

# FUNDAMENTAL ISSUES IN OPTICAL INTERCONNECTIONS

HALDUN M. OZAKTAS

*Bilkent University*

*Department of Electrical Engineering*

*TR-06533 Bilkent, Ankara, Turkey*

**Abstract.** We review some of the relatively fundamental work in the area of optically interconnected digital computing systems. We cover comparisons of optical interconnections with other interconnection media in terms of energy and interconnection density, studies determining the optimal combination of optical and electrical interconnections that should be used, work on free-space optical interconnection architectures, complexity studies, and work on physical and logical system architectures and algorithms. We exclude work on devices, components, materials, and manufacturing.

## 1. Introduction

A substantial amount of effort has been invested to determine whether, when, and how the use of optical interconnections may lead to improvements in digital computing systems. These efforts have been multi-faceted, ranging from the development of materials, devices, and components; to system-level studies comparing the relative strengths and weaknesses of optical interconnections; to the development of novel optical and system architectures.

In this work we attempt to provide a review of some of the relatively fundamental work in this area. We exclude fundamental work on devices, materials, and manufacturing techniques. We also limit our attention to digital systems, excluding analog systems and neural networks. We further exclude other important topics such as computer-aided design of optoelectronic or optically interconnected systems.

## 2. General considerations

The following considerations are not specific to optically interconnected systems but apply to all digital computers.

## 2.1. HEAT REMOVAL LIMITATIONS

The nonlinear elements (switches or gates) of which digital computing systems are universally made of, dissipate a certain amount of energy  $E_d$  per switching event. Furthermore, each bit of information that is sent from one point to another is represented by an amount of energy  $E$ , which is consequently dissipated. The value of  $E_d$  may very well be in the fJ range. On the other hand, the value of  $E$  will more likely be of the order of a pJ or more. Smaller values may be attainable for very short interconnections, but for systems whose overall size is of the order of 1 m, it seems fair to take  $E + E_d \approx E \sim 1$  pJ as an appropriate average for the energy per switching/transmission event.

If we let  $N$  denote the number of elements (switches or gates) in our system and  $k$  the average number of interconnections per switch, then the total power dissipated is  $\mathcal{P} = kNEB$ , where  $B$  is the bit-repetition-rate; the rate at which bits of information are being emitted into each interconnection. With  $N = 10^8$ ,  $k = 5$ ,  $B = 2$  Gb/s, we obtain  $\mathcal{P} = 1$  MW. If this system occupies a volume of  $\mathcal{L}^3 = (10 \text{ cm})^3$ , then it becomes necessary to remove  $10^8 \text{ W/m}^2 = 10 \text{ kW/cm}^2$  from the system. This seems to be the very upper limit of what can be achieved by fluid convection [1], and would probably be very difficult to achieve. Note that the amount of power we can remove from a system is proportional to the *cross-sectional area* (not the surface area) of the system [1].

Simple calculations will show that the space occupied by the elements and even the interconnections is small in comparison to that required by heat removal considerations. The moral is that heat removal will be the most important consideration in high performance future computing systems. It will be of paramount importance to employ the most aggressive heat removal methods and integrate heat removal into system design from the onset.

## 2.2. CONFINING ACTIVE DEVICES TO PLANES

For highly interconnected circuits, the following result is of relevance [2]: Provided the interconnections are allowed to be routed through three-dimensional space, there is no disadvantage in restricting the active devices to a plane. That is, the overall volume and signal delay and thus clock rate of a highly interconnected three-dimensional system in which the active devices are restricted to a plane will not be inferior to a system in which the active devices can be situated freely throughout the volume. This is because the volume is primarily determined by heat removal limitations or the space occupied by the interconnections, so that additional restrictions on the ac-

tive devices are of little or no significance. This result, which is based on purely geometrical considerations, is valid for all types of interconnections.

### 2.3. GLOBALLY INTERCONNECTED SYSTEMS

Globally interconnected systems are better than locally interconnected systems because global connections allow flexible and fast global transfer of information. Global interconnections take up much more space, but since heat removal considerations imply large separations between the elements anyway, this does not result in larger system size and delays [3].

## 3. Energy comparisons

Having concluded that heat removal limitations will become one of the most important, if not the most important consideration, it follows that reducing the energy per transmitted bit of information  $E$  is of primary importance. Various studies comparing  $E$  for isolated optical and electrical lines indicate that optics becomes more profitable for interconnections shorter than some break-even distance, estimated between 0.1 mm and 10 cm [4, 5, 6]. (These comparisons assume that the energy received is large enough to charge the receiver capacitance to the nominal logic level—area consuming amplifier circuits as used in long-distance communications would be unacceptable.)

Studies which account for a very large number of parameters exist, but the essence of the comparison can be captured by a simple argument. Since attenuation in optical media is very low (as evident from the distance information is able to travel in telecommunications fibers), the energy per transmitted bit can be assumed to be constant in optical systems. On the other hand, the energy for an unterminated (lumped RC) conducting interconnection is proportional to  $\ell^2$ , where  $\ell$  is the length of the line. Thus, beyond a certain value of  $\ell$ , optical interconnections are deemed favorable. If terminated lines are used, the energy is proportional to the length  $T$  of the shortest pulse that can propagate on the line with acceptable attenuation, which is in turn proportional to  $\ell^2/W^2$ , where  $W$  is the width of the line. The only way to prevent the energy from increasing with length is to increase  $W \propto \ell$ , but this increase in the thickness of lines and subsequent crowding will eventually result in an increase in the size of the system, increasing the line lengths  $\ell$ , which will require a further increase in  $W$ , and so on, thwarting the attempt to circumvent the intrinsic limitations associated with line resistance.

Miller has given an intriguing physical interpretation of the fact that optical interconnections are energetically more advantageous [6].

It has been shown that the energy per transmitted bit with superconducting interconnections is also constant and comparable to what may be

achieved with optical interconnections, for a nominal voltage level of 1 V. It may be further reduced if lower nominal voltages are employed at lower temperatures [2].

## 4. Interconnection density

### 4.1. COMPARISONS WITH OTHER TECHNOLOGIES

A fairly common misconception is that optical channels do not occupy any volume and are “free” in this sense. Whereas it is true that separate optical channels can cross through each other with little or no interference, optical channels nevertheless have a volume cost: The *minimum* volume occupied by an optical interconnection system with total interconnection length  $\ell_{\text{total}}$  is  $\sim \lambda^2 \ell_{\text{total}}$ , where  $\lambda$  is the wavelength of light [7]. Thus each independent spatial channel can be thought to have a constant transverse area of  $\lambda^2$ .

On the other hand, the cross-sectional area of conducting interconnections must increase with length to maintain acceptable attenuation. Thus beyond a certain length, optical interconnections become favorable. The break-even lengths cited are comparable to those obtained from energy considerations [2, 4, 8, 9].

Two-dimensional integrated optical guided wave circuits are not able to support a surface density of one spatial channel per wavelength because of the larger separations implied by crosstalk considerations. Such systems will probably not be favorable until extremely large system sizes [10].

### 4.2. TOPOLOGICAL CONSIDERATIONS FOR OPTICS

The word “minimum” appearing in the first paragraph of the preceding subsection is of significance. Indeed, it was first thought that a strong tradeoff between topological flexibility (the ability to implement arbitrary patterns of connections) and volume existed: that greater flexibility came at the cost of greater volume. However, it was later understood that these tradeoffs applied only to a broad class of architectures which may be referred to as multi-facet architectures, but that other architectures free from these limitations are possible. Thus arbitrary circuit diagrams (graphs) can be realized with an effective interconnection density of  $\sim 1/\lambda^2$ . This result assumes optical systems whose  $f$ -numbers (analogous to the  $F$ -stop on a camera) approach unity. In practice, it seems realistic to expect this density to be achieved within an order of magnitude. In most cases this will be sufficient, since heat removal and not interconnection density will be the limiting factor.

Publications investigating these issues include [11, 12, 13, 14]

### 4.3. ARCHITECTURES FOR FREE-SPACE OPTICS

A large variety of free-space optical interconnection architectures have been proposed and studied. In this section we mention works that deal with the optical interconnections only, without consideration of the overall computing system architecture [15, 16, 17, 18, 19, 20, 21]. These studies concentrate primarily on optical systems to image arrays of sources onto arrays of detectors in an efficient manner. Several authors have discussed the ultimate limitations of these systems; others have been led to investigations of mathematical and combinatoric issues regarding permutation and other types of networks. Shuffle-based systems are by far the most commonly studied, both because they have efficient optical implementations and because of their versatility.

## 5. System-level studies

### 5.1. AREA-VOLUME-TIME COMPLEXITY STUDIES

Here we are referring to area-volume-time studies either in the spirit of classical VLSI complexity theory, or the closely related pin-out/line-length distribution based studies [22]. Work on three-dimensional complexity theory includes [2, 23, 24]. Some of these works are classical complexity studies, being based on purely geometrical considerations and ignoring all constants of proportionality. Others have reintroduced physical constants and order-of-magnitude accuracy in an attempt to obtain results which are more physically relevant [2].

### 5.2. SYSTEM-LEVEL TECHNOLOGY COMPARISONS

These studies typically assume an array of elements of some degree of complexity (usually ranging from transistors to processors) and compare performance and cost measures resulting from the use of different interconnection media. These works differ from the comparison studies mentioned earlier in that they compare systems, rather than isolated interconnections [2, 25, 26].

Based solely on fundamental physical considerations, optical and superconducting interconnections are comparable to each other and superior to normally conducting interconnections, even when aggressive use of repeaters is made to circumvent capacitive and resistive problems. The possibility of three-dimensional circuits and freedom from termination problems seems to give optics an edge over superconductors.

The use of normally conducting interconnections results in severe limitations in large scale systems. Even with arbitrarily small scaling and arbitrarily fast devices, with or without termination, the maximum bit rates that can be supported sharply decrease with increasing system size. Likewise,

the bisection-inverse delay and bisection-bandwidth products are bounded from above. This is in contrast with the other media with which it is possible to arbitrarily increase  $B$  and the bisection-bandwidth product for any given  $N$ . The use of repeaters alleviates, but does not change the inferiority of normal conductors, while at the same time greatly increasing manufacturing complexity because of the need to support active devices at all levels of the interconnection hierarchy. The use of optics is comparably even more advantageous when large bandwidths are required.

### 5.3. OPTIMAL COMBINATION OF OPTICAL AND ELECTRICAL INTERCONNECTIONS

These studies discuss how optical and electrical interconnections can be used together. They essentially address the question: Beyond what level of the interconnection hierarchy should optics be used? Or, how should a large circuit be partitioned into optically interconnected electronic chips? That is, what is the optimal grain size? [10, 27, 28]

The general indication is that the system should be partitioned into clusters of  $\sim 10^4$ – $10^6$  transistors; for instance, in the form of an optically interconnected array of chips with this many transistors on each. It has been argued that the use of optical interconnections allows one to access points in performance space unachievable otherwise, however possibly at a significant cost of space and/or power [10]. The increase in performance may take the form of smaller signal delay and/or larger bandwidth in comparison to an all-electrical system. If minimization of the cost associated with size and power is the primary objective, rather than the maximization of performance, then the use of optics does not become beneficial until extremely large system sizes.

### 5.4. MULTI-FACETED SYSTEM-LEVEL STUDIES

Here we mention some studies which deal with many or all aspects of an optically interconnected system simultaneously, rather than concentrating on only one part (such as the smart pixel array or the optical interconnection architecture). Some of these studies are part of a specific development effort, but the results nevertheless have general applicability or implications. [29, 30, 31]

### 5.5. PHYSICAL SYSTEM ARCHITECTURE AND TECHNOLOGY PLATFORM DEVELOPMENT

Many different systems employing optical interconnections in one form or another have been or are being developed. Although these specific devel-

opment efforts may not be considered fundamental, they are important because they represent novelties or breakthroughs in design, architecture, packaging, or systems engineering [32, 33, 34].

It should be noted that considerable subjectivity is involved in determining what should be included in this or the preceding subsection.

## 5.6. ALGORITHMS AND LOGICAL ARCHITECTURES

The different physical architectures supported by optics has created interest in different logical architectures and associated algorithms which better match the abilities of optical interconnections. These often involve massive parallel flows of information [35]. In addition to exploration of various different paradigms (systolic architectures, connectionist systems, cellular systems, etc.), there has also been a revival of work in switching networks [32, 36, 37].

## 5.7. COMPONENTS AND DEVICES

While we have excluded components and devices from the scope of this paper, a number of studies deserve inclusion because of their fundamental implications from a systems viewpoint. For instance, we may mention device comparison studies such as [38, 39], and studies analyzing the scaling properties of refractive and diffractive lenses [40].

Other publications deal with the design of optical components specifically for optical interconnection applications. These include the design of diffractive optical components, microoptical assemblies, and spot array generators.

## 6. Technology evolution studies

These studies may help track the evolution of optoelectronic technology in the context of the history of development of digital computing systems in general, and may also shed light on why a technology with fundamental credentials as strong as optics, is nevertheless not easily able to find its way into mainstream systems [41, 42].

## 7. Conclusion

Although a consolidated body of knowledge and understanding has emerged in several areas, this is particularly not the case for full-blown system-level studies. The different starting points and assumptions makes it difficult to compare the many studies on a common footing. This may be viewed as

a strength as much as a weakness, as the community has not been locked into a specific direction which may be suboptimal.

This work was adapted from [43] (Copyright IEEE).

## References

1. Ozaktas, H. M., Oksuzoglu, H., Pease, R. F. W., and Goodman, J. W. (1992) Effect on scaling of heat removal requirements in three-dimensional systems, *Int J Electronics* **73**, 1227–1232.
2. Ozaktas, H. M. and Goodman, J. W. (1991) The limitations of interconnections in providing communication between an array of points, in S. K. Tewksbury (ed.), *Frontiers of Computing Systems Research, Volume 2*, Plenum Press, New York, pp. 61–130.
3. Ozaktas, H. M. and Goodman, J. W. (1993) Comparison of local and global computation and its implications for the role of optical interconnections in future nanoelectronic systems, *Opt Commun* **100**, 247–258.
4. Kostuk, R. K., Goodman, J. W., and Hesselink, L. (1985) Optical imaging applied to microelectronic chip-to-chip interconnections, *Appl Opt* **24**, 2851–2858.
5. Feldman, M. R., Esener, S. C., Guest, C. C., and Lee, S. H. (1988) Comparison between optical and electrical interconnects based on power and speed considerations, *Appl Opt* **27**, 1742–1751.
6. Miller, D. A. B. (1989) Optics for low-energy communication inside digital processors: quantum detectors, sources and modulators as efficient impedance converters, *Opt Lett* **14**, 146–148.
7. Ozaktas, H. M. and Goodman, J. W. (1990) Lower bound for the communication volume required for an optically interconnected array of points, *J Opt Soc Am A* **7**, 2100–2106.
8. Feldman, M. R., Guest, C. C., Drabik, T. J., and Esener, S. C. (1989) Comparison between optical and electrical interconnects for fine grain processor arrays based on interconnect density capabilities, *Appl Opt* **28**, 3820–3829.
9. Miller, D. A. B. and Ozaktas, H. M. (1997) Limit to the bit-rate capacity of electrical interconnects from the aspect ratio of the system architecture, *J Parallel and Distributed Computing* **41**, 42–52.
10. Ozaktas, H. M. and Goodman, J. W. (1994) Elements of a hybrid interconnection theory, *Appl Opt* **33**, 2968–2987.
11. Lohmann, A. W. and Marathay (1989) A. S. Globality and speed of optical parallel processors, *Appl Opt* **28**, 3838–3842.
12. Ozaktas, H. M., Amitai, Y., and Goodman, J. W. (1991) Comparison of system size for some optical interconnection architectures and the folded multi-facet architecture, *Opt Commun* **82**, 225–228.
13. Lohman, G. E. and Brenner, K.-H. (1992) Space-variance in optical computing, *Optik* **89**, 123–134.
14. Ozaktas, H. M. and Mendlovic, D. (1993) Multi-stage optical interconnection architectures with least possible growth of system size, *Opt Lett* **18**, 296–298.
15. Lohmann, A., Stucke, G., and Stork, W. (1986) Optical perfect shuffle, *Appl Opt* **25**, 1530–1531.
16. Brenner, K.-B. and Huang, A. (1988) Optical implementations of the perfect shuffle interconnection, *Appl Opt* **27**, 135–137.
17. Stirk, C., Athale, R. A., and Haney, M. W. (1988) Folded perfect shuffle optical processor, *Appl Opt* **27**, 202–203.
18. Song, Q. W. and Yu, F. T. S. (1988) Generalized perfect shuffle using optical spatial filtering, *Appl Opt* **27**, 1222–1223.
19. Shamir, J., Caulfield, H. J., and Johnson, R. B. (1989) Massive holographic inter-



- connection networks and their limitations, *Appl Opt* **28**, 311–324.
20. Giglmayr, J. (1992) Generation and evaluation of  $d$ -dimensional ( $d \geq 3$ ) optical shuffle patterns, *Appl Opt* **31**, 4662–4675.
  21. Marsden, G. C., Marchand, P. J., Harvey, P., and Esener, S. C. (1993) Optical transpose interconnection system architectures, *Opt Lett* **18**, 1083–1085.
  22. Ozaktas, H. M. (1992) Paradigms of connectivity for computer circuits and networks, *Opt Eng* **31**, 1563–1567.
  23. Barakat, R. and Reif, J. (1987) Lower bounds on the computational efficiency of optical computing systems, *Appl Opt* **26**, 1015–1018.
  24. Eshaghian, M. M., Panda, D. K., and Kumar, V. K. P. (1991) Resource requirement for digital computations on electrooptical systems, *Appl Opt* **30**, 928–935.
  25. Kiamilev, F. E., Marchand, P., Krishnamoorthy, A. V., Esener, S. C., and Lee, S. H. (1991) Performance comparison between optoelectronic and VLSI multistage interconnection networks, *J Lightwave Tech* **9**, 1674–1692.
  26. Camp, L. J., Sharma, R., and Feldman, M. R. (1994) Guided-wave and free-space optical interconnects for parallel-processing systems: a comparison, *Appl Opt* **33**, 6168–6180.
  27. Krishnamoorthy, A. V., Marchand, P. J., Kiamilev, F. E., and Esener, S. C. (1992) Grain-size considerations for optoelectronic multistage interconnection networks, *Appl Opt* **31**, 5480–5507.
  28. Fan, J., Catanzaro, B., Ozguz, V. H., Cheng, C. K., and Lee S. H. (1995) Design considerations and algorithms for partitioning optoelectronic multichip modules, *Appl Opt* **34**, 3116–3127.
  29. Drabik, T. J. (1994) Optoelectronic integrated systems based on free-space interconnects with an arbitrary degree of space-variance, *Proc IEEE* **82**, 1595–1622.
  30. Rolston, D. R., Robertson, B., Hinton, H. S., and Plant, D. V. (1996) Analysis of a microchannel interconnect based on the clustering of smart-pixel-device windows, *Appl Opt* **35**, 1220–1234.
  31. Ozaktas, H. M. (1997) Toward an optimal foundation architecture for optoelectronic computing, part I: regularly interconnected device planes; part II: physical construction and application platforms, *Appl Opt* **36**, 5682–5696 & 5697–5705.
  32. Hinton, H. S. (1993) *Introduction to Photonic Switching Fabrics*, Plenum, New York.
  33. Irakliotis, L. J., Feld, S. A., Beyette, Jr., F. R., Mitkas, P. A., and Wilmsen, C. W. (1995) Optoelectronic parallel processing with surface-emitting lasers and free-space interconnects, *J Lightwave Tech* **13**, 1074–1084.
  34. Louri, A. and Furlonge, S. (1996) Feasibility study of a scalable optical interconnection network for massively parallel processing systems, *Appl Opt* **35**, 1296–1308.
  35. Krishnamoorthy, A. V. and Miller, D. A. B. (1997) Firehose architectures for free-space optically-interconnected VLSI circuits, *J Parallel and Distributed Computing*, **41**.
  36. Ozaktas, H. M. Ozaktas and Goodman, J. W. (1992) Organization of information flow in computation for efficient utilization of high information flux communication media, *Opt Commun* **89**, 178–182.
  37. Cloonan, T. J., Richards, G. W., Lentine, A. L., McCormick, Jr., F. B., Hinton, H. S., and Hinterlong, S. J. (1993) A complexity analysis of smart pixel switching nodes for photonic extended generalized shuffle switching networks, *IEEE J Quantum Electronics* **20**, 619–634.
  38. Miller, D. A. B. (1990) Device requirements for digital optical processing, in *Digital Optical Computing, SPIE Critical Reviews*, SPIE, Washington.
  39. Fan, C. Mansoorian, B., Van Blerkom, D. A., Hansen, M. W., Ozguz, V. H., Esener, S. C., and Marsden, G. C. (1995) Digital free-space optical interconnections: a comparison of transmitter technologies, *Appl Opt* **34**, 3103–3115.
  40. Ozaktas, H. M., Urey, H., and Lohmann, A. W. (1994) Scaling of diffractive and refractive lenses for optical computing and interconnections, *Appl Opt* **33**, 3782–3789.

41. Krishnamoorthy, A. V. and Miller, D. A. B. (1996) Scaling optoelectronic-VLSI circuits into the 21st century: a technology roadmap, *IEEE J Selected Topics Quantum Electronics* **2**, 55–76
42. Ozaktas, H. M. (1997) Levels of abstraction in computing systems and optical interconnection technology, in A. Ferreira and P. Berthomé (eds.), *Advances in Optics and Computer Science: The Interface*, Kluwer Academic Publishers, Norwell, Massachusetts, pp. 1–18.
43. Ozaktas, H. M. (1997) Fundamentals of optical interconnections—a review, in *Proceedings of the Fourth International Conference on Massively Parallel Processing Using Optical Interconnections (MPPOI'97)*, IEEE Computer Society, Los Alamitos, California, pp. 184–189.