Optimization of Multiple-Stage Optical Interconnection Networks

Benjamin A. Small, *Student Member, IEEE,* and Keren Bergman, *Member, IEEE*

***Abstract*—An analysis of the cost and complexity of multiple-stage optical interconnection networks is presented. Because of the high cost of photonic switching elements from which these networks are composed, the radix (degree of the switching nodes) of a simple banyan network is important in determining total network cost. For most switching node cost models, it is shown that 3 × 3 switching nodes can actually be more efficient than 2 × 2 ones.**

***Index Terms*—Banyan networks, multistage interconnection networks (MINs), photonic switching systems.**

I. Introduction

NUMEROUS applications for interconnection networks with extremely high throughput and low latency have been presented, including high-performance computing and telecommunications core routing [1]–[4]. Optical interconnection networks (OINs) are able to meet many of these system requirements by leveraging the large bandwidth afforded by fiber-optic and photonic technologies [2]–[4]. Multiple-stage (or multistage) interconnection networks (MINs) offer topological advantages such as scalability and modularity owing to their composition from hundreds or thousands of similar switching node building blocks [1], [5], as shown in Fig. 1.

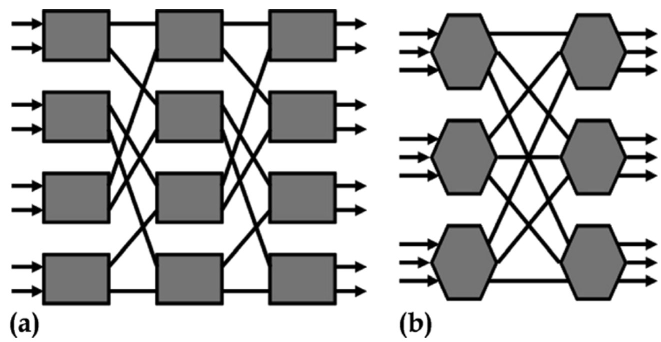

Fig. 1. Schematic of two MIN topologies: (a) binary 8 × 8 and (b) 3-ary 9 × 9 butterfly configurations with pathways according to the omega topology.

In OIN systems, most of the complexity and cost is attributable to the photonic switching elements. When multiple-wavelength data is employed, these switching elements must execute space switching in addition to wavelength switching. Many switching node designs based on this technique have been proposed, and most are implemented with semiconductor optical amplifier (SOA) switching gates [2]–[4], [6]–[12].

Previous work has analyzed various performance metrics and optimization techniques for MINs [1], [5], [13]–[16] but few of these results are specific to the particulars of photonic implementation. In this work, we address the salient issues of OIN designs that specifically incorporate photonic switching element cost and complexity. In designing multiple-stage OINs, it is important to account for these cost metrics while optimizing network performance metrics, like throughput, latency, and utilization [16].

II. Switching Nodes

SOA-based OIN switching nodes are appealing because of the large optical bandwidth supported by SOAs, and because the SOA gain can be adjusted so that it compensates for the losses incurred by couplers and other passive optical components within the switching node. Designs for a single-packet 2 × 2 self-routing switching node [11] and a nonblocking 2 × 2 self-routing switching node [12] have been implemented and investigated. The datapaths of these OIN switching nodes are comprised of SOA switching gates and fiberoptic couplers and splitters; filters, optoelectronic components, and basic electronics enable the self-routing functionality. The technological framework utilized for these prototypes can be extended to higher degree OIN switching node designs (Fig. 2).

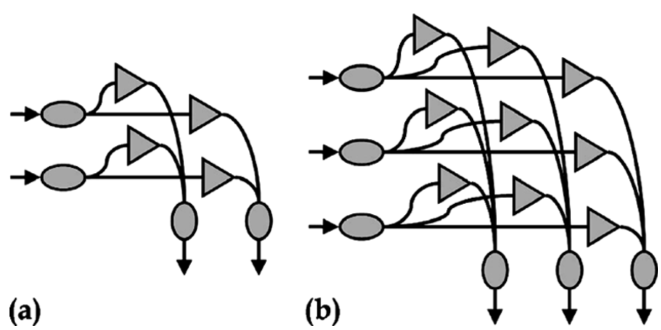

Fig. 2. Schematic of the optical data pathways of nonblocking 2 × 2 and 3 × 3 switching nodes, with SOAs as triangles and optical couplers as ovals.

The number of SOAs, passive optical splitters, and couplers depends on the degree of the switching node (i.e., the number of input and output ports). The design is fundamentally just a crossbar switch preceded by a fan-out splitter and followed by a fan-in coupler. Thus, the number of switching gates or elements scales as the square of the switching node degree: k^2 for k input and output ports. The cost and complexity of other components, such as the splitter and coupler, generally scale linearly with switching node degree k. Therefore, a cost function can describe the relative complexity of switching nodes of differing degree

$$c(k) = k^2 + \alpha k \tag{1}$$

Manuscript received September 15, 2005; revised October 23, 2005. This work was supported by the National Science Foundation under Grant ECS-0532762.

The authors are with the Department of Electrical Engineering, Columbia University, New York, NY 10027 USA (e-mail: bas@ee.columbia.edu).

Digital Object Identifier 10.1109/LPT.2005.862016

where α is a technology-dependent scaling parameter for the linear cost component. This cost function is arbitrarily normalized; only relative cost scaling is relevant.

Based on current technological trends as of 2005, SOAs are almost ten times more expensive than most passive optical components. For the technology generation based on discrete SOA chips and fused silica couplers, $\alpha \sim 0.3$ [17]. These trends are likely to continue even with further device integration [2], [8]–[10].

It is also important to note that the framework of the switching node cost function (1) is independent of the switching device technology. This form applies to any switching node composed of simple ON–OFF switching elements, as is the case with many photonic nonblocking node designs [3]–[9]. In advanced technologies, it may be possible to increase α further, i.e., to decrease the cost of the switching node components which scale quadratically.

Although we have emphasized the monetary cost of the switching node components, it is possible to represent other kind of costs or complexity metrics by the form (1). For example, in single-substrate integrated optics technology, the die area can become a valuable commodity, and switching node footprint scales quadratically with degree, as in [2], [8]–[10]. Electrical power consumption is yet another important cost metric which scales quadratically with switching node degree when active optoelectronic devices are used as switching elements [2].

III. Networks

Having established an analytic expression for the cost or complexity of a single switching node, we present a framework for evaluating whole multiple-stage OINs (as in Fig. 1) and their total cost as a function of the number of network terminals (input or output ports) and of the degree of the constituent switching nodes.

Consider that cost or complexity of an OIN can be quantified as the product of the number of switching nodes and the cost or complexity of each of the individual and identical switching nodes. Mathematically, these two quantities are a function of network size and radix

$$C(T,k) = N(T,k)c(k) \tag{2}$$

where T is the number of terminal nodes in the network (as for a $T \times T$ network), k is the radix of the network, and hence, the degree of the switching nodes, and the quantities N and c represent the number of switching nodes and the cost of an individual node, respectively.

For a banyan network, which is the general class containing the butterfly and omega (multiple-stage shuffle-exchange) topologies, neither of which are nonblocking, the switching nodes are arranged in a rectangular grid containing $\log_k T$ columns of height T/k, yielding

$$N(T,k) = \frac{T}{k}\log_k T. \tag{3}$$

Thus, the total cost function for a k-ary banyan network is

$$\begin{aligned} C(T,k) &= \frac{T}{k}\log_k T \cdot (k^2 + \alpha k) \\ &= T\log_k T \cdot (k+\alpha) \end{aligned} \tag{4}$$

where α represents the switching node's linear cost component, as discussed above for (1). The family of curves described by (4) are all upwardly concave with global minima within an order of magnitude of α (Fig. 3).

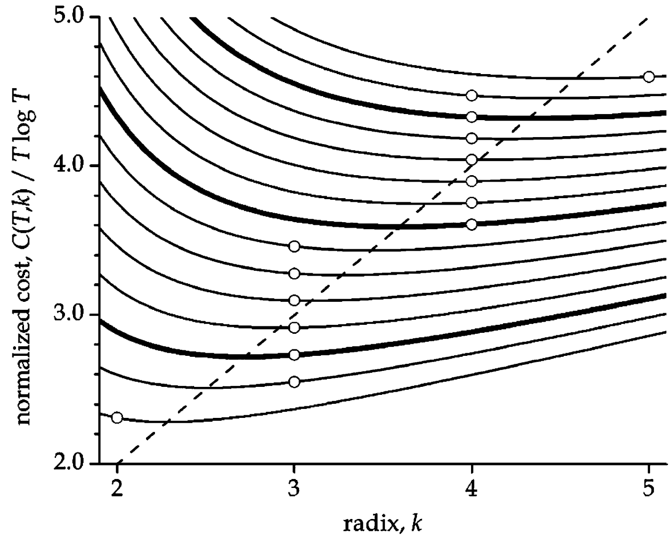

Fig. 3. Contours of normalized total cost function for values of α between -0.4 and 2.4, with heavy curves for $\alpha = \{0.0, 1, 0, 2.0\}$. The locus of global minima (- -) and the minimum integer values of $k(\bigcirc)$ are also shown.

Minimizing the total cost function requires solving for k when the derivative is set to zero

$$\frac{\partial}{\partial k}C(T,k)|_T = 0. \tag{5}$$

Thus, for the banyan network with total cost described by (4), the optimal value of k is

$$k^* = \frac{\alpha}{W\left(\frac{\alpha}{e}\right)} \tag{6}$$

where $W(\cdot)$ is Lambert's transcendental function (*product-log* or *omega function*) [18] and e is Euler's constant (i.e., the base of the natural logarithm).

Before proceeding further with this analysis, it is important to recall that the parameters T, k, and N, must all be natural numbers (positive integers). However, because T and N are often assumed to be many orders of magnitude greater than k, it can be assumed without much loss of precision that T and N are continuous variables and that only k is discrete.

First consider the case where $\alpha \to 0$, for which (6) gives $k^* = e$. Knowing that a switching node with $e \approx 2.7$ input ports and e output ports is untenable, consider instead which of $k = 2, 3$ is preferable

$$\begin{aligned} C(T,2) &= \frac{1}{\log 2}T\log T \cdot (\alpha+2) \\ C(T,3) &= \frac{\log 2}{\log 3}\frac{(\alpha+3)}{(\alpha+2)}C(T,2). \end{aligned} \tag{7}$$

Thus, for all $\alpha > \log(8/9)/\log(3/2) \approx -0.3$, the total cost of a switch composed of 3×3 switching nodes is less than that of 2×2 nodes. Knowing that for any real switching technology $\alpha > 0$, it is clear that an OIN of degree $k \geq 3$ is optimal. In fact, by an argument similar to (7), for $\alpha > \log(81/64)/\log(4/3) \approx 0.8$, 4×4 switching nodes are preferred; only when $\alpha > \log(1024/625)/\log(5/4) \approx 2.2$ do 5×5 switching nodes become the optimal solution (Fig. 4.).

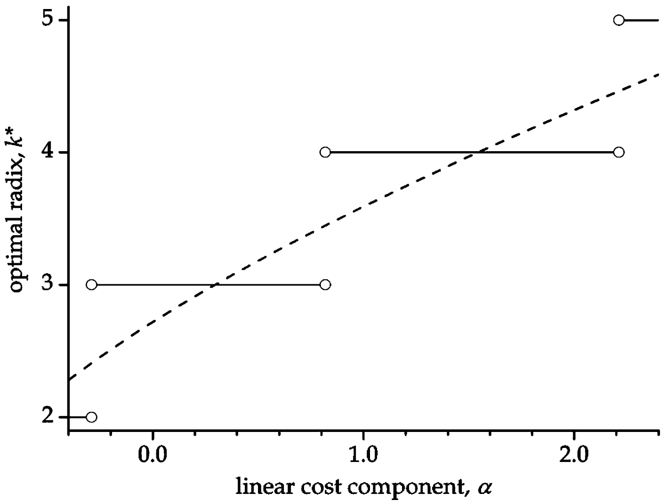

Fig. 4. Plot of the optimal value of the network radix and node degree k, solved continuously (- -) based on (6) and discretely (◯).

As an example, consider a switch with 2048 input and output terminals. With 2 × 2 switching nodes, $\log_2 2048 = 11$ stages must contain $2048/2 = 1024$ nodes each, yielding a total of 11 264 switching nodes. Each 2 × 2 SOA-based OIN switching node discussed above contains four SOA switching gates, so this 2048 × 2048 binary banyan OIN requires 45 056 SOAs, which matches (4). If, however, 3 × 3 switching nodes are used, seven stages (rounding up from $\log_3 2048 \approx 6.9$) of 683 (from $2048/3 \approx 682.7$) nodes each implies a total of 4781 switching nodes. With nine SOAs per node, the 2049 × 2049 3-ary banyan OIN has a total of 43 029 SOA switching gates. Thus, the OIN based on 3 × 3 nodes requires 2029 or 4.7% fewer SOAs than the 2 × 2 case while providing an additional terminal. The cost benefits of the 3 × 3-based topology are even more pronounced for $\alpha \gg 0$.

It should also be noted that, primarily because the number of stages decreases for larger node radices, the blocking probability also decreases [1], so this and other performance metrics should be considered along with cost during network design.

We have chosen to focus on the blocking banyan topologies for this analysis because it offers interesting and practical methods of cost optimization. Other classes of topologies, such as torus (cube) structures and Clos networks, offer no inherent cost advantages to varying the radix of the component switching nodes. Varying the dimensionality of tori (hypercubes) has no cost effect except for the number of pathways [1], [5] which are relatively inexpensive in OINs. On the other hand, the base of a Clos network is set to optimize performance and has no interesting implications on total network cost [1], [5]. Although all simple banyan topologies like the ones discussed here are not nonblocking, they are nevertheless common in a variety of computing applications due to their other advantages, including logical simplicity and minimal footprint [1].

IV. Conclusion

We have presented a method of analyzing the cost of multiple-stage OINs. The total cost of a banyan network composed of photonic switching nodes is then analyzed algebraically. An optimal configuration is determined from a model of the cost of individual switching nodes, which emphasizes the photonic switching elements, independent of network performance metrics. It is then shown that the total cost for this class of networks can be minimized by utilizing 3 × 3 and higher-degree switching nodes instead of the conventional binary 2 × 2 ones.

References

[1] W. J. Dally and B. Towles, *Principles and Practices of Interconnection Networks*. San Francisco, CA: Morgan Kaufmann, 2004.

[2] G. Zhou, Y. Zhang, and W. Liu, "Optical fiber interconnection for the scalable parallel computing system," *Proc. IEEE*, vol. 88, no. 6, pp. 856–863, Jun. 2000.

[3] R. S. Tucker and W. D. Zhong, "Photonic packet switching: an overview," *IEICE Trans. Electron.*, vol. E82-C, pp. 202–212, Feb. 1999.

[4] G. I. Papadimitriou, C. Papazoglou, and A. S. Pomportsis, "Optical switching: switch fabrics, techniques, and architectures," *J. Lightw. Technol.*, vol. 21, no. 2, pp. 384–405, Feb. 2003.

[5] A. Pattavina, *Switching Theory*. New York: Wiley, 1998.

[6] J. D. Evankow and R. A. Thompson, "Photonic switching modules designed with laser diode amplifiers," *IEEE J. Sel. Areas Commun.*, vol. 6, no. 7, pp. 1087–1095, Aug. 1988.

[7] M. S. Borella, J. P. Jue, D. Banerjee, B. Ramamurthy, and B. Mukherjee, "Optical components for WDM lightwave networks," *Proc. IEEE*, vol. 85, no. 8, pp. 1274–1307, Aug. 1997.

[8] *Issue on Integrated Optics and Optoelectronics, Part 1, IEEE J. Sel. Topics Quantum Electron.*, vol. 11, no. 1, pp. 1–296, Jan./Feb. 2005.

[9] *Issue on Integrated Optics and Optoelectronics, Part II, IEEE J. Sel. Topics Quantum Electron.*, vol. 11, no. 2, pp. 297–564, Mar./Apr. 2005.

[10] M. A. Mentzer, *Principles of Optical Circuit Engineering*. New York: Dekker, 1990.

[11] B. A. Small, A. Shacham, and K. Bergman, "Ultra-low latency optical packet switching node," *IEEE Photon. Technol. Lett.*, vol. 17, no. 7, pp. 1564–1566, Jul. 2005.

[12] A. Shacham, B. G. Lee, and K. Bergman, "A wideband, nonblocking, 2 × 2 switching node for a SPINet network," *IEEE Photon. Technol. Lett.*, vol. 17, no. 12, pp. 2742–2744, Dec. 2005.

[13] J. Giglmayr, "Architectural framework for optical frequency division multiplexing systems based on regular space-frequency interconnections," *J. Lightw. Technol.*, vol. 12, no. 7, pp. 1291–1306, Jul. 1994.

[14] A. Agarwal, "Limits on interconnection network performance," *IEEE Trans. Parallel Distrib. Syst.*, vol. 2, no. 4, pp. 398–412, Oct. 1991.

[15] F. Kiamilev, P. Marchand, A. V. Krishnamoorthy, S. C. Esener, and S. H. Lee, "Performance comparison between optoelectronic and VLSI multistage interconnection networks," *J. Lightw. Technol.*, vol. 9, no. 12, pp. 1674–1692, Dec. 1991.

[16] C. Guan and V. W. S. Chan, "Topology design of OXC-switched WDM networks," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 8, pp. 1670–1686, Aug. 2005.

[17] "OFC/NFOEC buyer's guide," presented at the Optical Fiber Commun. Conf., Anaheim, CA, Mar. 2005.

[18] J. H. Lambert, "Observations variae in mathesin puram," *Acta Helvitica*, vol. 3, pp. 128–168, 1758.