provide routing, grooming, and restoration at the wavelength level as well as wavelength-based services. Customers are demanding more services and options and are carrying more and different types of data traffic. Optical networks provide the required bandwidth and flexibility to enable end-to-end wavelength services and meet all the high-capacity and varied needs [2,11].

Optical fiber offers much higher bandwidth than conventional copper cables. A single fiber has a potential bandwidth on the order of 50THz [12]. Meanwhile, it has low cost, extremely low bit error rate (typically 10^{-12}, compared to 10^{-6} in copper cables), low signal attenuation and low signal distortion. In addition, optical fibers are more secure from tapping, since light does not radiate from the fiber and it is nearly impossible to tap into it secretly without being detected. As a result, it is the preferred medium for data transmission with bit rate more than a few tens of megabits per second over any distance more than one kilometer. It is also the preferred means of realizing short distance (a few meters to hundreds of meters), high-speed (gigabits per second & above) interconnection inside large systems. In the past few decades, optical fibers have been widely deployed in all kinds of telecommunications networks. Optical fiber has been used in two generations of optical network. In the first generation, it was essentially used for transmission and simply to provide capacity, since it provides lower bit error rates and higher capacities than copper cables. All the switching and other intelligent network functions were handled by electronics. Thus, the bandwidth was limited by the electronics at the fiber endpoints. Currently, transmission rates are restricted to 10 Gb/s (OC-192) in commercially available systems. Examples of the first generation optical networks are SONET and SDH networks.

In the second-generation optical networks, optical layer handles switching and some of the routing. The fiber bandwidth is further exploited by a technique called wavelength division multiplexing (WDM), where the optical bandwidth is partitioned into a large number of channels on different wavelengths (or, equivalently, colors), and each channel works at peak electronic rate. These wavelengths do not interfere with each other as long as the channel space is large enough. Other than providing a huge bandwidth, WDM networks can also provide data transparency in which the network may accept data at any bit rate and any protocol format within the limits. Data transparency may be realized through all-optical (or single-hop) transmission and switching of signals. In an all-optical network, data is transferred from source to destination in optical form, without undergoing any optical-to-electrical conversion. Keeping the signal in optical form eliminates the "electronic bottleneck" of communications networks with electronic switching [10].

4.2 Switching in Optical Networks

In optical networks, circuit switching is used. Packet switching is not possible with Optical Multistage Interconnection Networks. If packet switching is used, the address information in each packet must be decoded in order to determine the switch state. In a hybrid MIN, it means it require conversions from optical signals to electronic ones, which could be very costly. For this reason, circuit switching is usually preferred in optical MINs. So assume that circuit switching is used.

Table 4.1 Characteristics of optical and electronic multistage networks

Characteristics	Electronic Multistage Networks	Optical Multistage Networks
Speed	Less	High
Energy Transmitted	Electricity	Light
Bandwidth	Used for less bandwidth applications	Used for high bandwidth applications
Latency	High	Less
Error Probability	High	Less
Weight	More	Less
Cost	Less	More
Switching	Packet Switching	Circuit Switching
Path	Provide Multi path from source to destination.	Provide single path from source to destination
Complexity	More Complex	Less Complex
Structure considered	2-dimensional	3-dimensional

There are lots of benefits of optical networks over the electronic ones. The main benefit of the optical networks over the electronic network is the high speed of the Optical signals. In the Optical networks light is transmitted which has a very good speed but in the electronic Multistage interconnection networks electricity is used which has very slow speed.

The second advantage is the bandwidth. These days applications in communication require high bandwidth. The optical networks give combination of very high bandwidth and low latency. That is why they have been used in the parallel processing applications. Optical MINs are also used in wide area networks which require less error probability and very high bandwidth. Fiber optic transmission distance is significantly greater than the electronic ones. The signal needs not to be regenerated in optical networks. Optical fiber has very less weight in comparison to electronic MINs. Thus Optical networks give the combination of high bandwidth and low latency.

4.3 Various Optical MINs

Crossbar network is a squared (or rectangular), wide-sense no blocking network without Crossover. It has first order switch crosstalk and requires N^2 switches.

The Partitioned Optical Passive Star (POPS) network is a SIMD interconnection network that uses multiple optical passive star couplers. OPS coupler is an all-optical passive device which is capable of receiving an optical signal from one of its d sources and broad cast it to all of its d destinations. Being a passive all-optical technology, it benefits from a number of characteristics such as no opto-electronic conversion, high noise immunity, and low latency.

Omega networks implemented by free-space digital optics have been widely used in optical computing and photonic switching systems because of its flexibility, simplicity, and identity of link functions.

Double crossbar network consists of two crossbar-like structures placed on top of each other.it is a strictly no blocking network without first order switch crosstalk. But it has crossover and uses $2N^2$ switches.

Banyan network has the best results on the number of switch elements ((NlogN)/2) and the path length (logN). But it is a blocking network with first order switch crosstalk.

Cantor network is one of the symmetrical, strictly nonblocking networks. Unlike the unique-path and rearrangeable networks, strictly nonblocking networks have the ability to provide paths among the inputs and the outputs without disturbing the ongoing communications. For N inputs, Cantor network can be constructed using $4N \log N$ cross-points and $2\log_2 N + 1$ depth (stages).

Benes network can be considered as a Banyan network concatenated with a reverse Banyan network and the middle stage overlapped. It is a rearrangeable nonblocking network with first order switch crosstalk and complicated control algorithm.

Dilated Benes network vertically stacks benes network and adds one extra at the end of it DBN solves the first order switch crosstalk problem by following the restriction that at most one input at every switch element can be active at a time. It is a rearrangeable nonblocking network built using recursive construction.

Extended baseline network is constructed by adding a cross-connect stage after the recursive baseline network. This cross-connect stage provides one more path.

4.4 MIN Switches

There are two types of switches: blocking or nonblocking. A switch is said to be nonblocking if an unused input port can be connected to any unused output port. Thus a nonblocking switch is capable of realizing every interconnection pattern between the inputs and the outputs. If some interconnection pattern(s) cannot be realized, the switch is said to be blocking. One of the popular blocking networks is Banyan network.

4.4.1 Banyan Network

Banyan networks were first introduced by Goke and Lipovski [4]. Banyan network is a multistage interconnection network (MIN), which usually consists of a number of switching elements (SEs) grouped into several stages interconnected by a set of links. The traditional cross bar has two states, namely, cross state and bar state. There is another type of cross bar, which is often referred as non-traditional, where one input (output) can be directed to (form) both of the outputs (inputs). This type is more powerful but also is more expensive and therefore not popular. The formal definition of Banyan networks is as follows,

- It has N inputs, N outputs, $\log N$ stages and N/2 SEs in each stage.
- There is a unique path between each input and each output.
- Let u and v be two SEs in stage i, and let $S_j(u)$ and $S_j(v)$ be two sets of SEs to which u and v can reach in stage j, $0 < i + 1 = j \leq n$. Then $S_j(u) \Lambda S_j(v) = \emptyset$ or $S_j(u) = S_j(v)$ for any u and v.

Banyan networks are widely used as switch networks or interconnection networks due to its nice properties such as (uniformed connection pattern, self-routing, and short network diameter). There are several well-known Banyan networks, such as Omega, Shuffle Exchange, Butterfly, and Baseline networks.

One of the properties of Banyan network is that it is self-routing. And the routing is decided by the destination. For example, a binary form of a destination is $d_{n-1}\ d_{n-2} \ldots d_0$. When making a routing decision, if the (n-i)th bit of the destination equals to 0, the input of the SE on the connection path in stage i is connected to the SE's upper output link; if it equals 1, the input of the SE on the connection path in stage i is connected to the SE's lower output link. As can be seen in Figure 4.1, the connection is from 001 to 111, since the destination is 111, the path will take the lower, lower, and lower output at each of the states it goes through.

Banyan network has some very nice properties, such as short path and uniform path length. The length of the path is $\log_2 N$, and each path has the same length. Moreover, the number of SEs, or cross point, is $N \log_2 N$. What makes Banyan network even more appealing is its self-routing property, which makes local routing decision making possible. However, Banyan network has a very serious drawback, it being blocking network. For example, in Figure 4.2, two connections, (001-111) and (011-110), both require the lower output link in the second stage, exhibiting a conflict.

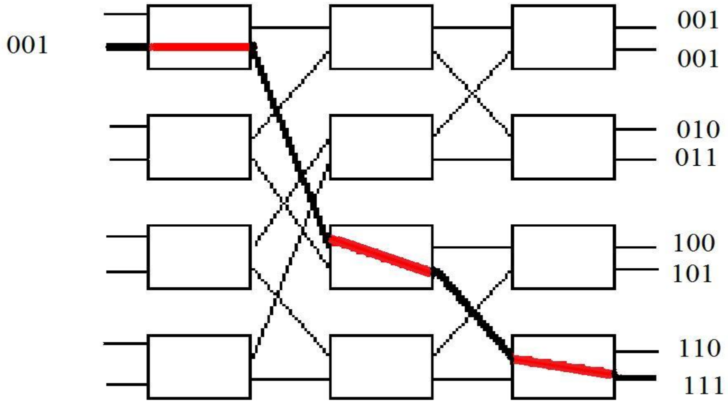


Figure 4.1: Connection path from 001 to 111

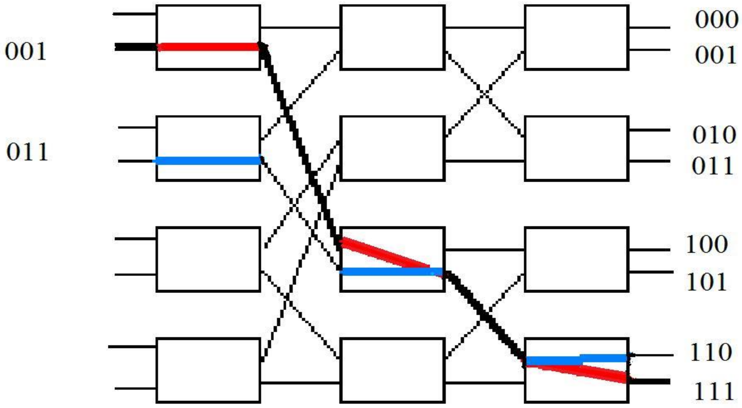


Figure 4.2: Blocking in Banyan Network

As previous noted, there are three types of nonblocking, namely, strictly nonblocking, wide sense nonblocking, and rearrangeably nonblocking, each of which is elaborated further.

4.4.2 Strictly Nonblocking Networks

A strict nonblocking switch allows any unused input to be connected to any unused output regardless of how previous connections were made through the switch. One example of strict nonblocking network is Clos network [13], Clos network has three stages of SEs, which can be implemented by crossbars. The first stage contains r_1 SEs, each of which has n_1 inputs and m outputs, and each of the m outputs goes to each of the m $r_1 \times r_2$ SEs in stage 2; each of the r_2 outputs in stage 2 in turn goes to one of the r_2 $m \times n_2$

inputs in the third stage. A Clos network is usually denoted by $C(n_1, r_1, m, n_2, r_2)$, and if $n_1=n_2=n$ and $r_1 = r_2 = r$, Clos network can be denoted as C(n, m, r), as in Figure 4.3.

If m >= 2n –1, then Clos network is strict nonblocking. The number of SEs used in Clos network is $O(N^{3/2})$ [13]; each path in Clos network goes through the same number of SEs, therefore the signal energy loss along the path is uniformed.

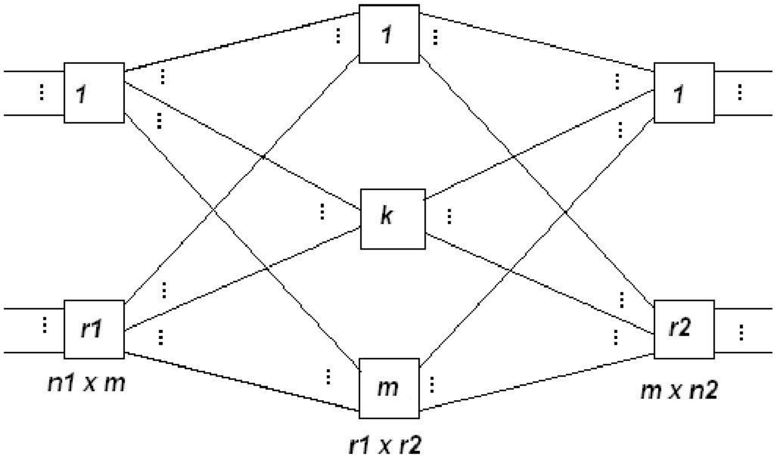


Figure 4.3: 3-stage Clos network[13]

4.4.3 Wide Sense Nonblocking Networks

A switch is said to be wide-sense nonblocking if any unused input can be connected to any unused output, without requiring any existing connection to be rerouted. Wide-sense nonblocking switches usually make use of specific routing algorithms to route connections so that future connections will not be blocked. An example of wide sense nonblocking network is crossbar. A crossbar consists of a matrix of N × N SEs, as in Figure 4.4, to connect input i to output j, the path taken traverses the SE in row i till it reaches column j and then traverses the switches in column j till it reaches output j. Thus the SE on this path in row i and column j must be set appropriately for this connection to be made. An N × N crossbar requires n^2 SEs. The shortest path length is 1 and the longest path length is 2N – 1. Obviously, paths consist of different number of SEs and therefore the signal energy loss along different paths may be different and that is not desirable.

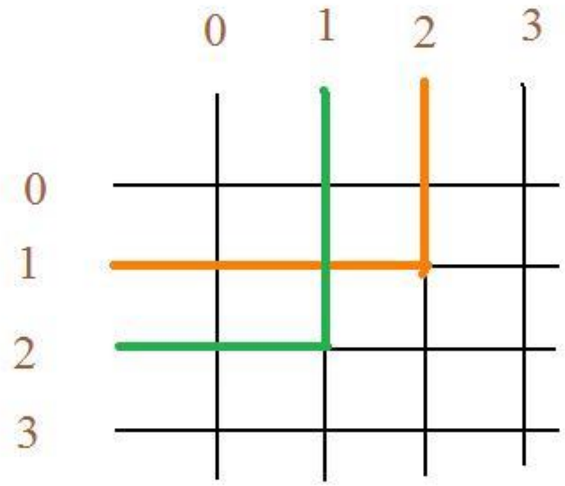
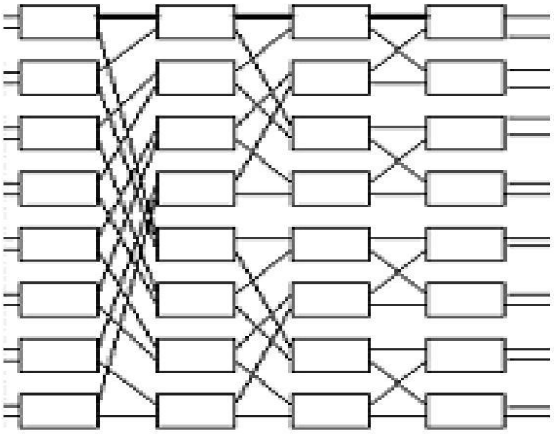


Figure 4.4: A 4 x 4 Crossbar

4.4.4 Rearrangeably Nonblocking Networks

A nonblocking switch that may require rerouting of connections to achieve the nonblocking property is said to be rearrangeably nonblocking. The Benes network is rearrangeably nonblocking switch architecture, and is one of the most efficient switch architectures in terms of the number of 2×2 switches it uses to build larger switches. Banyan-type networks have a single path between an input–output pair. A common design technique for creating alternate paths is to append x extra stages to the back of a regular Banyan-type network in which case the number of paths between an input–output pair becomes 2^x (see Figure 4.5). The maximum number of stages that can be added to such network is (log N – 1), which corresponds to the Banyan network.



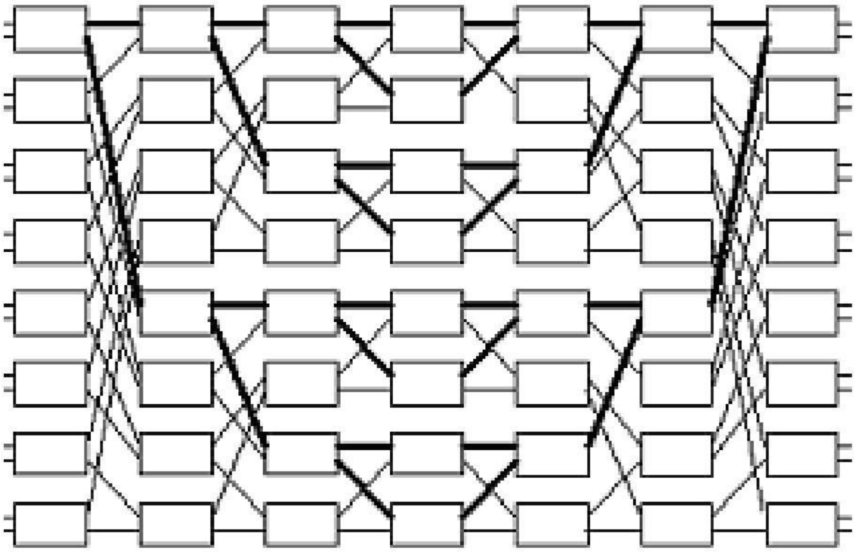


Figure 4.5: By appending extra stages, alternative routes are available

4.5 Problems in OMIN

Fiber optic communications promise to meet the increasing demand of communication systems, and received much attention in parallel processing community as well [10]. Although optical MIN has great promises and has some advantages over the electronic MIN, it leads to some other problems too. Optical MINs suffer with problems such as path loss, conversion of the signal at the switch and crosstalk [11].

4.5.1 Bottlenecks at the Switch Element

An optical cross-connect switch may be thought of as a black box with multiple input and output fibers carrying network traffic. The optical cross-connect switches used in today's networks rely on electronic cores. An optical signal arriving at a switch input port is converted to an electronic signal by a high-speed photo detector (receiver). Electronic circuits in the switch core then direct the signal to the desired output port. A final electrical-to-optical conversion is performed by a laser diode, transforming the signal back into light for onward transmission on the fiber network [11]. The fundamental problem with these electronic cores is that they do not scale well to large port counts (numbers of input and output channels) and are costly to replace for network upgrades to the higher data rates needed for the growing demand for bandwidth. In order to avoid this problem, the need is to develop all optical switching technologies with low-optical-loss switching and extremely high reliability.

4.5.2 Path Dependent Loss

Path dependent loss means that optical signals become weak after passing through an optical path. In a large MIN, a big part of the path-dependent loss is directly proportional to the number of couplers that the optical path passes through. Hence, it depends on the architecture used and its network size. Hence, if the optical signal has to pass through more no of stages or switches, the path dependent loss will be more.

4.5.3 Crosstalk in OMIN

One of the most serious problems is optical crosstalk in optical MIN. This crosstalk occurs when two signal channels interact with each other. When a crosstalk happens, a small fraction of the input signal power may be detected at another output although the main signal is injected at the right output. For this reason, when a signal passes many switching elements, the input signal will be distorted at the output due to the loss and crosstalk introduced on the path [10]. This was not too big an issue in electrical MINs, but because the more stringent bit error rate in optical network, it has become a big problem. There are two ways in which optical signals can interact in a planar switching network. The channels carrying the signals could cross each other in order to embed a particular topology. Alternatively, two paths sharing a SE will experience some undesired coupling from one path to another within a SE [10]. This is shown in Figure 4.6

Each switching element can be in two connecting schemes as shown in Figure 4.6. Since these two ways in Figure 4.6 will cause crosstalk, what need to do is to prevent these situations from occurring in all the switching elements.

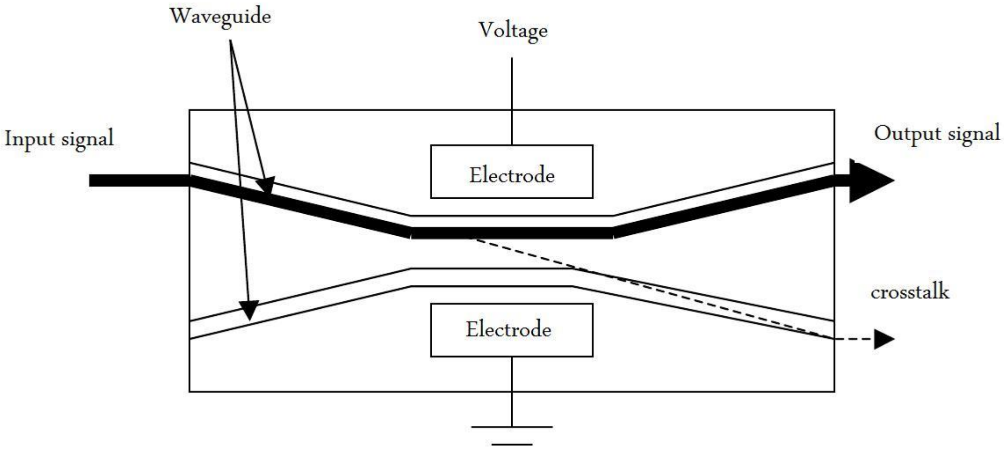


Figure 4.6: Undesired coupling of one path to another within a SE

Crosstalk problem is more dangerous than the path-dependent loss problem with current optical technology. Thus, switch crosstalk is the most significant factor that reduces the signal-to-noise ratio and limits the size of a network. Luckily, first-order crosstalk can be eliminated by ensuring that a switch is not used by two input signals simultaneously. Once the major source of crosstalk disappears, crosstalk in an optical MIN will have a very small effect on the signal-to-noise ratio and thus a large optical MIN can be built and effectively used in parallel computing systems.

4.6 Ways to Solve Crosstalk Problem

4.6.1 Space Domain Approach

One way to solve the crosstalk problem is a space domain approach [10,14], where a MIN is duplicated and combined to avoid crosstalk. The number of switches required for the same connectivity in a network with space domain approach is slightly larger than twice that for the regular network. This approach uses more than double the original network hardware to achieve the same. Thus for the same permutation the hardware or can say the no of switches will be double. Thus cost will be more with the networks using space domain approach. In all the four cases only one input and only one output is active at a given time so that no cross talk occurs. With the space domain approach, extra switching elements (SEs) (and links) are used to ensure that at most one input and one output of every SE will be used at any given time.

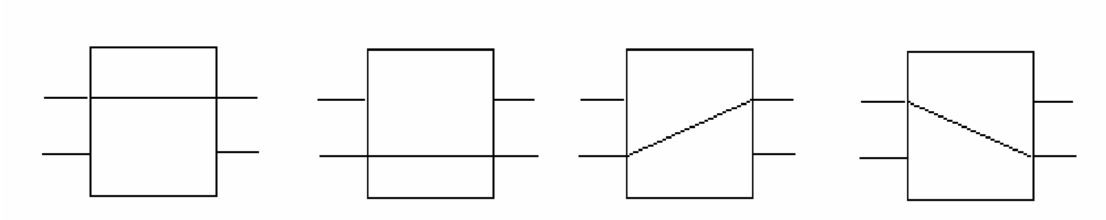


Figure 4.7: Ways to Avoid Crosstalk in the Network using Space Domain Approach

4.6.2 Time Domain Approach

Another way to solve the problem of crosstalk is the time domain approach [10,15]. With the time domain approach, the same objective is achieved by treating crosstalk as a conflict; that is, two connections will be established at different times if they use the same SE. Whereas the messages to be sent to the network should be distribute into several groups, a method is used to find out which messages should not be in the same group because they will cause crosstalk in the network. A set of connections is partitioned into several subsets such that the connections in each subset can be established simultaneously in a network. There is no crosstalk in these subsections. This approach makes importance in optical MINs for various reasons [16]. First, most of the multiprocessors use electronic processors and optical MINs. There is a big mismatch between the slow processing speed in processors and the high communication speed in networks carrying optical signals. Second, there is a mismatch between the routing control and the fast signal transmission speed. To avoid crosstalk, TDM approach is used, which is to partition the set of messages into several groups such that the messages in each group can be sent simultaneously through the network without any crosstalk. If the crosstalk is not allowed in the network that is if the crosstalk in the network is zero the performance of the optical networks will be less. But if crosstalk is allowed in the network in the earlier stages the performance of the networks improves. The performance of the optical networks will be very less than the non optical networks if the crosstalk is zero. But the performance starts increasing if limited crosstalk allowed in the earlier stages of the networks.

CHAPTER 5

Vertically Stacked Optical Banyan Network

A large-scale optical switch is usually composed of numerous basic switching elements (SEs) grouped in multiple stages along with the optical links arranged in a specified interconnection pattern. The basic SEs and the interconnecting optical links in an optical switching device will perform a pre-defined switching function such that the optical flow at an input can be transported to a specific output of the switch. Here the interconnection pattern of the optical links refers to the basic SEs and the input/output ports of the switch, as the network of optical switches. The basic 2×2 SE in optical switching systems is usually a directional-coupler (DC) that is made of two waveguides close to each other [13,17]. DC's can switch multiple wavelengths at the same time, and also at high speed (switching time in the order of ns), which is important for the future optical cross-connects (OXCs). It is notable that DC suffers from an intrinsic crosstalk problem [17,18], in which a portion of optical power in one waveguide of a DC will be coupled into the other waveguide unintentionally when two input optical flows pass through the DC at the same time no matter it is in a BAR or a CROSS status. This undesirable coupling effect is called first-order crosstalk, which may propagate downstream stage by stage, leading to a higher order crosstalk in each downstream stage with a decreasing magnitude. A cost effective solution to the crosstalk problem is to make sure that only one signal passes through a DC at a time such that the first-order crosstalk can be eliminated. Banyan type (e.g. banyan, baseline, omega, shuffle-exchange etc.) networks [4,5] are a class of attractive switching structures for constructing DC-based optical switches, because they have a smaller and exact same number of SEs along any path between an input-output pair such that absolute loss uniformity and smaller attenuation of optical signals are guaranteed in this class of switching networks. A typical N×N banyan network consists of $\log_2 N$ stages, each containing N/2 2×2 switches, and the link connections between adjacent stages are implemented by recursively applying the butterfly interconnection pattern as shown in Figure 5.1.

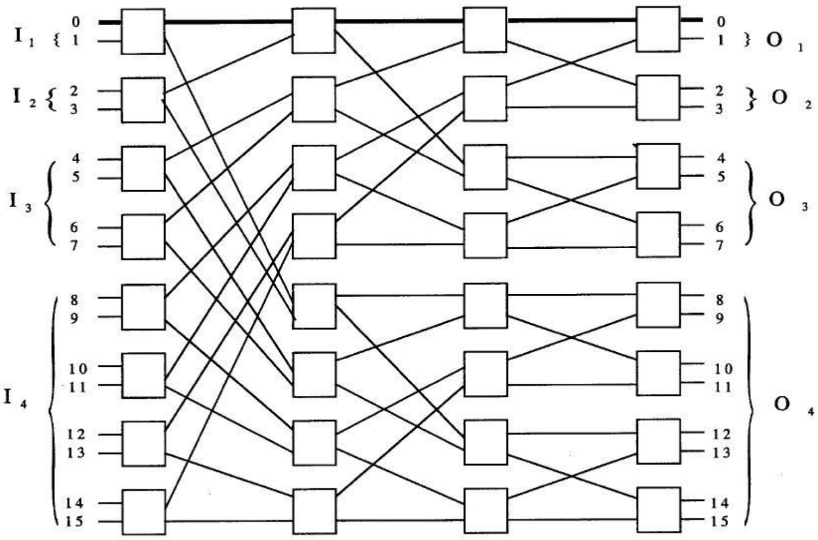


Figure 5.1: Banyan network with butterfly interconnection pattern [24]

However, the banyan topology has only a unique path from each network input to each network output, and cannot connect all the inputs to all the outputs at any time, for which the network is simply degraded as a blocking one. To deal with this situation, it is an effective approach to make the whole network nonblocking by vertically stacking multiple copies of an optical banyan network [19]. This class of networks is called vertically stacked optical banyan (VSOB) networks (Figure 5.2). Numerous studies have been reported for the VSOB networks [20,21,22] with the focus upon the determination of the minimum number of planes required for achieving the nonblocking characteristic. These studies showed that the adoption of the vertical stacking scheme, although attractive, will significantly increase the hardware cost. Khandker, Jiang, Horiguchi, Hong [23] focused on determining the minimum number of stacked copies (planes) required for a nonblocking VSOB networks without link failures if packing strategy is used for routing a request to a plane. It has been shown that required number of planes is minimum when the VSOB network is rearrangeably nonblocking. Due to the increasing importance and requirement for fault-tolerance in optical switches for large mesh WDM networks, performance analysis on VSOB networks at the presence of probability of link failures becomes critical for the practical adoption of the VSOB networks in the current internet applications. The blocking probability of VSOB networks having link failures has been determined by X. Jiang et al. [24]. When there are no link-failures in the switch network, the lower bound on number of planes required to make the network nonblocking is the same as that required for a rearrangeably nonblocking network. However, when there exist some link-failures in the network, it is a challenging job to find the number of planes required to make the network rearrangeably nonblocking. Proof of rearrangeably nonblockingness requires a routing algorithm which can make the network nonblocking allowing rearrangement of the existing connections. In this case the existing routing algorithm (Euler's split algorithm) is not applicable. Therefore, the packing strategy is used for routing signals through the network in the simulation. In packing strategy a banyan plane (of a VSOBN) is packed with maximum number of connections so that all N connections can be established using minimum number of banyan planes. Since new connections are accommodated by rearrangement of the existing connections, the results achieved by this algorithm will be a close approximation of the actual number of planes to make VSOB networks rearrangeably nonblocking when some links fail or are broken in a banyan network.

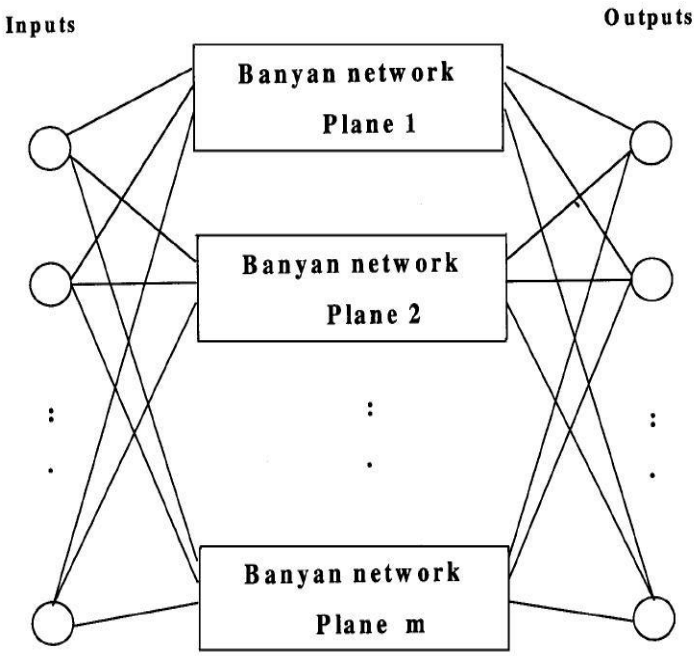


Figure 5.2: Illustration of Vertically Stacked Optical Banyan (VSOB) Network [19]

5.1 VSOB Networks

Based on the vertical stacking scheme, the conditions for a banyan-type network to be rearrangeably nonblocking (free of first-order crosstalk in SEs; refer to this as crosstalk-free hereafter) have been determined [21].

$$P \geq \begin{cases} \sqrt{N} \ \ if \ \log_2 N \ \ is \ even \\ \sqrt{2N} \ \ if \ \log_2 N \ \ is \ odd \end{cases}$$

The blocking probability of a VSOB networks is defined as the probability that a feasible connection request is blocked, where a feasible connection request is a connection request between an idle input port and an idle output port of the network. Due to the topological symmetry, all paths in banyan networks have the same property in terms of blocking. Without loss of generality, focus is on the path between the first input and the first output (which is termed as the tagged path hereafter). All the SEs and links on the tagged path are called tagged SEs and tagged links respectively. The stages of SEs are numbered from left (stage 1) to right (stage $\log_2 N$) and the stages of links are also numbered from left (stage 1) to right (stage $\log_2 N + 1$). For the tagged path, an input intersecting set $I_i = \{2^{i-1}, 2^{i-1} +1, \ldots, 2^i -1\}$ at stage i is defined as the set of all inputs that intersect a tagged SE at stage i. Likewise, an output intersecting set $O_i = \{2^{i-1} , 2^{i-1} +1, \ldots, 2^i -1\}$ associated with stage i contains all the outputs that intersect a tagged SE at stage $\log_2 N - i + 1$.

In VSOB networks, blocking happens when two connections intend to use the same link or SE. However, if a connection is not allowed to pass through a SE to avoid cross-talk in the signal, it automatically resolves the link contentions. That means, a crosstalk free network will be free from link blocking.

5.1.1 Lower Bound on Number of Planes without Link-Failure

To route a feasible connection through a VSOB networks, a routing algorithm must be adopted to find a path from an input port to an output port. To get the lower bound on the blocking probability (i.e. the minimum possible blocking probability) of VSOB network, Jiang-et-al. [24] has considered a packing strategy. Under the packing strategy for a VSOB network, a connection is realized on a path found by trying the least free plane of the network first and most free plane last [25,26]. This routing strategy only guarantees that each of these requests that block a tagged SE will be in a distinct plane in such a way that the minimum number blocked planes can be achieved. If no plane can satisfy the connection request of the tagged path, this request is blocked. This also gives the minimum number of planes required to make the network nonblocking (in other words rearrangeably nonblocking). The lower bound on the number of planes required to make VSOB networks nonblocking without link-failures is shown in Figure 5.3. It is interesting to see that the minimum number of planes required to make a VSOB networks without link failures nonblocking is the same as that required to make the VSOB networks rearrangeably nonblocking. This has been possible because, in packing algorithm, during the routing of new connections, all existing connections are cleared and re-established with new connections altogether. Therefore, some rearrangements of existing connections are taking place in the packing algorithm.

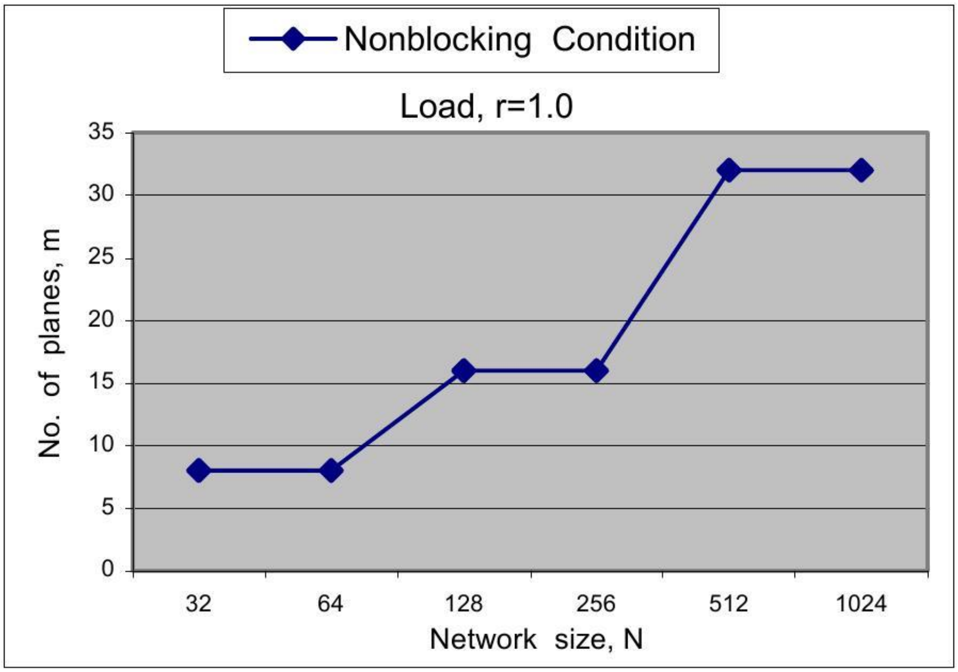


Figure 5.3: Minimum number of planes to make VSOB networks nonblocking without link failures[24]

5.2 VSOB Networks with Link Failures

A connection request may also be blocked by a link failure in a faulty VSOB network, which is referred to as the failure-blocking. Let assume that links in VSOB networks may fail independently and these failures are permanent. Thus, both crosstalk-blocking and failure-blocking should be fully considered in the blocking analysis of a faulty VSOB networks as illustrated in Figure 5.4 for a 8×8 network.

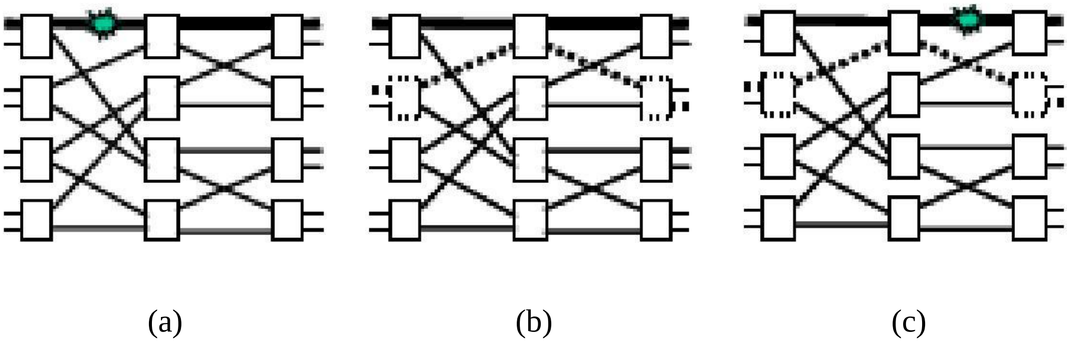

(a) (b) (c)

Figure 5.4: Blocking in a VSOB network: (a) Failure-blocking, (b) Crosstalk-blocking, (c) Combination of failure- blocking and crosstalk-blocking [27]

5.2.1 Minimum Number of Planes for Rearrangeably Nonblocking Networks

The rearrangeably nonblocking condition on number of planes allowing link-failures means the minimum number of planes required to make the network nonblocking when some links are failed in the network. The number of planes required to make the network nonblocking naturally increases if some links in the network are failed or broken. In the simulation it should verify that the minimum number of planes required to make a VSOB networks nonblocking without link failures (i.e. when the link failure parameter, pfr = 0) is the same as that required to make the VSOB networks rearrangeably nonblocking found in the literature. From this result assume that our simulated minimum number of planes for the case when link failures are taken into account is also a close approximation to the actual number of planes to make a VSOB networks rearrangeably nonblocking having link-failures. The network simulator [27,28] developed consists of six major modules as shown in Figure 5.5.

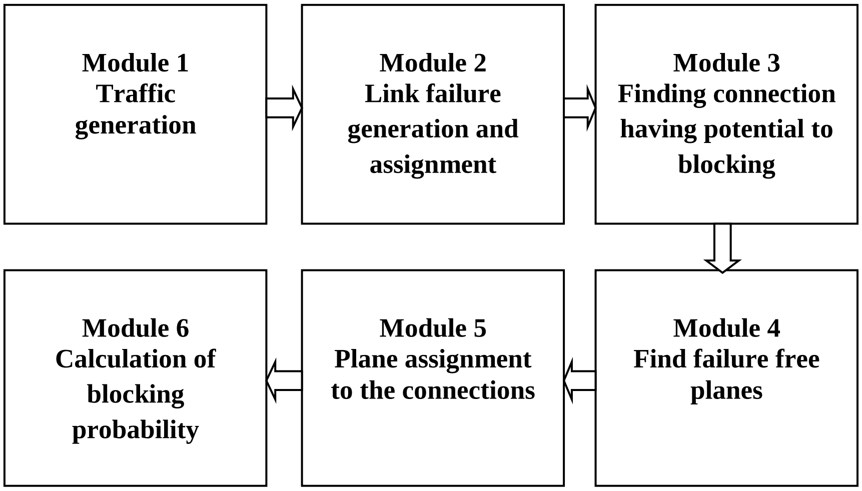


Figure 5.5: Block diagram of network simulator

Here the permutation request is considered as the traffic. A permutation request is such a pattern of requests in which no two inputs request for the same output. Therefore, blocking, if happens, will be due to the internal link-blocking of the switch structure. Due to the symmetric architecture of VSOB (N, T) network, every connection request has the same probability to be blocked. In this simulation, fix the connection request of input-output pair 0-0 and investigate the blocking probability of this connection request only that may result by other contentious connections. The traffic generation module randomly generates a permutation request for the VSOB (N, T) network based on the workload r (here workload r is defined as the occupancy probability of a port). The link failure generation and assignment module generates link failures based on the given pfr (here pfr is defined as the probability that a link is failed or broken) and then assigns those failures randomly to different links.

Here some definitions are given which are used in the discussion of the simulation procedure:

- Set of blocking connections, C_{bc}: The connections that are potential to blocking the tagged path.

- Crosstalk-blocking connections, C_{cbc}: The connections those are potential to crosstalk-blocking the tagged path.
- Set of failure free planes, P_{bc}: The planes those are free from link-failures on the path of each C_{bc}.

- Least free plane: It is defined as the plane which is able to establish minimum number of C_{bc}. For example, let plane $|P_{bc}| = 2$, plane 1 be failure free for 5 C_{bc} and plane 2 be failure Then plane 2 is the least free plane.

- Maximum allowed crosstalk, C_m: C_m is defined the maximum allowed crosstalk along the tagged path.

- Allowed crosstalk, c: c is defined the allowed crosstalk along the tagged path at any time.

Module 1: This module generates random permutations. Each permutation request pass through all the subsequent modules, and the last module checks if the permutation reaches the outputs successfully.

Module 2: This module generates failure pattern with a given probability. These failures are assigned randomly to different links in the switch networks.

Module 3: This module finds C_{bc} which is determined by the following relation:

$$I_i + O_j < \log_2 N \qquad (1)$$

$$I_i + O_j < \log_2 N + 2 \qquad (2)$$

Here, input i is originated from input intersecting group I_i and destined to output j that belongs to output intersecting group O_j (see Figure 5.1).

$$I_i + O_j = \log_2 N + 1 \qquad (3)$$

Module 4: First it checks all the planes if there is a failure on the tagged path, and make a list of planes, say P_{tagged}, in which no links on the tagged path is failed or broken. Then it constructs the list P_{bc} from P_{tagged} and sorts it in the ascending order such that the first plane in the list is least free plane by C_{bc}.

Module 5: The plane assignment module attempts to assign connection requests to different planes using the packing strategy. Plane assignment module groups the input and output connections as per equations (4) and (5).

$$G_i = \begin{cases} \frac{i}{\sqrt{N}} & if\ \log_2 N\ is\ even \\ \frac{i}{\sqrt{2N}} & if\ \log_2 N\ is\ odd \end{cases} \qquad (4)$$

$$G_o = \begin{cases} \frac{o}{\sqrt{N}} & if\ \log_2 N\ is\ even \\ \frac{o}{\sqrt{2N}} & if\ \log_2 N\ is\ odd \end{cases} \qquad (5)$$

Here G_i is the input group for input i and G_o is the output group for output o. If $X(i_1, o_1)$ and $Y(i_2, o_2)$ are two connection requests arrive at inputs i_1, i_2 respectively and destine to o_1, o_2 respectively, then their path is completely disjoint if $G_{i1} \neq G_{i2}$ and $G_{o1} \neq G_{o2}$. Therefore, requests X and Y arrive at the output successfully [29]. The plane which is least free plane (that is the first plane in the sorted list of P_{bc}) for C_{bc} is assigned first to a connection chosen from its list. As soon as the plane is assigned it is marked as ‘busy’ for the input and output groups of that connection’s input and output. The other blocking connections belonged to these input and output groups are not assigned to this busy plane. However this plane can be assigned to a connection having different input and outputs groups. Then the second plane from the sorted list is picked and one or more blocking connections are assigned to the plane as mentioned above. This plane assignment algorithm ensures the maximum use of a banyan plane; thereby ensures the use of minimum number of planes for routing a permutation request. At last it is tried to assign the tagged path to a free plane in list P_{tagged}. If no such free plane is left for the tagged path, the connection request pattern is recorded as a blocked connection pattern.

Module 6: The blocking probability is then estimated by the ratio of number of permutation requests in which the 0-0 request is blocked to the total number of permutation requests generated. It is notable that, in the next permutation request, some inputs request may be the same the previous one and some may be different. However, all previous connections are cleared first and then the next permutation is routed.

The following tables shows the number of planes required to make the network rearrangeably nonblocking with and without link failures, different blocking probabilities and with given crosstalk.

Table 5.1: Minimum number of planes required in VSOB (N,K) with given crosstalks, different workloads and BP[26]

		N=512				**N=1024**			
		C=1	C=2	C=3	C=4	C=1	C=2	C=3	C=4
BP=0	r=1.0	41	39	37	36	63	57	55	53
BP<0.0001%	r=0.5	21	19	16	13	30	29	26	24
	r=0.75	28	23	21	19	37	33	29	24
	r=1.0	33	26	22	21	39	36	32	28
BP<0.01%	r=0.5	17	14	12	11	22	20	19	17
	r=0.75	20	18	16	15	26	22	21	19
	r=1.0	21	19	18	18	29	27	25	24
BP<1%	r=0.5	13	12	11	10	15	13	12	12
	r=0.75	15	13	12	10	18	17	17	16
	r=1.0	17	15	15	14	20	18	18	16

The following table shows the number of planes required to make the vertically stalked optical banyan (VSOB) network nonblocking with link failures. This does not consider the effect of crosstalk on the blocking probability.

Table 5.2: Lower bound of number of planes for nonblocking VSOB network with link-failures[27]

	N=32	**N=64**	**N=128**	**N=256**
pfr=0.002	8	8	16	16

pfr=0.007	10	10	16	16
pfr=0.012	11	12	16	17
pfr=0.017	12	12	18	18
pfr=0.022	13	14	19	20
pfr=0.027	14	14	20	21
pfr=0.032	15	16	21	23
pfr=0.037	16	16	22	23
pfr=0.042	17	18	23	24

Table 5.3: Upper bound of number of planes for nonblocking VSOB network with link-failures[28]

	N=32	**N=64**	**N=128**	**N=256**
pfr=0.002	11	15	23	31
pfr=0.007	13	16	23	31
pfr=0.012	14	18	23	31
pfr=0.017	16	20	24	31
pfr=0.022	17	21	25	31
pfr=0.027	19	23	26	31
pfr=0.032	20	24	27	32
pfr=0.037	21	25	28	33
pfr=0.042	23	26	28	34

CHAPTER 6

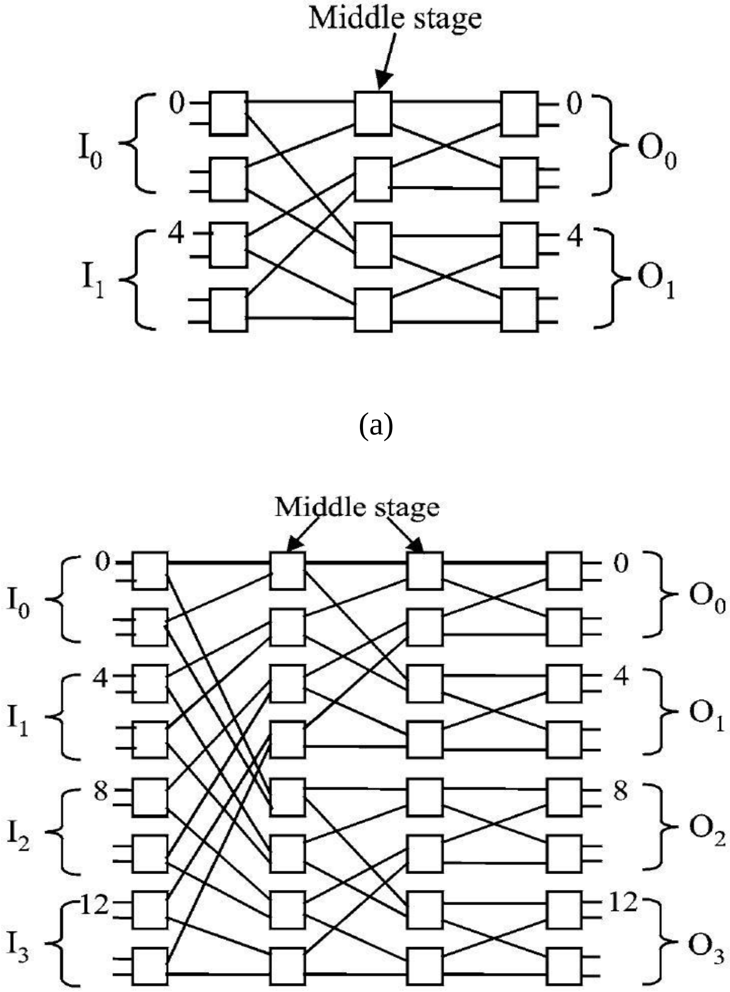
Extended Pruned VSOB (EP-VSOB) Networks

Since banyan-type switching structure has small depth and loss uniformity, the optical versions of this switch network has proposed for eliminating the crosstalk problem by allowing one input of a DC be active at any instant of time and using vertical stacking scheme. To design an N×N crosstalk-free banyan network at least T (T = $2^{\lfloor \log_2 N+1)/2 \rfloor}$) planes are needed to guarantee rearrangeably nonblocking property with an algorithm that has time complexity O($\log_2 N$). Considering that not always a rearrangeably network is essential, a structure with less number of banyan planes (to reduced hardware cost) allowing small amount of blocking can also be designed. However, routing of signals has higher time complexity, either O($N\sqrt{N}\log_2 N$) (for centralized control) or O($\sqrt{N}\log_2 N$) (distributed control). A self-routing vertically stacked optical banyan networks has proposed [23] in which all inputs of at least $\sqrt{N}$ ($\sqrt{2}N$, when $\log_2 N$ is odd) pins apart are permanently tied to a plane and therefore, does not need any plane selection algorithm. This switch structure has the optimum routing time complexity O($\log_2 N$) and the blocking probability that increases to a single digit. A pruned banyan structure (called extended pruned vertically stacked banyan or EP-VSOB in short) has been proposed in [30,31,32]. In this structure, two kinds of banyan planes have been used: pruned and regular banyan. A plane is pruned by eliminating some of its SEs as those are never used in the system. Similar to [23], $\sqrt{N}$ ($\sqrt{2}N$, when $\log_2 N$ is odd) pruned banyan planes and K regular (non-pruned) banyan planes (called extra planes) are stacked vertically. Each input is provided with a connection to one pruned plane and K extra planes. Each input is given chances to search through these extra planes only once in case they fail to find a route in their corresponding pruned planes. This strategy reduces the blocking probability dramatically with only three extra planes, therefore, the time complexity remains almost the same as O($\log_2 N$). Pruning banyan planes also saves hardware cost.

In this the blocking behavior of EP-VSOB network with variable number of pruned planes in addition to extra planes [31] are studied. This approach gives more flexibility to tradeoff among different performance metrics while choosing an EP-VSOB network.

6.1 Extended Pruned VSOB (EP-VSOB) Networks

A Grouping of Inputs for an N×N banyan network [29], the parameter T (where T = $2^{\lfloor \log_2 N+1)/2 \rfloor}$) is used to group its input set I={0,1,...,N-1} into the following disjoint groups I_i as illustrated in Figure 6.1.



(a)

(b)

Figure 6.1: Illustration of inputs and outputs grouping. (a) Network with odd number of stages. (b) Network with even number of stages [31]

It is notable that any two connections generated from two different input groups are crosstalk-free up to the middle stage [24].

6.2 Pruned-VSOB or P-VSOB (N, T) Networks

In this architecture there are T planes, each accepts N/T connections selected from N according to equation (2). Let P_i be the group of inputs connected to plane i, if $X_i \in I_j$ is an input belongs to this group, then

$$P_i = \{X_i | X_i \in I_j, \forall 0 \leq j < N/T\} \qquad (2)$$

It can be seen easily from equation (2) that all N inputs are always evenly distributed among the T planes. Since these N/T connections are fixed at their corresponding inputs in the plane, only I input of a group will be active; all other inputs will remain unused. Therefore, input switches connected to unused inputs and switches in the successive stages corresponding to these unused inputs are redundant. All these redundant switching elements are eliminated. Figure 6.2 explains the idea.

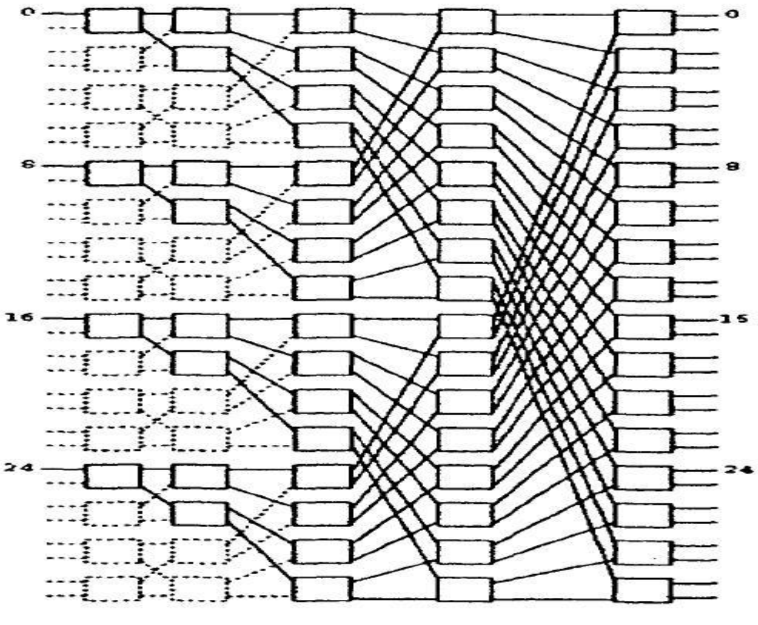


Figure 6.2: Illustration of pruning of a Banyan network. Redundant SEs are drawn in dashed lines [32]

6.3 GENERALIZED EP-VSOB NETWORKS

6.3.1 Input Grouping

The input groups are redefined as shown in equation (3).

$$I_i = \{ iD, iD+1, \ldots, iD+(D-1)\}, 0 \le i \le \frac{N}{D}, D = 2^x, 0 \le x \le \log_2 N \qquad (3)$$

From theorem 4 [24], a banyan network generates zero crosstalk in its first $\log_2 D$ stages if the set P_i connections offered is such that atmost one input from each group D is active and permanently tied to a plane. Equation (2) creates this set P_i of connections which is permanently tied to the plane I. Practically for group size, D = 1, 2 (i.e. x= 0, 1), none of the planes can be pruned because all SEs are used. Pruning comes into effect for group size larger or equal to 4. Figure 6.3 shows a EP-VSOB (16, 8+2) network.

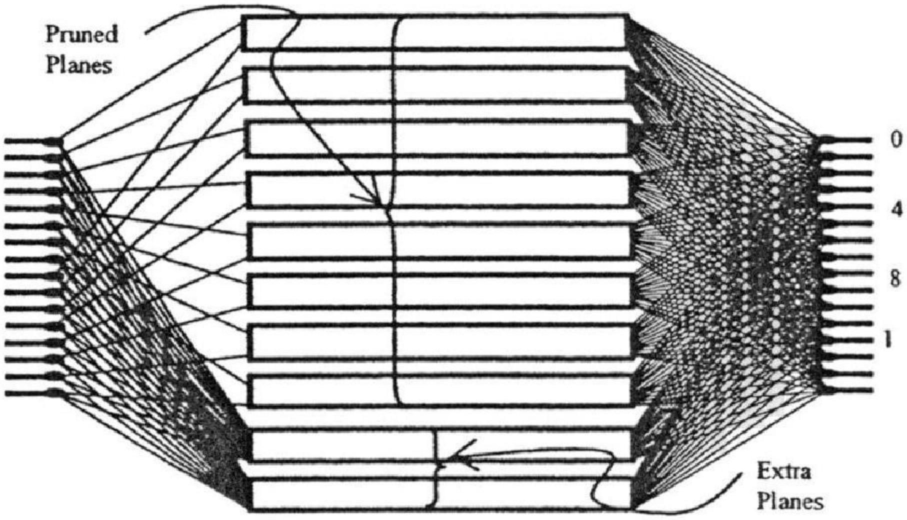


Figure 6.3: The structure of an EP-VSOB (16, 8+2) network [32]

6.3.1 Plane Fixed Routing (PFR):

In the PFR algorithm [30], it selects one input from each input group and tie them with a plane of VSOB (N, T) network permanently. Since all the inputs tied with a plane are group disjoint. One possible plane assignment of PFR algorithm can be performed as following. First define the following subsets of input set I = { 0, 1,..., N− 1}

$$g_i = \left\{i, i+T, i+2T, \ldots, i+\left(\frac{N}{T}-1\right)T\right\}, 0 \leq i \leq T-1.$$

all inputs are tied in set g_i to the plane i of the VSOB (N, T) network (suppose the planes are numbered as 0, 1,...,T−1). For example, the subsets g_i ($0 \leq i \leq 3$) for VSOB (16, 4) network are determined as g_0 = { 0, 4, 8, 12} , g_1 ={ 1, 5, 9, 13} , g_2 ={ 2, 6, 10, 14} ,and g_3 ={ 3, 7, 11, 15} , and the final plane assignment of inputs is illustrated in figure 6.4. A switching network is said to be self-routing if each input-signal can determine its path to the destined output by its destination address, regardless of the destination addresses of other connections. Since each plane of VSOB (N, T) is a self-routing banyan network, the VSOB (N, T) network employing PFR scheme is actually a self-routing network because there exists only a unique path between each input–output pair in such network. Therefore, a VSOB (N, T) network employing PFR scheme has the routing time complexity of O(log N) that is proportional to number of stages of the network, and this time complexity is optimum because it need at least O(log N)time to establish a path in a log N – stage banyan network.

It is notable that for a VSOB (N, T) network employing PFR scheme, it only guarantee the nonblocking (crosstalk-free) property in the first half stages, and blocking may happen in the last half stages. However, results indicate that the blocking probability is reasonably low for a large VSOB (N, T) network adopting PFR scheme.

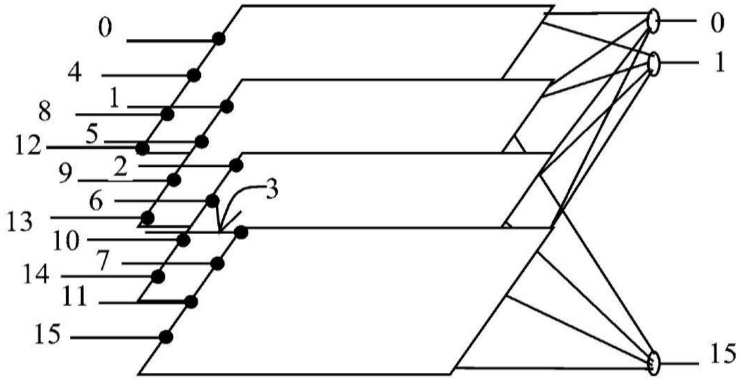


Figure 6.4: Plane assignment for inputs in VSOB (16, 4) network employing PFR algorithm [29]

Depending on how a plane is chosen from the extra planes, two routing algorithm presented for fast connection establishment in an EP-VSOB (N,T+K) network, namely PFR with linear search algorithm and PFR with random routing algorithm.

6.3.1.1 PFR with Linear Search (PFR_LS)

In the PFR_LS algorithm for an EP-VSOB (N, T+K) network, each connection now has (1+K) chances to be established through the network. Whenever a request arrives at an input, first it is sent to the fixed plane of that input. If the request cannot be established in the fixed plane, the input searches for a free plane among the extra K planes. The searching starts with the first plane and continues orderly up to the last one so that each plane is being searched at most one time by an input. If the connection request still fails to find a free plane among all K extra planes, it is considered as a blocked request. The time complexity of the PFR_LS algorithm O($\log_2 N$) when K is constant and much smaller compared to T.

6.3.1.2 PFR with Random Selection (PFR_RS)

In the PFR_RS algorithm for an EP-VSOB (N, T+K) network, each connection has two chances to be established through the network. The first chance is to establish the connection through its fixed plane. If the connection request is blocked in its fixed plane, it still has a second chance to be established through another plane selected randomly from K extra planes. If the connection request still could not be established through the randomly selected plane, it is considered as a blocked request. It is easy to see that the time complexity of the PFR_RS algorithm remains the optimum O($\log_2 N$).

6.4 Blocking Probability of EP-VSOB (N, D+K) Network

The permutation request is considered as the traffic since a permutation does not have output contention, and therefore, gives real blocking probability of the switch network only. Due to the symmetric architecture of VSOB network, every connection request has the same probability to be blocked. Here the connection request of input-output pair 0-0 is fixed and investigate the blocking probability of this connection request only. The traffic generation module randomly generates a permutation request for a network based on r, the occupancy probability of a port. The plane selection module attempts to assign connection requests to different planes using the routing algorithm. The number of fixed planes is decided by D. Connections assigned to a plane are established by self-routing algorithm. The blocking probability is then estimated by the ratio of the number of permutation requests in which the 0-0 request is blocked to the total number of permutation requests generated. Blocking probabilities of an N×N network with different fixed planes and different extra planes are simulated.

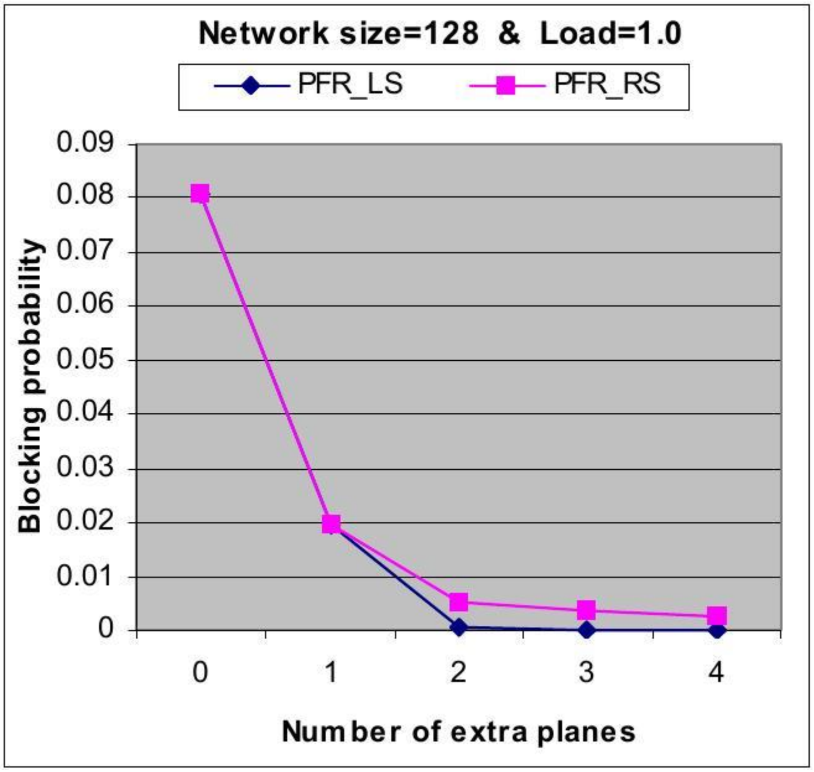


Figure 6.5: Blocking probability EP-VSOB networks with PFR_LS and PFR_RS algorithm[31]

6.5 EP-VSOB Networks with Link Failures

In this the blocking behavior of EP-VSOB networks having link-failures [33] using PFR_LS and PFR_RS algorithm is analyzed.

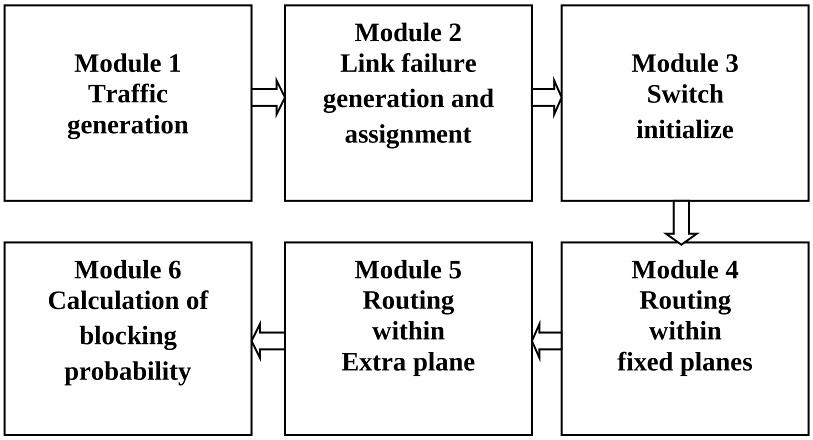


Figure 6.6: Block diagram of network simulator

The permutation request here is considered as the traffic since a permutation does not have output contention, and therefore, gives real blocking probability of the switch network only. Due to the symmetric architecture of P- VSOB (N, T) network, every connection request has the same probability to be blocked. In simulation, the connection request of input–output pair 0–0 is fixed and investigates the blocking probability of this connection request only.

Module 1: This module randomly generates a permutation request for the VSOB (N, T) network based on the workload r (here workload r is defined as the occupancy probability of a port).

Module 2: This module generates link failures based on the given pfr (here pfr is defined as the probability that a link is failed or broken) and then assign those failures randomly to different links.

Module 3: This module initializes the switches with the permutation request.

Module 4: This module attempts to assign connection requests to different planes using PFR algorithm. It only consider the plane that contains the tagged path. It tries to establish the tagged path in the selected plane. If the tagged path is not established, then try to assign other connections in their selected planes.

Module 5: If the tagged path is not established in the selected plane then try to establish the tagged path (along with other connections that are not established in their selected plane) in the extra planes. In case of PFR_LS algorithm, the searching starts with the first plane and continues orderly up to the last one. If the tagged path still fails to find a free plane among all extra planes, it is considered as a blocked request. In case of PFR_RS algorithm, it randomly selects a plane from extra planes. If the tagged path still could not be established through the randomly selected plane, it is considered as a blocked request.

Module 6: In this module the blocking probability is estimated by the ratio of number of connection requests in which the 0–0 request is blocked to the total number of connection requests generated.

6.5.1 PFR_LS algorithm

The simulation results of EP-VSOB networks using PFR_LS algorithm with link failures are given below in Table 6.1.

Table 6.1: Blocking probability of extended pruned VSOP network using PFR_LS algorithm[33]

	Blocking Probability			
	Ep=0	Ep=1	Ep=2	Ep=3
Pfr=.005	0.09	0.025	0.001	0
Pfr=.01	0.14	0.030	0.010	0.005
Pfr=.015	0.17	0.060	0.020	0.010
Pfr=.02	0.20	0.070	0.030	0.020
Pfr=.025	0.22	0.060	0.040	0.025
Pfr=.03	0.26	0.065	0.050	0.025

Pfr=.035	0.29	0.160	0.060	0.030
Pfr=.04	0.32	0.190	0.080	0.040

It is interesting to note that the blocking probability not always increases with the increase of link failure; blocking probability decreases for certain range of link failures and then increases again. The reason for decreasing the blocking probability may be as follows.

When there are some links failed on the path of potential blocking connections, they cannot interfere with the tagged path. This phenomenon increases the tagged path's chance of being successful. Table 6.1 shows the results of blocking probability for given link failure probability (pfr) with different number of extra planes (Ep).

6.5.2 PFR_RS algorithm

The simulation results of EP-VSOB networks using PFR_RS algorithm with link failures are given below. Table 6.2 only present the results of extra planes 2 & 3 because for extra plane 1 the blocking behavior of both PFR_LS & PFR_RS algorithm is the same.

Table 6.2 shows that blocking probability decreases linearly with the number of extra planes (Ep). For N=128, pfr=0.01 & r=1; if it adds 2 extra planes to the networks then the blocking probability decreases 81% and 85% for 3 extra planes.

Table 6.2: Blocking probability of extended pruned VSOP network using PFR_RS algorithm[33]

	Blocking Probability	
	Ep=2	Ep=3
Pfr=.005	0.018	0.018
Pfr=.01	0.028	0.025
Pfr=.015	0.040	0.040
Pfr=.02	0.062	0.038
Pfr=.025	0.060	0.060
Pfr=.03	0.090	0.070
Pfr=.035	0.138	0.085
Pfr=.04	0.145	0.10

Conclusion

- Vertically Stacked optical Banyan (VSOB) networks are attractive for serving as optical switching system due to the good properties of banyan structures (such as the small depth and self-routing capability), but this usually require a high hardware cost to build. The study of VSOB networks from probabilistic view and both the upper and lower bounds on blocking probabilities can guide the network designers to estimate the maximum blocking probability of the VSOB networks in which different routing strategies may be applied. These bounds provide quantitative measurements for tradeoffs between network hardware cost and blocking probability.
- The Vertically Stacked optical Banyan (VSOB) networks can accommodate some link failures without degrading the performance. That means a rearrangeably nonblocking VSOB network can still be nonblocking even when some links are broken or failed. The blocking probability does not monotonously increase with the increase of link failure probability.
- The plane fixed routing algorithm (PFR) is a self-routing algorithm, so far the first of this kind for VSOB network, with low blocking probability, optimum routing complexity and attractive hardware saving property. The performance of PFR_ LS algorithm is better than PFR_RS algorithm.

Future Scope

- New techniques can be developed to reduce the blocking probability.

- The routing algorithms for faster connections and making the VSOB networks having link-failures rearrangeably nonblocking can be develop.

- The concept of vertical stacking can be used for fault-tolerant multipath optical MINs.

- The study can be extended to irregular optical networks.

References

[1] Anujan Varma and C. S. Raghavendra, “Interconnection Networks for Multiprocessors and Multicomputers, Theory and Practice,” IEEE Computer Press, 1994.

[2] J. Duato, S. Yalamanchili, L. Ni, “Interconnection Networks,” Morgan Kaufman, 2002.

[3] N. Das, B. B. Bhattacharya, R. Menon, A. Sarkar, “Permutation Routing in Optical MINs with Minimum Number of Stages,” Journal of Systems Architecture, vol. 48, pp. 311-323, 2003

[4] L. R. Goke and G. J. Lipovski, “Banyan networks for partitioning multiprocessing systems,” Proceedings of the First International Symposium on Computer Architecture, pp. 21-28, 1973.

[5] J. H. Patel, “Performance of processor-memory interconnections for multiprocessors,” IEEE Transactions on Computers, vol. C-30, pp. 771-780, October 1981.

[6] Kruskal C.P., Shir M., “Performance of multistage interconnection networks for multiprocessors,” IEEE transactions on computers, vol. C 32, December 1983.

[7] Charles Chenggong Charles and Bruck Jehoshua, “Tolerating Multiple Faults in Multistage Interconnection Networks with Minimal Extra Stages,” IEEE Transactions on Computers, vol. 49, no. 9, pp. 998-1004, September 2000.

[8] Sengupta J., Bansal P. K. and Gupta Ajay, “Permutation and Reliability measures of Regular and Irregular MINs,” International Conference IEEE, pp. 531-536, 2000.

[9] Bhogavilli Suresh K. and Abu-Amara Hosame, “Design and Analysis of High Performance Multistage Interconnection Networks,” IEEE Transactions on Computers, vol. 46, no. 1, pp. 110 -117, January 1997.

[10] Yi Pan, C. Qiao, and Y. Yang, “Optical Multistage interconnection Networks: New Challenges and Approaches,” IEEE Communications Magazine, vol. 37, no. 2, pp. 50-56, February 1999.

[11] Regis J. Bates, “Optical Switching and Networking Handbook,” McGraw-Hill, February 2001.

[12] C. Siva Ram Murthy and Mohan Gurusamy, “WDM optical networks, concepts, design, and algorithms,” Prentice Hall, Upper Saddle River, N.J., 2002.

[13] Rajiv Ramaswami and Kumar N. Sivarajan, “Optical networks, a practical perspective,” Morgan Kaufmann Publishers, 2002.

[14] K. Padmanabhan and A. N. Netravali, "Dilated networks for photonic switching,” IEEE Transactions on Communication, vol. 35, no. 12, pp. 1357-1365. December 1987.

[15] Qiao C., “A Time Domain Approach for avoiding Crosstalk in Optical Blocking Multistage Interconnection Networks,” Journal of Lightwave Technology, vol. 12, no. 10, pp. 1854-1862, 1994.

[16] Jianchao Wang and Yi Pan “Permutation Capability of Optical Multistage Interconnection Networks,” August 24, 2000.

[17] H. S. Hinton, “An introduction to Photonic Switching Fabrics,” Plenum, New York, 1993.

[18] V. R. Chinni, T. C. Huang, P. K. A. Wai, C. R. Menyuk, and G. J. Simonis, “Crosstalk in a lossy directional coupler switch,” IEEE/OSA Journal of Lightwave Technology, vol. 13, no. 7, pp. 1530-1535, July 1995.

[19] C. T. Lea, “Muti-log2N networks and their applications in high speed electronic and photonic switching systems,” IEEE Transactions on Communication, vol. 38, pp. 1740-1749, October 1990.

[20] G. Maier and A. Pattavina, “Design of photonic rearrangeable networks with zero first-order switching-element-crosstalk,” IEEE Transactions on Communication, vol. 49, no. 7, pp. 1268-1279, July 2001.

[21] M. M. Vaez and C.T. Lea, “Strictly nonblocking directional-coupler-based switching networks under crosstalk constraint,” IEEE Transactions on Communication, vol. 48, no. 2, pp. 316-323, February 2000.

[22] M. M. Vaez and C. T. Lea, “Space-wavelength tradeoff in the design of nonblocking directional coupler based network under crosstalk constraint,” Journal of. Lightwave Technology, vol. 16, pp. 1373-1379, August 1998.

[23] M. M. R. Khandker, X. Jiang, P. H. Ho, S. Horiguchi and H. T. Mouftah, “Performance of Fast Routing Algorithms in Large Optical Switches Built on the Vertical Stacking of Banyan Structures,” Cluster Computing, vol. 7, no. 3, pp. 219-224, July 2004.

[24] X. Jiang, M. M. R. Khandker, H. Shenand, S. Horiguchi, “A Nonblocking Optical Switching Network for Crosstalk-free Permutation,” IEICE Transactions. on Communication, vol. E86-B, no. 12, December 2003.

[25] X. Jiang, M. M. R. Khandker and S. Horiguchi, “Nonblocking Optical MINs Under Crosstalk-free Constraint,” Proceedings of the 2001 IEEE Workshop on High Performance Switching and Routing, pp. 307-311, Dallas, USA, May 2001.

[26] X. Jiang, P. H. Ho, H. Shenand and S. Horiguchi, “Fault Tolerance Analysis of Optical Switching Systems Built on the Vertical Stacking of Banyan Network,” Proceeding, IEEE Workshop on High Performance Switching and Routing, 2004.

[27] Basra Sultana, M. M. R. Khandker, “On number of Planes of Rearrangeably Nonblocking Optical Banyan Networks with Link Failures,” Journal of Scientific Research, vol. 1, no. 1, pp. 43-45, December 2008.

[28] Basra Sultana, M. M. R. Khandker, X. Jiang and S. Horiguchi, “Blocking Probability of Vertically Stacked Optical Banyan Networks with Link Failures,” Proceeding. The International Workshop on High Performance and Highly Survivable Routers and Networks, Tohoku University, Sendai, Japan, pp. 157-169, March 14, 2007.

[29] M. M. R. Khandker, X. Jiang, P.H. Ho, and S. Horiguchi, “Blocking Probability of Self-routing Vertically Stacked Optical Banyan Switching Networks with given Crosstalk constraints,” 3rd International Conference on Electrical & computer Engineering ICECE, December 2004.

[30] M. M. R. Khandker, X. Jiang, Masuru Fukushi and S. Horiguchi, "Pruned Optical Banyan Networks on Vertical Stacking Scheme with Faster connection Setup Algorithms," Optics Communications, The Netherlands. 26 September 2005.

[31] M. M. R. Khandker, X. Jiang, Masuru Fukushi, S. Horihuchi, “Pruned optical banyan networks on vertical stacking scheme for faster connection establishment,” Graduate School of Information Sciences, Tohoku University, Sendai, Japan, 8 September 2005.

[32] Basra Sultana, M. M. R. Khandker, “Blocking Behavior of Crosstalk-free Pruned Optical Banyan Networks,” IEEE Malaysia International Conference on Communications and IEEE International Conference on Networks (MICC & ICON 2005), Kuala Lumpur, Malaysia, November 2005.

[33] Basra Sultana, M. M. R. Khandker, “Blocking Behavior Analysis of Extended Pruned Vertically Stacked Optical Banyan Networks with Link Failures,” 5TH International Conference on Electrical and Computer Engineering ICECE-2008, December 2008.

Papers Published

K. Nagarjun, Rinkle Aggarwal, "Behavior Analysis of Optical Banyan Networks on Vertical Stacking," National Conference on Advances in Computer Networks & Information Technology NCACNIT-09, Gurujambheswar University, Hisar, 24-25 March 2009.