Feasibility of Optical Interconnection Networks Theory and Case Study

Valeria Bertacco, Ujval Kapasi, Brucek Khailany, Edward Lee

Abstract — **Interconnection networks have traditionally been implemented with electrical signaling. However, recent developments in optical interconnect technology has made it a competitive alternative. There are many fundamental advantages to using optical interconnect as the communication medium. Although optical interconnect technology is still inferior to electrical interconnect technology in terms of cost and maturity, it does make sense in certain situations where electrical signaling might not do so well. This paper addresses some of these situations and discusses several architectures specifically tailored for them.**

1. Introduction

Optical interconnect has long been employed in long-distance, high-bandwidth telecommunications. However, for interconnection networks that span relatively short distance (e.g. chip-to-chip, board-board, cabinet-to-cabinet), optical interconnections are still relatively rare. That we have not used optical communications more up to now is due to many sound reasons: optics may have been impractical because the necessary devices, though physically possible, were either not realized or not easily manufactured; it may be unfamiliar to designers; applications requiring sufficiently large interconnection networks to justify the use of optics may not have emerged; architectures best suited for optical interconnection networks may not have been studied as thoroughly as those for the electrical counterparts; electrical interconnections have been working in a very broad range of applications and symbolized cost-effective and mature technology; a significant amount of opto-electric and electro-optical conversion overhead is required because the capability of optical switching devices are not even close to that of the electrical counterparts.

Nonetheless, there are many physical and architectural reasons for using optical interconnections. It is not the purpose of this paper to argue for either optical or electrical interconnections for all cases. Rather, the objective here is to provide a comparative study and discuss reasons why we would want to use optics-based interconnection networks in some situations. Section 2 provides some background on the state-of-art electrical and optical signaling and discusses the pros and cons of each. Section 3 describes some network architectures specifically targeted for optical interconnections and compares them with the more traditional electrical interconnection networks. Concluding remarks are given in Section 4.

2. Electrical and Optical Signaling Comparison

2.1 Latency

It is worth noting first that both electrical and optical signals are electromagnetic in nature. This point is often misunderstood in the case of electrical signaling: it is not electrons or other charge carriers that carry the signal; rather, it is the electromagnetic wave. In typical electrical cables (LC lines), signals travel essentially at the speed of light (or slightly slower due to higher dielectric constant) rather than the speed of electrons. To say that optical interconnects have lower latency is grossly incorrect. In fact, signals typically travel slightly slower in optical fibers than in coaxial cables because the dielectric used in cables has a lower dielectric constant than glass. Furthermore, with the current state-of-art, most of the computations in a system have to be done in the electrical domain; therefore, there is a significant latency involved in domain conversion if optical interconnects are employed.

In the case of on-chip electrical connections, signals do move slower because the wave is moving in a *dissipative* fashion. The typical speed is 10^7m/s [3]. However, in a typical interconnection network, wire delay is dominated by off-chip LC lines.

2.2 Bandwidth

The bandwidth (or data rate) of an electrical medium is limited by frequency-dependent (both skin-effect and dielectric loss) attenuation. In particular, frequency-dependent attenuation causes a phenomenon known as inter-symbol interference (ISI) in which signals both in the past and in the future alter the current signal in such a way that it is no longer reliably detectable. For most of electrical cables, attenuation is dominated by skin-effect, and the following metric gives the bandwidth of a cable given its length and allowable attenuation [1]:

$$Bd^2 = 2f_1\left(\frac{A_T}{A_1}\right)^2 \quad (1)$$

where B is bandwidth in b/s, d is the length of the wire in m, A_1 is the attenuation of a 1m wire at frequency f_1. For a

given allowable attenuation, (1) indicates that Bd^2 is constant. For example, Bd^2 for an unequalized 24AWG pair is 4×10^{10} bits×m^2/s (40Gb/s over 1m or 400Mb/s over 10m). Better equalization can improve this figure by allowing a larger A_T. If we assume perfect equalization, a transmit level of 1V, no coding, a minimum of 20dB SNR for reliable signaling, and that the Johnson noise is the only fundamental noise source and the noise temperature is 300K, then it can be shown that the maximum capacity (note that this is different from Shannon Capacity) of a 24AWG pair is 1Gb/s over 230m, 10Gb/s over 63m, or 100Gb/s over 17m [1].

Optical interconnects, on the other hand, exhibit no significant frequency-dependent attenuation. The carrier frequency of light is so high compared to the modulating (information-bearing) frequency that any modulation of the light beam makes essentially no difference to the propagation of light. There is some dispersive effect and propagation loss over long distance. However, they are so small that they can be safely ignored for fibers with distance up to several kilometers. Also, optical fiber can be made extremely small in diameter (e.g. 125 microns). As a result, optical interconnections can readily exceed the bit rate capacity of electrical interconnections by 9 orders of magnitude for the same cross-sectional area and length [3].

Note that the above bandwidth limitation only considers the physical media. There is obviously a constraint on how fast the on-chip switches can go. At present, the fastest signaling system achieved with standard CMOS technology is 10Gb/s over 10m of PE-142LL cable (about 6.48dB of loss) [4]. It is expected that both the electrical and optical signaling rate will improve with Moore's law due to faster on-chip electrical switches [1] [2].

2.3 Packaging and Layout

Optical signaling allows the possibility of making *free space* interconnects. Figure 1 shows an example of this [6].

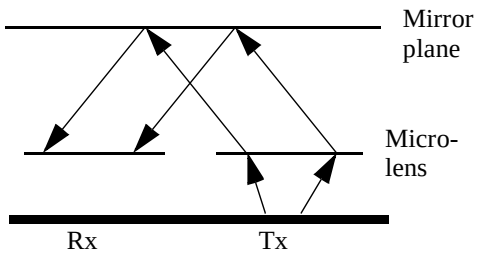

Figure 1: Free space optical interconnects. Through microlens refraction and mirror plane reflection, it is possible to route from arbitrary points in the transmitter to arbitrary points in the receiver.

A benefit of optical interconnects is that all of the connections are very well synchronized with one another, hence avoiding any signal skew. Free space interconnects have the additional benefit that signals can be routed in the third dimension and cross each other without any interference. These features are useful in very global interconnection network topologies like perfect-shuffle, butterfly, and banyan. Furthermore, as we shall see later, because free space optical interconnects allow the possibility of 3 dimensional routing, bisection bandwidth, a factor derived from the planar nature of electrical interconnects, is no longer a suitable constraint of the network, and as a result higher dimension torus and mesh could be built with less penalty using optical interconnects [6].

One could argue that free space communications have been done in electrical domain (essentially wireless communication). However, a significant difference between optical and electrical signaling is that optical light beam has a much shorter wavelength. Wavelengths of electrically-driven signals are usually comparable to or larger than a circuit board, and certainly larger than a chip. The law of diffraction says that it is difficult to focus a wave to a dimension smaller than its wavelength. Wavelengths of optically-driven signals, however, are much smaller. It is possible to have more than 200,000 free space connections on a 12cm×12cm board [6].

Since optical interconnects operate on the principle of detecting photons, it is possible to generate voltages and currents at the receiving end without any direct electrical connection with the light beam source. In other words, optical interconnects give us perfect electrical isolation between two circuits, and the absolute voltage levels between different parts of the system are irrelevant. This solves the problem of voltage variation in electrical systems and allow much larger systems to be built [3].

2.4 Power

Power consumption of an electrical signaling system is limited by the amount of the current necessary to drive a typical 50-ohm line. If the minimal allowable signal swing is 50mV and the power supply is 2.5V, it would require 2.5×(50m/50) = 2.5mW of power. Notice that this does not take into account the power consumption of doing additional pre-emphasis on the line. Optical signal, however, works on a completely different principle from electrical signals, and the voltage generated in a photodetector bears no relation to the classical voltage in the light beam. It is possible to generate 1V in a photodetector from a light beam with 600mV of classical voltage, a consequence of photo-electric effect [3]. Photon energy of light (~2eV) is also much larger than the photon energy of electrically-induced waves (40neV - 40ueV). Traditional optical interconnects have not been able to exploit this fundamental advantage because optical drivers alone consume more

energy than a 50-ohm resistor. However, the emergence of quantum well modulator technology has led to practical low power optical output devices with < 6mW of power operated at 375Mb/s. [3]

Besides the issue of energy-inefficient optical drivers, optical interconnection networks also suffer from losses in the signal paths. Although an optical fiber has almost no propagation loss (as mentioned previously), it does suffer from losses when passing through optical connectors or switches (if an all-optical network is implemented). Although inserting solid state amplifiers (SOA, the equivalent of repeaters) could solve this problem, they are very expensive and take up a good deal of board area. Therefore, a good deal of power has to be budgeted for this kind of loss (30-40dB) [6].

2.5 Technology Summary

Both electrical and optical signaling have approximately the same signal propagation delay because both operate with electromagnetic waves. Where optical signaling might do better is in the areas of bandwidth, packaging/layout, and power. All of these come from the underlying physical differences between light and electricity. Optical signaling has orders of magnitude more data rate capacity than electrical signaling. The opportunity of free space optical interconnects also allows more flexible layout and packaging. Finally, because the photon energy of light is so much larger than electrically-induced waves, optical interconnects can operate with much lower power consumption with appropriate devices. All of the above advantages could potentially be exploited to build an interconnection network system with a much better performance/area/power combination than the existing systems.

3. Impact of Optical Technology on Interconnect Architectures

Optical interconnect technology differs in many respects from conventional electronic interconnect technology, as shown in the previous section. Since the architecture of an interconnect is highly dependent on the underlying technology, it is interesting to explore the sorts of interconnect architectures that are feasible with optical technology, and whether these provide significant benefits over electronic interconnects. The following are examples of interconnection networks that previous researches have proposed (and in some cases implemented). Each one will be discussed in the context of the aspects of optical interconnect technology that motivate each network.

3.1 Parallel Electrical / Optical Networks [6]

This system (Gemini interconnection network) uses the high-bandwidth optical technology to target applications that are bandwidth limited, such as Space-Time Adaptive Processing (STAP). However, since control of the optical circuitry and optical memory technology are not readily available, the control has to be done electrically. Thus, the Gemini interconnection network uses a electrically controlled optical network and a new memory system designed to provide high bandwidth to the network.

Short messages are sent over the electrical interconnect, and long messages are sent over the optical interconnect. The optical interconnect uses circuit switching, and the messages to set up the optical paths are sent over the same electrical interconnect as the short messages. Thus, this design works well when a significant percentage of the messages sent over the network are large.

In order to provide enough bandwidth from the memory system to the network, the network visible memory is interleaved, and the network interface has gather/scatter capabilities. This allows large amounts of data to be transferred over the network, and efficiently executes operations such as large matrix transforms.

3.2 Optical Time Division Multiplexing (OTDM) [5]

This technology also takes advantage of the high bandwidth provided by optical link technology. The basic premise of OTDM technology is to time multiplex several lower bandwidth channels onto a single high bandwidth fiber. The key to this technology is the Terahertz Optical Asymmetric Demultiplexer (TOAD) device [5]. This device is essentially an optical AND gate that allows effective demultiplexing of the optical signal.

Laboratory experiments have indicated that using the TOAD device, bisection bandwidths of up to 250 Gbits/sec can be achieved, and that the technological limit is above 5 Tbits/sec [5].

The authors asserts that the OTDM technology can be used to effectively implement a large crossbar interconnect. Large crossbars are very difficult to implement with electronic technology. This is mainly because the number of switches needed is $O(N^2)$. Further, crossbars do not map well to electronic packaging technology since they are completely pin-limited.

In contrast to electronic technology, OTDM technology effectively implements large crossbars by time multiplexing all the node traffic onto one high bandwidth link. Then, the signal on this link is received at every node and demultiplexed appropriately. Each receiving node has an assigned channel (i.e., time slot) assigned to it in the base case. Fancier channel assignment schemes can be used to allow broadcasting and multi-casting. Furthermore, the bandwidth of each node can be increased by assigning more than one channel to it and replicating the demultiplexing circuitry.

The authors also presents an arbitration scheme which allows only one transmitter on each time-multiplexed channel every packet slot. Fairness is implemented using a form of rotating priorities. This scheme provided an effective throughput of 60% with large numbers of nodes. A slightly more complicated scheme allows two arbitration cycles every packet slot, allowing all transmitters that were denied a channel on the first arbitration to submit a request for a different channel. This is useful when nodes are capable of queueing packets to be sent and when out-of-order packet delivery is acceptable [5]. Using this scheme, throughputs of approximately 75% can be achieved.

An interesting application of this network, the paper shows, is to multicomputers. Message passing multicomputers benefit from the multi-casting and broadcasting capabilities, which are normally difficult to implement in electronic interconnect networks. Shared memory multiprocessors benefit greatly from the "globally visible event ordering" in OTDM. This allows common synchronization operations to be implemented using the OTDM interface directly [5]. Implementation of cache-coherency directory protocols can also be simplified since all events are ordered and occur in the span of a packet cycle. In other words, certain issues can be ruled out, such as pending messages and out of order delivery, etc.

3.3 Free-Space Optics [7]

Free space optical technology presents the possibility of using the third dimension to route signals, as opposed to the planar electronic routing. In [7] an analysis is given of $\underline{k}$-ary n-cube networks implemented with free-space optical technology.

Free space optical interconnects are implemented using a planar mirror and a plane of microlenses (see Figure 1). Each microlens is associated with a transmitter or receiver. A microlens has a diameter of 125μm in the technology available at the time the paper [7] was written. The transmitter and receiver circuit arrays are integrated with the router chips using flip chip bonding.

The paper analyzes the performance of k-ary n-cube networks and claims that the correct metric for free-space optical interconnects is not bisection bandwidth. Although bisection bandwidth is a useful notion for planar VLSI technology, it is not representative of the 3D nature of free-space interconnects. In this case, the authors of the paper propose a new metric called *connection capacity*. Connection capacity is a measure of the total connections needed in a network, and is simply :

$$C = nk^n W \tag{2}$$

Here, W is the width of each channel. Thus, now the effect of higher dimension networks on the channel width of a network can be analyzed by normalizing the channel capacity to that of a hypercube with the same number of nodes and channels with a width of 1. That results in a channel width of :

$$W(k, n) = \log_2 k \tag{3}$$

Thus, as the *aryness* of a network is increased (or its dimensionality reduced), the channel width increases logarithmically. For electrical interconnects, which are limited by bisection bandwidth, the channel width increases linearly with k. This linear factor in electrical networks motivates designers to decrease the dimensionality of the network in order to reduce the latency due to the serialization of the message and to increase the speedup of the network (assuming a constant bisection bandwidth).

In contrast, free-space optical interconnect technology more efficiently implements higher dimensional networks. It is clear that lowering the dimensionality of a free-space optical network is not as large a win (logarithmic vs. linear), and the increase in hop count that accompanies decreasing dimensionality might have a larger impact on the overall performance. In addition, their calculations indicated that networks implemented with fairly aggressive electrical and optical technology differ in that the optical one can support wider channels (the network fit on a single PCB board in both cases). Thus optical network latency is less dependent on message length to begin with, so moving to lower dimensional networks is not as important.

4. Discussion

4.1 Gemini Interconnect

The first thing to consider is what all of the above case studies gain from using optical technology. The Gemini network uses circuit switched optical paths for long messages. Thus, if the bandwidth of the optical path is higher than that can be achieved by electrical paths, the optical network is better. The lasers that the authors considered in the paper could operate at 2.5Gb/s, and they mentioned that 10 to 20Gb/s lasers would be affordable soon [5].

However, the bandwidth of a single channel is not all that matters. For example, the optical switches used in this network did not include repeaters. Yet any non-ideal passive device must have a loss associated with it. The authors calculated that with the current power budget (1mW laser source and -30dBm receiver sensitivity) only 32 node switches could be built. They predicted 256 node switches by the year 2002. Thus, the stability of the current technology is suspect, unless a low latency, cheap, and small active optical switch is developed.

Finally, the ease of integration of the optical technology should be considered. The optical switch used by the authors was 1.5cm by 12cm in the technology that was current when the paper was written. The switch is perhaps a

little smaller now, but is still much greater than the size of a similar switch implemented in CMOS. The question then is, can you implement a switch that performs the same function as the optical switch and provides just as much bandwidth in the same area, with the same power budget, and the same price?

If electrical signalling and switching can give the same or nearly the performance, using the optical technology cannot be justified. However, the authors unfortunately did not run the proper experiments and provide the proper results. They compared the performance of the dual electrical and optical networks with an all electrical network. However, all the messages in the electrical network only used the packet-switched portion of the Banyan network they were using. It would have been nice if they had at least two more data sets : 1) the performance of the best electrical interconnect they could build with the same or lower cost and power; and 2) the performance of a network in which all the optical components are replaced with electrical ones -- in other words a similar network with both the packet-switched and circuit-switched portions implemented with electrical signalling and switching.

4.2 OTDM Crossbar

The OTDM network presented above in section 3.2 seemed to be a more interesting and viable idea. This network used a truly optically controlled switch, the TOAD. The free-space interconnect just used static optical routes, and the Gemini paper used optical switches, but even these were controlled electrically, and thus bound to the same frequency limitations that all electronic components are bound to.

In contrast, the OTDM makes better use of the potential bandwidth available on a single optical fiber link. The backbone of the network is simply one fiber with several (1 - 1024) electrically modulated signals time multiplexed onto it. This usage of the optical technology seems to be better in almost every way than the methods used in the other two architectures. The free-space interconnect gets most of its gain from the fact that optical signals can be routed through free space and that more optical connections can be made per unit area of chip space. The Gemini network plans to use wavelength division multiplexing, which requires a separate laser for each channel that is multiplexed.

Not only is the use of the optical technology good with OTDM, but the authors implement the best network topology possible: a crossbar. The arbitration scheme the authors use is clever, but not optimal. Finally, the only real critique currently is the size and cost of the network. Currently the size of a TOAD device is approximately 10 in^2 as estimated from the photo they provide, and the cost is approximately $20,000 per node. However the TOAD device shown was only experimental and production model will surely be more compact, and the authors expect the price per node to drop to less than $1,000 in the future.

Overall, their application of optical technology targets a class networks which are not viable in the electrical interconnect domain. With the right optical technology a decent crossbar can be built, and this is a reasonable motivation for continued research into the technology presented in the paper.

4.3 Free-Space Optical *k*-ary *n*-cube networks

The paper that discussed the analysis of free-space interconnect networks. The arguments presented for building higher dimensional networks with free-space technology make sense. However, the original premise behind using free-space optical interconnect technology is not perfectly sound. A free-space optical *k*-ary *n*-cube does provide lower latencies and higher bandwidths than their electrical counterparts. However, free-space interconnects have stricter packaging constraints than electrical interconnects. First, nothing must be allowed to interfere with the free-space optical signals. Second, the packaging does not seem scalable. Some solution must be found to scale the optical technology better, for instance, how to make connections between two boards that use free-space optics. Also, they made a comparison to similar networks implemented in planar VLSI technology. However, it would have been nice to have seen some sort of reference or comment on the possibility or feasibility of a 3D electrical interconnect.

5. Conclusion

We have shown that there are fundamental advantages to using optical interconnects over electrical interconnects. These include links with much higher bandwidth, longer signaling distances, lower power, and more efficient packaging. Theoretically, these improvements over electrical signaling could lead to much larger and higher performance interconnection networks.

We have also shown that there are interconnection network architectures that could take advantage of optical interconnects [5] [6] [7]. These architectures have exploited optical interconnects for high bandwidth links and free-space optical interconnects for more heavily connected topologies. In addition, one could also envision building very large interconnection networks connected by optical fibers. These networks could have many more nodes than current systems because link distance is no longer as important an issue.

However, the main reasons we have not seen optical interconnection networks built in volume is due to engineering constraints. One such engineering constraint is cost. For example, the current OTDM hardware has a cost of around $20,000/node [5] (as compared to hundreds of dollars per node for an electrical network). Another con-

straint is technology maturity. Devices that efficiently exploit the advantages of optical interconnects and interface easily to electrical systems are either not easily manufactured in volume or just have not been realized yet.

In addition to the engineering constraints just mentioned, more research needs to be done to find which network architectures which are best suited to take advantage of the increased raw performance provided by optical technology.

Some progress has been made in recent years to overcome these shortcomings and improve the availability of optical interconnect devices [2] [3] [5]. As these trends continue, it will become more cost-effective to build optical interconnection networks. Fundamentally, optical signaling has advantages over electrical signaling that will allow for larger, higher-performance interconnection networks in the future.

6. References

[1] W. Dally, M.-J. E. Lee, F.-T. An, J. Poulton, S. Tell, "High performance Electrical Signaling," *MPPOI98*

[2] David A. B. Miller, "Dense Optical Interconnections for Silicon Electronics," in *Trends in Optics: Research, Developments, and Applications*, Vol. 3, 207-222, Ed: A. Consortini (Int'l Commission for Optics, Academic Press, 1996).

[3] D. A. B. Miller, "Physical Reasons for Optical Interconnection," *Special Issue on Smart Pixels*, Int'l J. Optoelectronics 11 (3), 155-168 (1997).

[4] R. Farjad-Rad, C.-K. K. Yang, M. A. Horowitz, T. H. Lee, "A 0.4-um CMOS 10Gb/s 4-PAM Pre-Emphasis Serial Link Transmitter," *IEEE JSSC*, vol. 34, no. 5, May 1999.

Case Studies:

[5] A. G. Nowatzyk, P. R. Prucnal, "Are CrossBars Really Dead? The Case for Optical Multiprocessor Interconnect Systems," *ISCA95*

[6] R. D. Chamberlain, M. A. Franklin, R. R. Krchnavek, B. H. Baysal, "Design of an Optically-Interconnected Multiprocessor," In *Proc. of the 22nd International Symposium on Computer Architecture*, June 1995, pp. 106-115.

[7] M. Raksapatcharawong andT.M. Pinkston, "An Optical Interconnect Model for k-ary n-Cube Wormhole Networks," *Proc. 10th Int'l Parallel Processing Symosium*, pp 666-672, Apr. 1996.