Hierarchical Optical Ring INterconnection (HORN): A Scalable Interconnection-Network for Multiprocessors and Massively Parallel Systems

Ahmed Louri and Rajdeep Gupta

Department of Electrical and Computer Engineering
The University of Arizona
Tucson, AZ 85721

Abstract

A new interconnection network for massively parallel computing is introduced. This network is called a Hierarchal Optical Ring INterconnection (HORN). HORN consists of a single-hop, scalable, constant degree, strictly non-blocking, fault-tolerant interconnection topology that utilizes Wavelength Division Multiple Access (WDMA) to provide better utilization of the terahertz (THz) bandwidth offered by optics. The proposed optical network integrates attractive features of hierarchical ring interconnections, e.g., simple node interface, constant node degree, better support for locality of reference, fault-tolerance, with the advantages of optics. This paper presents the HORN topology, analyzes its architectural properties and presents an optical design methodology for it. Furthermore, a brief feasibility study of HORN is conducted. The study shows that the topology is highly amenable to optical implementation using commercially available optical elements.

1 Introduction

Parallel processing systems are a proposed solution to the increasing demands for processing power and computation speeds. These systems can consist of thousands of processing elements (PEs) interconnected via an interconnection network such as in Massively Parallel Processing (MPP). Due to the large number of PEs contained in these systems, the interconnection network usually determines performance and cost. Such an network must have low interconnection complexity (such as a low node degree, thus low cost and ease of implementation), relatively small diameter for such a large number of PEs, a high degree of scalability and expandability, and most importantly, efficient support for both local and remote communications. Recent studies [1, 2] have shown that efficient implementation of local communications (spatial locality) is a fundamental requirement for interconnection networks since PEs engage in data transfers more frequently with nearby neighbors than with more distant PEs.

It is proving to be very difficult for flat interconnection networks to satisfy the above requirements, especially scalability to a large number of PEs, while still maintaining a small diameter and low cost. Recently there has been strong interest in hierarchical interconnection networks [3, 4] that can provide a high degree of scalability while still maintaining a low network latency. The rationale behind hierarchical networks is based on the locality of reference found in the communication profiles of many parallel processing applications. Therefore, it is desirable to have cluster-based interconnection networks where a cluster is comprised of a relatively small number of PEs. The intra-cluster level should efficiently support local communication, while global communication will take place at the inter-cluster level. An additional advantage of hierarchical networks is modularity, but as the number of PEs increases and the performance of each PE increases, the demand for higher communication bandwidths and higher interconnect densities also increases. There are, however, some serious technical challenges to making these systems a reality.

A possible solution to the realization of interconnection networks for large parallel processors and MPPs is the use of optical technology [1, 5, 6, 7]. Optics provides many features such as parallelism, large bandwidth, low power requirements, reduced crosstalk, and better isolation than semi-conductor electronics can provide. To exploit the terahertz (THz) bandwidth of optics for large parallel processors, Wavelength Division Multiple Access (WDMA) techniques that enable multiple multi-access channels to be realized on a single physical channel can be utilized. In a WDMA network, the optical spectrum is divided into many different logical channels, each channel corresponding to a different wavelength. These channels can be carried simultaneously on a small number of physical channels, e.g., optical fibers. Additionally, each network node is typically equipped with a small number of transmitters and receivers (transceivers), some of these being dynamically tunable to different wavelengths. For a single-hop packet transmission to occur, one of the transmitters of the sending node and one of the receivers of the destination node must be tuned to the same wavelength for the duration of the packets' transmission.

Several WDMA-based network architectures have been introduced recently [8, 9, 10, 11, 12]. This list is by no means complete but gives us a broad outlook on types of WDMA networks. Some of these architectures

0-8186-7591-8/96 $5.00 © 1996 IEEE
Proceedings of MPPOI '96

are not size-scalable to large numbers [8, 9, 11], while other architectures are multi-hop [9] where a packet may not remain completely in the optical domain between source and destination. This incurs a major delay at each intermediate node due to optical/electrical conversions and processing of the packet for routing and re-transmission. Some of these networks [10] require tunable transmitters and receivers at each PE which is very costly at this time. Other networks [10, 11, 12] suffer from splitting losses incurred from star couplers. Finally, no distinction is made between local and remote communications in any of these networks which has significant performance implications..

The above considerations have led us to look into optical hierarchical networks to circumvent the disadvantages of current WDMA-based networks listed above. To this end, we present a novel interconnection topology, known as the Hierarchical Optical Ring INterconnection (HORN). HORN is based on a ring of buses hierarchical paradigm and consists of a single-hop, scalable, nonblocking topology. Cost-savings is accomplished through the low node degree while maintaining scalability and achieving excellent performance through the use of WDMA and single-hop techniques. Packets are sent from the source node on a distinct wavelength and arrive at the destination node with the same wavelength. No wavelength reconfiguration is required for changes in traffic. A distinction is made between local and remote communications where both are implemented independently of one another. Through wavelength re-use, we are able to implement both local and remote communications efficiently. PEs consist of a single, (slow) tunable transmitter and a small set of non-tunable receivers, consequently not requiring tunability at both ends. A connection between any two nodes does not require PEs to forward packets, and no optical to electrical (o/e) and electrical to optical (e/o) converters are required during routing, hence a single-hop architecture. Finally, fault tolerance is enforced through the use of dual rings.

2 Topology of HORN

In this section we define the structure of HORN including the wavelength assignment used, message routing, diameter, link complexity, fault tolerance, and an example of a multiple access protocol.

2.1 Definition of HORN

It has been shown that a PE engages in data transfer more frequently with nearby neighbors (local communication) than with more distant nodes (remote communications) [2, 13]. Therefore, the interconnection topology must be designed so that it can efficiently support local data transfers (spatial locality). This emphasis has led us to consider a hierarchical interconnection network topology in which the lower level network supports local communications very efficiently. We have chosen the snowflake topology [2, 13] because it is well suited for this type of communication.

HORN is an optical interconnect approach that achieves the architectural objectives of snowflakes while also providing significant performance improvements in elements such as unity diameter, fault tolerance, nonblocking capability, and scalability through the use of WDMA and wavelength re-use. The dual rings of HORN are used strictly for routing and for fault tolerance. WDMA is used to achieve multiple logical channels without requiring multiple physical links.

Figure 1 shows a diagram of a three-level HORN where all PEs are located in the first hierarchical level $H(1)$. PEs in this figure are identified as black filled circles and switching nodes are identified as gray filled circles. Switching nodes are located at $H(i)$, where $2 \leq i \leq 3$, and are used for routing purposes. The notation $HORN(i)$ is used to characterize HORN where i represents the number of hierarchies. Figure 1 shows a diagram of a $HORN(3)$ where all groups are labeled using $H(n, g)$ notation such that $H(n, g)$ is used to identify individual groups of HORN, where n refers to the hierarchy and g refers to a group at hierarchy n. The dual rings of HORN can be seen in this figure.

Two types of communication are possible in HORN: local and remote. In both cases, a packet undergoes o/e conversion only at the source and destination and no further o/e conversion is required during routing. Local communication takes place when both the source and destination PEs are in the same hierarchical group $H(1, g)$ where, in Figure 1, $1 \leq g \leq 18$. By contrast, remote communication takes place when the source and destination PEs are in different hierarchical groups. We have separated local and remote communications from one another in order to provide a more efficient implementation for both communications. Local communication employs the inner ring and remote communication employs the outer ring as shown in Figure 1. Switching nodes, therefore, are not used for local communication but are used in remote communication. Efficient implementation of local and remote communications is accomplished through the novel wavelength assignment which is discussed in the next section.

2.2 Optical Wavelength Assignment for HORN

Assigning a unique wavelength to all PEs would be an ideal solution since it would make packet routing a trivial task. However, there is a limited number of available wavelengths, restricting the interconnection size [10, 11]. The number of wavelengths available determines the number of logical channels that is supported by a single line of the interconnection. While this number may be large when considered from an information-capacity point of view, it may not be large enough to support the number of PEs needed for a massively parallel architecture. One method of overcoming this limitation is to re-use wavelengths. Wavelengths are re-used in HORN by allowing those used in local communication to be used in remote communication.

The number of wavelengths employed for local communication equals the maximum number of PEs located in the rings of the first hierarchy of HORN (N_1).

$$N_1 = Max[|H(1, i)|\forall i] \qquad (1)$$

Figure 1 shows an example of the wavelength assignment of the $H(1, 8)$ group. The wavelengths located next to each PE correspond to the wavelength that each PE receives. This same wavelength assignment applies to all rings located in the $H(1)$ hierarchy.

An ideal wavelength assignment for remote communication would be one such that a unique wavelength is assigned to all distinct rings of HORN. Figure 2 shows an example of such a wavelength assignment for Figure 1. Notice that the wavelength assignment for local communication is also shown for completeness. Communication takes place with a source PE sending data packets on the wavelength assigned to the destination PE's ring. For example, PEs wanting to send to group $H(1,12)$ do so by sending on λ_{12}, with all PEs in this respective ring consuming the packet. PEs, therefore, receive packets sent on the wavelength assigned to their ring as well as on the wavelengths assigned to higher level rings. For example, PEs located in $H(1,1)$ receive packets sent on λ_1, λ_{19}, and λ_{22}. Consequently, multicast and broadcast capabilities are very naturally handled in HORN. PEs wanting to multicast to groups $H(1,7)$, $H(1,8)$, $H(1,9)$, $H(1,10)$, $H(1,11)$, and $H(1,12)$ do so by sending on λ_{20}, and PEs wanting to broadcast to all PEs do so by sending on λ_{22}.

All PEs consist of a number of fixed tuned receivers that correspond directly to the number of wavelengths they receive on. Assuming the wavelength assignment shown in Figure 2, each PE is assigned one wavelength for local communication, and h wavelengths for remote communication. Consequently, $h+1$ receivers are required for each PE. The number of wavelengths on which each PE receives is a small subset of the total number of available wavelengths in HORN and is equal to 18% for the wavelength assignment of figure 2.

2.3 Message Routing in HORN

The design of an interconnection network must permit efficient routing. PEs must be able, at any point in time, to establish a route to an intended destination. The interconnections need each of the PEs to communicate with the intended destinations. That this communication be established is an essential parameter in the design of the interconnection. In HORN, PEs communicate with destination nodes through either local or remote communication.

One of the novel features of the HORN routing protocol is the separation of local and remote communications. By physically separating the two protocols, the routing paths do not coincide with one another as shown in Figure 1.

Local communication takes place when both the source and destination PEs are in the same hierarchical group, $H(1,a)_{source} = H(1,a)_{destination}$. The source PE tunes its transmitter to the preassigned wavelength of the destination PE and transmits. The destination PE subsequently consumes the packet. Moreover, a simple WDMA concept is employed, and a diameter of one is achieved for local communication.

Remote communication takes place when the source and destination PEs are not in the same hierarchical group, $H(1,a)_{source} \neq H(1,a)_{destination}$. The key to the routing used in remote communication is the use of acousto-optic tunable filters (AOTFs) [14, 15], located in the switching nodes, that are able to route on individual wavelengths. AOTFs, therefore, can be thought of as optical switches. Configuration of the AOTF to a given routing algorithm is accomplished by sending an appropriate acoustic wave. Once the AOTF is configured, optical packets experience no delay other than the propagation delay through the acousto-optic cell. Thus, AOTFs are able to operate as transparent optical switches.

2.4 Multiple Access Protocols

HORN requires a multiple-access protocol in order to prevent packets of the same wavelength from colliding with one another. Examples of multiple access protocols include CSMA (Carrier Sense Multiple Access), CSMA/CD (Carrier Sense Multiple Access with Collision Detection), ALOHA, Slotted ALOHA, TDMA (Time Division Multiple Access), and Arbitration [16]. Variations of HORN are manifested through changes in Multiple Access Protocols. In this paper, and due to page limitations, we only discuss HORN TDMA as an example multiple access protocol.

TDMA with WDMA is a very powerful means of sharing the enormous bandwidth (THz range) of optics. Time slots are assigned to each of the individual wavelengths providing a two-dimensional sharing of the THz bandwidth among multiple users. A fixed time slot is assigned to each PE where PEs send only during their preassigned time slot. It is a cyclic process where PEs wait for materialization of the time slot for transmission when previous materializations of the time slots are completed.

In order to provide for a more efficient implementation of TDMA, two time slot assignment protocols which are completely independent of one another are employed in HORN. PEs can choose to use either one or both. One time slot protocol is utilized for local communication while the other protocol is utilized for remote communication. Therefore, it is possible for a PE to send information using remote communication and send local communication at the same time if both time slots are active. This is the advantage of having two independent time slot protocols over one.

Figure 3 shows the time slot assignment protocol employed for local communication for the $H(1,1)$ ring of Figure 1. The vertical axis shows the wavelengths employed for local communication and the horizontal axis shows the time scale. The time slot assignment is constructed to allow each PE a chance to send on each of the available wavelengths for local communication in one cycle. This scheme was originally proposed by Dowd et al. [10] for the Flat Hierarchical Architecture (FHA) interconnection network. The number of communication channels required for the interconnection is independent of the number of PEs. Each PE has a chance to send on all wavelength channels in one cycle. This time slot assignment guarantees a strictly non-blocking configuration because it is inherent in the TDMA protocol.

Figure 4 shows the time slot assignment protocol employed for remote communication. The vertical axis once again lists the wavelengths available for remote communication and the horizontal axis shows the time scale. It is exactly the same protocol as employed for local communication but incorporates all PEs of HORN interconnection. N identifies the number of PEs and k identifies the number of wavelengths used for remote communication. For example, for the HORN interconnection of Figure 1, N would be equal to 234 and k would equal 22. Each PE, once again, has a chance to send on all wavelengths in one cycle, hence a strictly non-blocking configuration. Both Figures 3 and 4 show one time cycle with other cycles occurring in a round-robin cyclic process as time slots progress.

3 Scalability issues of HORN

In this section we discuss size scalability, cost scalability, and optical scalability issues for the HORN interconnection architecture. Size scalability refers to the property that the size of the network (e.g., the number of PEs) can be increased with minor or no changes made to the existing configuration. Also, the increase in system size is expected to result in a proportional increase in performance. Cost scalability is measured in terms of hardware required. For a system to be considered cost-scalable, the number of physical hardware components should not grow faster than $O(N^2)$. Optical scalability refers to Power Loss and Dynamic Range calculations of a single ring. For the HORN interconnection, power loss and dynamic range calculations are simplified since packets are regenerated as they progress up and down the hierarchies. These calculations for HORN, therefore, degenerate to those of a single ring. Complexity scalability refers to cases when performance does not keep up with complexity of the interconnection as the number of processing nodes increase. These issues are detailed next.

3.1 Size Scalability

The overall HORN interconnection structure resulting from the composition of groups $H(n,g)$ can be enlarged modularly to construct $H(n+1,g)$ groups according to the following two approaches. The first approach is to make $H(n,g)$ a hierarchy of level n, with $H(n+1,g)$ becoming a new structure created by adding to $H(n,g)$ the number of groups needed to build the next level of the hierarchy. This construction enlarges the hierarchical interconnection in a uniform and regular manner with no changes made to existing $H(n,g)$ groups. When new hierarchies are formed, AOTFs have to be reconfigured to reflect changes in the HORN topology. The second approach, to be used when incremental growth is desired, is to add new PEs to existing $H(1)$ groups. The maximum number of PEs that can exist in an $H(1)$ group is limited to the number of wavelengths (α) available for maintaining local communication. Reconfiguration of the AOTFs, however, is not required for incremental growth since the HORN topology in the macro level does not change. For both expansions, neither the number of links nor the node degree of existing PEs or switching nodes changes.

The increase in size is reflected by a proportional increase in communication channels. The number of communication channels (Λ) available for local communication is equal to:

$$\Lambda = N \tag{2}$$

This is a direct consequence of the requirement of assigning a unique wavelength to all PEs in the $H(1)$ rings as shown in Figures 2. Figure 2 also shows the number of communication channels available for local communications for the $H(1,11)$ ring. A local communication channel (i.e. a unique λ) is available to all PEs at all times as shown in Figure 3. It can be seen from the figure that an increase in the number of PEs in HORN, results in a proportional increase in communication channels.

3.2 Cost Scalability

For a system to be considered cost-scalable, the cost should be less than $O(N^2)$. By this measure, a full crossbar is not considered cost-scalable. Beyond that, a system may scale better or worse than another regarding cost. This section evaluates the cost complexity of HORN which is related to the number of transmitters, receivers, AOTFs, passive couplers, Taps, and fiber links required as the number of PEs increase. Table 1 shows a listing of these physical components where N specifies the number of PEs and $|H(i,j)|$ specifies the total number of nodes whether they are PEs or switching nodes of the j'th group at hierarchy i . The first three rows describe the overall PE complexity while the remaining rows represent the switching node complexity.

A mathematical expression was derived by Dowd et al [10] that can be used in calculating the number of nodes at a given hierarchy of a given hierarchical interconnection. This expression can be used in calculating the number of switching nodes (S) in HORN:

$$S = N\{\frac{1}{|H(1,1)|} + \frac{1}{|H(1,1)||H(2,1)|} + \ldots + \frac{1}{N}\} \tag{3}$$

The number of switching nodes is necessarily lower than $\frac{2N}{|H(1,1)|}$, and the number of fiber links is necessarily lower than $[4N + 8\frac{2N}{|H(1,1)|}]$ for an arbitrarily large number of hierarchical levels. The 4 and 8 correspond to the node degree of the processing and switching nodes, respectively. By looking at Table 1 and looking at the limit expressions for the switching nodes and fiber links, we were able to conclude that HORN can be classified as being of $O(N)$ in terms of cost complexity. This is due to the simple node interfaces and inter-ring connections that hierarchical rings provide. Cost expressions for HORN are rather small when compared to other conventional networks [13].

3.3 Optical Scalability of HORN

Two important parameters for optical scalability are power loss and dynamic range of the received signals. The above two parameters can limit the size of an interconnection network [17]. In HORN, however, both calculations degenerate to those of a single ring as discussed in section 3. Power loss in HORN is defined as any losses associated with a ring such as coupling losses from the fiber to a node, coupling losses from a node to a fiber, connector insertion losses, and fiber attenuation losses. Dynamic Range for a receiver is defined as the maximum received power to the minimum received power [18]. A receiver in HORN receives various signals from different processing nodes on the same network, which is strictly dependent on the location of the source and destination nodes. This is important since receivers only receive signals with a narrow dynamic range. More quantitative discussions on power loss and dynamic range calculations for HORN are given below.

Let us assume for a ring in HORN that the power coupling from the bus to a node is x $(0 < x < 1)$ and α

is the coupling loss of the Tap. Coupling losses in the Tap for N nodes can be defined as:

$$10\log_{10}(\frac{P_{in}}{P_{out}}) = \alpha N \qquad (4)$$

where P_{in} is the power transmitted by the source PE and P_{out} is the power received by the destination PE. Solving for $\frac{P_{out}}{P_{in}}$ yields $10^{-\alpha N/10}$. Therefore, the ratio of the output power to the input power assuming two nodes to a ring equals $(1-x)10^{-\alpha/10}$ where the extra $1-x$ accounts for the coupling loss from the bus to a node. For N nodes the ratio of the output power to the input power is equal to:

$$\begin{aligned}\eta_{ring} &= x^2(1-x)^{N-2}10^{-\alpha N/10} \\ &= 10(N-2)\log_{10}(1-x) + 20\log_{10}x - \alpha N(dB) \qquad (5)\end{aligned}$$

If we take the derivative with respect to x and maximize η_{ring} we get:

$$x_{optimum} \approx \frac{2}{N} \qquad (6)$$

Combining equations (5) and (6) yields:

$$\eta_{ring,optimum} \approx \frac{1}{e^2}(\frac{2}{N})^2 10^{-\alpha N/10} = -[2.6 + 6log_2 N + \alpha N]dB \qquad (7)$$

Power budget for HORN can subsequently be calculated using $\eta_{ring,optimum}$. The power budget for HORN must ensure that enough power will reach the receiver for reliable performance during the entire system lifetime. Power budget must incorporate all losses in the system. Therefore, power budget in HORN is related to $\eta_{ring,optimum}$ by the following equation:

$$\begin{aligned}PowerBudget &= P_{tx} - P_{min} > [2.6 + 6log_2 N + \alpha N] \\ &= \eta_{ring,optimum} \qquad (8)\end{aligned}$$

where P_{tx} specifies the transmission power and P_{min} specifies the minimum required receiving power in dBs. Figure 5 shows a graph of minimum power budget values (in dB) for different values of N assuming α is equal to 1 dB.

In HORN, the maximum received signal occurs when the source and destination PEs are located counterclockwise of each other, respectively. The maximum received power is:

$$P_{max} = P_0 x^2 10^{-2\alpha/10} \qquad (9)$$

where P_0 is the transmission power If, on the other hand, the source and destination PEs are located clockwise of each other the minimum received power is:

$$P_{min} = P_0 x^2 (1-x)^{N-2} 10^{-N\alpha/10} \qquad (10)$$

The dynamic range is then P_{max}/P_{min}:

$$DR = \frac{10^{(N-2)\alpha/10}}{(1-x)^{N-2}} = (N-2)[-10\log_{10}(1-x) + \alpha]dB \qquad (11)$$

Figure 6 shows plots for different numbers of processing nodes in a ring at different values of x.

To demonstrate the feasibility of the HORN implementation, we have calculated the total power losses for our proposed system shown in Figure 1. We will be assuming a 1 dB loss occurs from the insertion of the laser signal into the fiber and a 1 dB detector loss. Furthermore, the fiber is assumed to be at most 1 m in length at a mean operating wavelength of 960 nm. At this wavelength the fiber has an attenuation of 3.5 dB/km. Thus the fiber loss for the system is 0.0035 dB. Assuming that there are 16 PEs in a ring further equates to a power loss of about 43 dB (Figure 5). Therefore, the total power losses from the input laser diode to the output photo diode is calculated to be approximately 45 dB. For a Laser Diode, from NEC (NDL7513P1), [19] operating at 110 mW and a InGaAs Photo diode, also from NEC (NDL5461P/P1), that can receive at 10 μW, a power budget of about 80 dB is attained. Consequently, this is well within our power loss requirements for even 16 nodes to a ring.

4 Optical Implementation of HORN

Figure 7(a) and (b) shows a block diagram of the composition of the processing and switching nodes of HORN, respectively. The processing node, shown in figure 7(a), consists of an EDFA (ErbiumDoped Fiber Amplifiers), Tap (passive coupler), receiver, transmitter, and a passive coupler. EDFA's, shown in dotted lines in the figure, are optional components for the processing nodes. EDFAs are only required if power budget calculations of a ring are not met as discussed in section 3.3. The Tap is used to splice the signal from the ring to the receiver where it can be detected depending on the wavelength of the signal and the wavelength that the receiver is tuned to. The use of EDFAs and Taps for a ring was originally proposed in [20]. The switching node shown in figure 7(b) consists of a passive coupler, EDFA, and an AOTF. AOTFs are optical switches that can route on individual wavelengths.

The top ring in figure 7(a) is used for local communication and the bottom ring is used for remote communication. The R_xs shown are an array of fixed tuned receivers used for receiving packets from both local and remote transmissions. Consequently, by setting up the receivers in this manner, HORN can simultaneously receive from both local and remote communications. The transmitter (tx) is connected to a passive coupler that is able to switch between the local and remote ring depending on the communication required.

The AOTF shown in figure 7(b) is used to route on individual wavelengths. A packet is either routed to the hierarchy $i+1$ ring or the hierarchy i ring. The AOTF is an electronically tuned optical filter that operates on the principle of acousto-optic diffraction. One salient reason for using an AOTF is its electronic control where no optical processing is required. All AOTFs are initially configured, and no further reconfigurations are required. Other features that make AOTFs ideal for interconnection networks are its electronic tuning with a fast scan rate and a wide tuning range without secondary passbands allowing it to route on multiple wavelengths [14, 15]. Moreover, the combined capabilities of a wide tuning range and relatively large throughput of acousto-optic tunable filters make them favorable for the HORN interconnection.

Due page limitations, a more detail description of the optical implementation will be given at the conference.

5 Conclusion

We have put forward in this paper an optical hierarchal interconnection topology, HORN, that is scalable and has a diameter of one. Efficient implementation of routing is done for both local and remote communications by virtue of wavelength reusability. Furthermore, fault tolerance and non-blocking routing are other characteristics that have been shown for HORN. We have conducted a detailed scalability analysis for the HORN interconnection to show the viability of using it for multiprocessors and massively parallel systems.

Finally, we have presented an optical design methodology for the proposed network and showed that the architecture is highly amenable to optical implementation. The physical components required are tunable transmitters, fixed tuned receivers, EDFAs, and passive couplers. We have shown the feasibility of these components as it relates to the HORN interconnection. Consequently, simple and cost-efficient optical implementation of the proposed network with existing optical hardware is possible.

Acknowledgments
This research is supported by an NSF grant No. MIP-9310082.

References

[1] J. W. Goodman, F. J. Leonberger, S. Y. Kung, and R. A. Athale, "Optical interconnections for VLSI systems," *Proc. IEEE*, vol. 72, pp. 850–866, July 1984.

[2] S. P. Dandamudi and D. Eager, "Hierarchical Interconnection Networks for Multicomputer Systems," *IEEE Transactions on Computers*, vol. 391, pp. 786–797, 1990.

[3] M. Holliday and M. Stumm, "Performance evaluation of hierarchical ring based shared memory multiprocessors," *IEEE Transactions on Computers*, vol. 43, pp. 52–67, January 1989.

[4] Z. G. Vranesic, M. Stumm, D. M. Lewis, and R. White, "Hector: A hierarchically structured shared-memory multiprocessor," *IEEE Computer*, pp. 72–79, January 1991.

[5] F. Kiamilev, P. Marchand, A. Krishnamoorthy, S. Esener, and S. Lee, "Performance comparison between optoelectronic and VLSI multistage interconnection networks," *Journal of Lightwave Technology*, vol. 9, pp. 1665–1674, December 1991.

[6] A. Louri and H. Sung, "An Optical Multi-Mesh Hypercube: A Scalable Optical Interconnection Network for Massively Parallel Computing," *Journal of Lightwave Technology*, vol. 12, pp. 704–716, April 1994.

[7] M. Murdocca, *A Digital Design Methodology for Optical Computing*. Cambridge: MIT Press, 1990.

[8] A. E. Willner, C. H. Chang-Hasnain, and J. E. Leight, "2-D WDM Optical Interconnections Using Multiple-Wavelength VCSELs for Simultaneous and Reconfigurable Communication Among Many Planes," *IEEE Photonics Technology Letters*, vol. 5, pp. 838–841, 1993.

[9] M. I. Irshid and M. Kavehrad, "A Fully Transparent Fiber-Optic Ring Architecture for WDM Networks," *Journal of Lightwave Technology*, vol. 10, pp. 6425–6437, Nov. 1993.

[10] P. Dowd, K. Bogineni, K. Aly, and J. A. Perreult, "Hierarchical Scalable Photonic Architectures for High-Performance Processor Interconnection," *IEEE Transactions on Computers*, vol. 42, pp. 1105–1120, September 1993.

[11] L. G. Kazovsky and P. T. Poggiolini, "A Multigigabit-per-second Optical LAN Utilizing a Passive WDM Star," *Journal of Lightwave Technology*, vol. 11, pp. 1009–1026, June 1993.

[12] J. Bannister, M. Gerla, and M. Kovacevie, "An All-Optical Multifiber Tree network," *Journal of Lightwave Technology*, vol. 11, pp. 997–1008, June 1993.

[13] D. A. Reed and H. D. Schwetman, "Cost-Performance Bounds for Multicomputer Networks," *IEEE Transactions on Computers*, vol. 32, pp. 83–95, Jan. 1983.

[14] R. J. Berinato, "Acousto-optic Tapped delay-line filter," *Applied Optics*, vol. 32, pp. 5797–5809, Oct. 1993.

[15] R. B. Jenkins and B. D. Clymer, "Acousto-optic comparison switch for Optical Switching Networks with analog addressing techniques," *Applied Optics*, vol. 31, pp. 5453–5463, Sept. 1992.

[16] G. Lundy, "Analyzing a CSMA/CD protocol through a systems of communicating machines specification," *IEEE Transactions on Communications*, vol. 41, pp. 447–450, March 1993.

[17] B. E. A. Saleh and M. C. Teich, *Fundamentals of Photonics*. New York, NY: A Wiley-Interscience Publication, 1991.

[18] G. P. Agrawal, *Fiber-Optic Communication Systems*. New York, NY: A Wiley-Interscience Publication, 1992.

[19] NEC, "Electron Components: Optical Semiconductor Devices," *A Data Sheet Pack for Optical Semiconductor Devices*, pp. 1–200, May 1995.

[20] A. E. Willner and S.-M. Hwang, "Transmission of Many WDM Channels Through a Cascade of EDFA's in Long-Distance Links and Ring Networks," *Journal of Lightwave Technology*, vol. 13, pp. 802–816, May 1995.

	HORN
Transmitters (tunable)	N
Fixed Tuned Receivers	$N * h + 1$
Tap	$2N$
AOTF	$\sum_{i=2}^{h} \lvert H(i,j) \rvert$
Fiber Links	$2\sum_{i=1}^{h} \lvert H(i,j) \rvert$
Passive Couplers	$\sum_{i=1}^{h} \lvert H(i,j) \rvert$

Table 1: Cost scalability expressions for HORN where N is the number of PEs,h identifies the number of hierarchies, and $|H(i,j)|$ refers to the number of PEs or switching nodes at the $j'th$ group at hierarchy i.

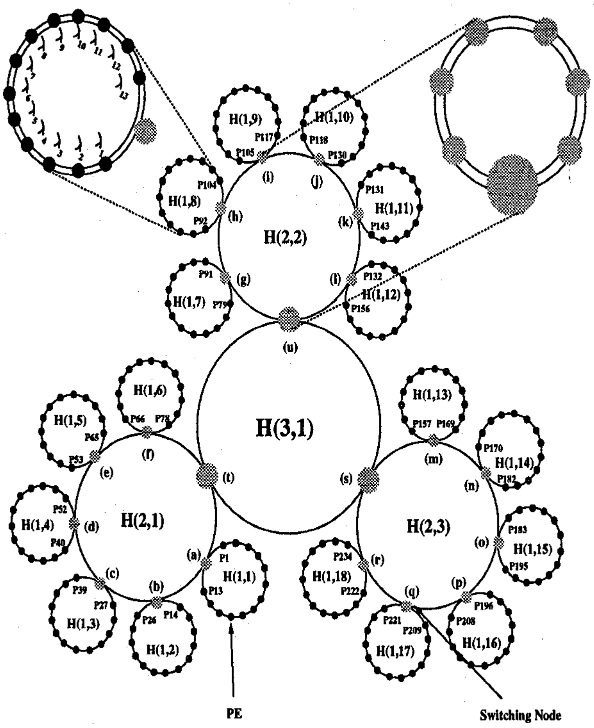

Figure 1: Diagram of a $HORN(4,3)$ where PEs are indicated by black filled circles and switching nodes are indicated by gray filled circles. All hierarchical groups are labeled using $H(n,g)$ where n identifies the hierarchy and g identifies a unique group at hierarchy n.

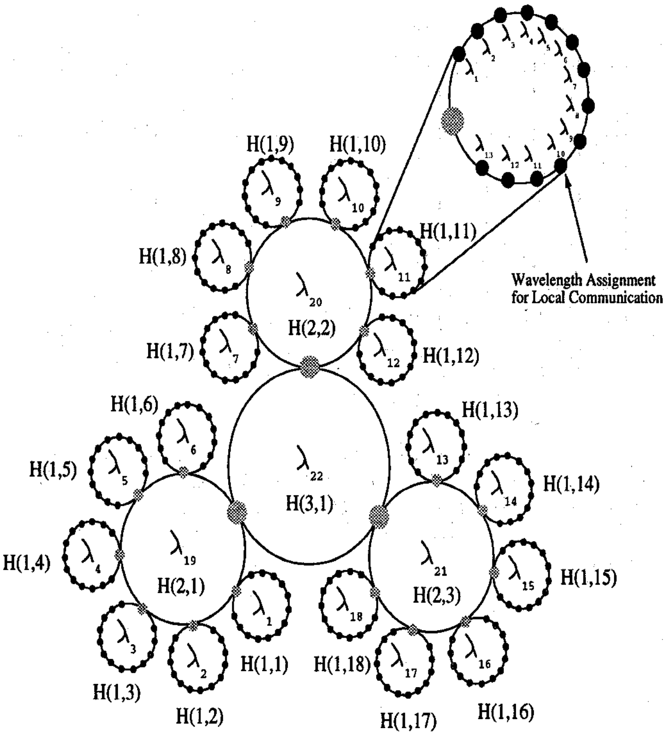

Figure 2: A wavelength assignment where there are 22 wavelengths available for remote communications. All rings of all hierarchies are assigned a unique wavelength. The local wavelength assignment for $H(1,11)$ ring is shown with all other $H(1)$ rings assigned in the same manner.

λ_1	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9	P_{10}	P_{11}	P_{12}	P_{13}
λ_2	P_{13}	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9	P_{10}	P_{11}	P_{12}
λ_3	P_{12}	P_{13}	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9	P_{10}	P_{11}
λ_4	P_{11}	P_{12}	P_{13}	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9	P_{10}
λ_5	P_{10}	P_{11}	P_{12}	P_{13}	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9
λ_6	P_9	P_{10}	P_{11}	P_{12}	P_{13}	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8
λ_7	P_8	P_9	P_{10}	P_{11}	P_{12}	P_{13}	P_1	P_2	P_3	P_4	P_5	P_6	P_7
λ_8	P_7	P_8	P_9	P_{10}	P_{11}	P_{12}	P_{13}	P_1	P_2	P_3	P_4	P_5	P_6
λ_9	P_6	P_7	P_8	P_9	P_{10}	P_{11}	P_{12}	P_{13}	P_1	P_2	P_3	P_4	P_5
λ_{10}	P_5	P_6	P_7	P_8	P_9	P_{10}	P_{11}	P_{12}	P_{13}	P_1	P_2	P_3	P_4
λ_{11}	P_4	P_5	P_6	P_7	P_8	P_9	P_{10}	P_{11}	P_{12}	P_{13}	P_1	P_2	P_3
λ_{12}	P_3	P_4	P_5	P_6	P_7	P_8	P_9	P_{10}	P_{11}	P_{12}	P_{13}	P_1	P_2
λ_{13}	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9	P_{10}	P_{11}	P_{12}	P_{13}	P_1
	TS 1	TS 2	TS 3	TS 4	TS 5	TS 6	TS 7	TS 8	TS 9	TS 10	TS 11	TS 12	TS 13

One Time Cycle

Figure 3: Time slot allocation for local communication. The vertical axis shows the wavelengths and the horizontal axis shows the time slots. One time cycle is shown. In one cycle each PE can send on all wavelengths.

λ_1	P_1	P_2	P_3	P_4	P_5		P_{N-4}	P_{N-3}	P_{N-2}	P_{N-1}	P_N
λ_2	P_2	P_3	P_4	P_5	P_6		P_{N-3}	P_{N-2}	P_{N-1}	P_N	P_1
λ_3	P_3	P_4	P_5	P_6	P_7	• • •	P_{N-2}	P_{N-1}	P_N	P_1	P_2
λ_4	P_4	P_5	P_6	P_7	P_8		P_{N-1}	P_N	P_1	P_2	P_3
λ_5	P_5	P_6	P_7	P_8	P_9		P_N	P_1	P_2	P_3	P_4
			⋮						⋮		
λ_{k-4}	P_{k-4}	P_{k-3}	P_{k-2}	P_{k-1}	P_k		P_{k-9}	P_{k-8}	P_{k-7}	P_{k-6}	P_{k-5}
λ_{k-3}	P_{k-3}	P_{k-2}	P_{k-1}	P_k	P_{k+1}		P_{k-8}	P_{k-7}	P_{k-6}	P_{k-5}	P_{k-4}
λ_{k-2}	P_{k-2}	P_{k-1}	P_k	P_{k+1}	P_{k+2}	• • •	P_{k-7}	P_{k-6}	P_{k-5}	P_{k-4}	P_{k-3}
λ_{k-1}	P_{k-1}	P_k	P_{k+1}	P_{k+2}	P_{k+3}		P_{k-6}	P_{k-5}	P_{k-4}	P_{k-3}	P_{k-2}
λ_k	P_k	P_{k+1}	P_{k+2}	P_{k+3}	P_{k+4}		P_{k-5}	P_{k-4}	P_{k-3}	P_{k-2}	P_{k-1}

One Time Cycle

Figure 4: Time slot allocation for remote communication. The vertical axis shows the wavelengths and the horizontal axis shows the time slots. One time cycle is shown. In one cycle each PE can send on all wavelengths.

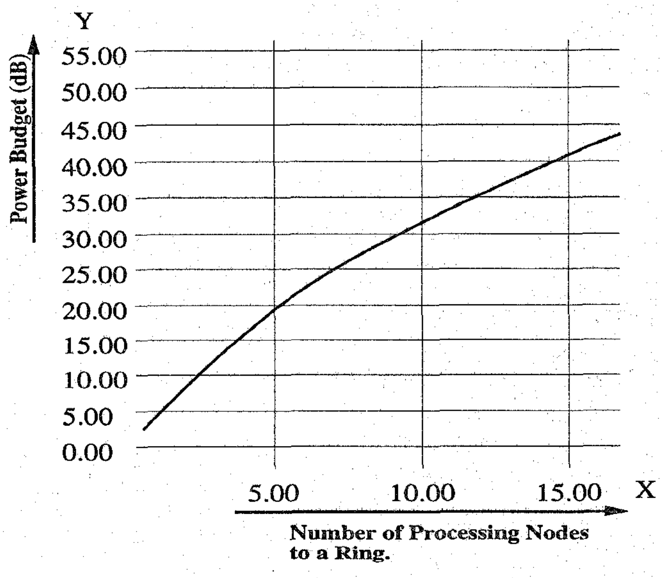

Figure 5: Shows a graph of the minimum Power Budget required for different numbers of processing nodes in a ring.

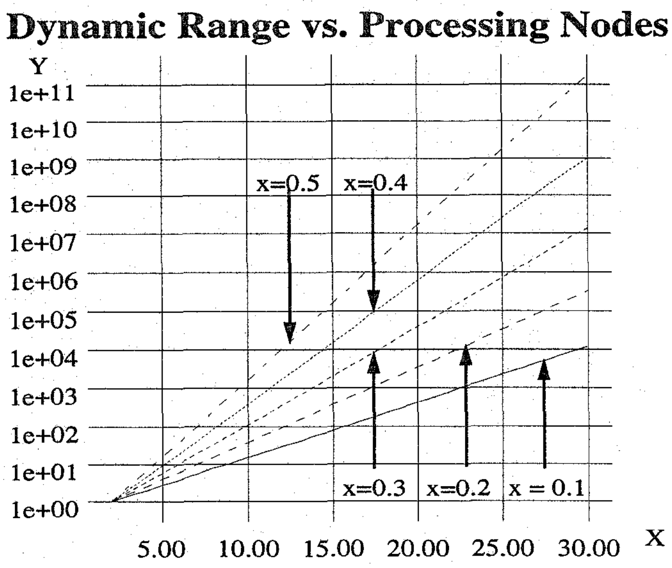

Figure 6: Shows a graph of Dynamic Range vs Processing Nodes for different values of power coupling loss from the bus to a node (x).

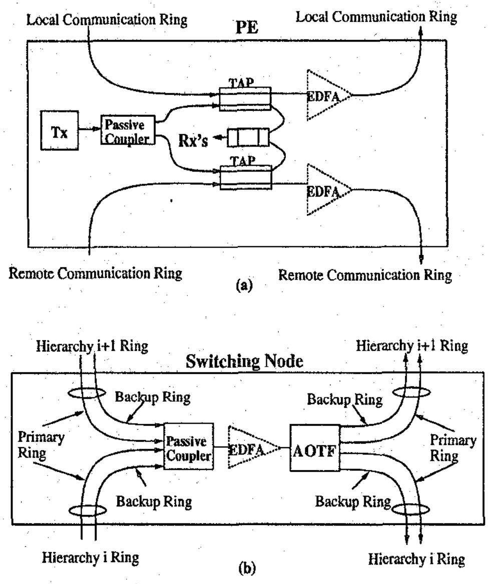

Figure 7: Electro/optical components of HORN. (a) shows the composition of a PE. Rx's represents an array of fixed tuned receivers and Tx represents a single tunable transmitter. EDFAs are shown in dotted lines since they only need to be used if power budget calculations of a single ring are not met. (b) shows the composition of a switching node.