Original manuscript formatted following Applied Optics prescriptions

Architectural approach to the role of Optics in monoprocessor and multiprocessor machines

Jacques Henri Collet, Daniel Litaize, Jan Van Campenhout, Chris Jesshope,
Marc Desmulliez, Hugo Thienpont, James Goodman, and Ahmed Louri

The relevance of introducing optical interconnects (OI) in mono and multiprocessors is studied from an architectural point of view. We show that perhaps the major explanation for why optical technologies have nearly been unable to penetrate *into* computers, is that optical interconnects generally do not shorten the memory access time, which is the most critical issue of today's stored-program machines. In monoprocessors, the memory access time is dominated by the electronic latency of the memory itself. Thus, implementing OIs inside the memory hierarchy without changing the memory architecture cannot dramatically improve the global performance. In strongly coupled multiprocessors, the node bypass latency dominates. Therefore, the higher the connectivity (possibly with optics), the shorter the path to another node, but the more expensive the network and the more complex the structure of electronic nodes. This relation leaves the choice of the "best" network open in terms of simplicity and latency reduction. The bottlenecks and the benefice of implementing OI are discussed in symmetric multiprocessors, rings and distributed shared-memory supercomputers.

OCIS Codes: 200.4650, 200.2610

I. Introduction

Although numerous studies are in progress all over the world for developing short-distance optical interconnects, it clearly emerges from the literature that most of them are based on technological arguments, and that the global operation of the targeted architecture is not fully analyzed. The proceedings of the conferences that constitute Ref. [1,2] may provide a non-exhaustive presentation of the current state of the art in this field. Many studies attempt to improve some part of the machines, but offer no insurance that this progress improves the global operation of the whole system.

In this work, we follow a different approach that consists of analysing the role of short-range OIs in monoprocessor and multiprocessor machines from the architectural point of view. Most of the discussion throughout the paper focuses on the relevance of OIs in reducing memory access latency (MAL), which is the most critical and permanent issue in stored-program computer architectures. The consideration of optics leads to the following paradox: On the one hand, OIs extend the communication bandwidth but generally do not *directly* change (or address) the access latency to the memory, which is dominated by electronic processing times, both in monoprocessors and tightly bound multiprocessors. On the other hand, OIs may increase the connectivity of inter-processor networks (thus reducing the path that separates remote nodes) at the expense of reinforcing the latency problem to the electronic domain and in particular to the design of efficient node circuits. The choice of the "best" network is therefore an open issue in terms of implementation simplicity and latency reduction. These difficulties may partly explain in part why OIs today are not deployed in general-purpose machines and are considered for use potentially in dedicated processors (that do not execute instructions or fetch data stored in a memory).

J.H. Collet (Collet@laas.fr) is with the Laboratoire d'Analyse et d'Architecture des Systèmes du Centre National de la Recherche Scientifique, 7 av. du colonel Roche, F-31077 Toulouse CEDEX, France.D. Litaize (Litaize@IRIT.fr) is with the Institut de Recherche en Informatique, Université Paul Sabatier; 118 route de Narbonne F-31062 Toulouse France. J. Van Campenhout (Jan.VanCampenhout@rugac.be) is with the Vakgroep Universteit Gent, St. Pieternieuwstraat 41, B-9000 Gent, Belgium. C.Jesshope (c.r.jesshope@-Massey.ac.nz) is with the Institute of Information Sciences and Technology, Massey University, New Zeeland. M. Desmulliez (m.desmulliez@hw.ac.uk) is with the Computing & Electrical Engineering, Heriot-Watt University, Riccarton EH14 4AS, UK. H. Thienpont (hthienpo@vub.ac.be) is with the Laboratory for Photonics, Vrije Universiteit Brussel Pleinlaan 2, B-1050 Brussels, Belgium. J. Goodman (goodman@cs.wisc.edu) is with the Department of Computer Sciences, University of Wisconsin, Madison, WI 53706,USA. A. Louri (Louri@ece.arizona.edu) is with the Electrical and Computer Engineering, University of Arizona, Tucson, AZ85721.

The contents of the paper are the following. The current state of the application of optical technologies to communication and interconnection networks is reviewed in Section II. The memory issue and its influence on the architecture of today's machines are reviewed in Section III. In Section IV we discuss the relevance of implementing OIs in monoprocessors. No dramatic increase in global performance is expected in such systems as the intrinsic memory latency is the dominant latency. The potentially stronger impact of OIs in multiprocessor machines is discussed in Section V. Simplicity of implementation is often preferred for small multiprocessor machines, making the introduction of OIs particularly attractive in symmetric multiprocessors (SMP's) and ring architectures for connecting about 100 nodes. In these architectures, OIs can provide a huge bandwidth that can minimize contention latency (related to the traffic saturation) as is also explained in Section IV while they maintain the simplicity of the node structure. In supercomputers, providing a global shared view on a physically distributed memory places a heavy burden on the interconnection network and, in particular, on low-latency high-connectivity electronic nodes. The introduction of OIs in new reconfigurable architectures and in dedicated processors is briefly discussed in Sections VI and VII, respectively.

II. Brief review of the Role of Optics in Communication Networks

The current state of the competition between optics and electronics for the processing and transmission of information is reviewed here as a function of the communication distance.

A. Telecommunication networks

Optical communications have won the battle for long-distance transmissions in wide area networks (WAN's) and metropolitan area networks (MAN's). There are at least three reasons for this: 1) The bandwidth limitation of OIs is much less pronounced than that of electrical transmission [3], losses are much lower, and in future systems the effects of non-linear dispersion can be countered by use of solitons. 2) Parallel transmissions are not usable over long distances because skew makes the synchronisation of the different reception channels particularly complex. 3) Multiwavelength (optical) transmissions make possible the extension the transmission bandwidth at almost no cost for the network infrastructure. Thus one permanent objective of long distance communications consists of increasing the transmission bandwidth through a single mono-mode fibre.

B. Local Area Networks

Local Area Network were first designed for data transmission between computers. We may distinguish between company networks and industrial networks that operate in a hostile environment with real-time constraints. Each computer (PC, workstation) in a company network is connected to a hub through a few tens of meters of links and operates with an Ethernet protocol. Hubs are themselves interconnected by means of high-throughput links operating under various protocols such as Ethernet, fiber distributed data interface (FDDI), and Asynchronous Transfer Mode (ATM). Links from computers to hubs are generally implemented with the preinstalled metallic cables of the building network, whereas serial optical (FDDI) links are mostly used for connections between hubs. Industrial networks also exhibit a hierarchical structure. Each level may use a specific field bus as the Profibus [4], the IEC Fieldbus [5], the Controller Area Network (CAN, developed by Bosch for cars [6]), aviation industry standards like ARINC 429 or 629 (developed by the Airline Electronic Engineering Committee), the Manufacturing Automotive Protocol (MAP, developed by General Motors), the interbus S [7], etc.

C. Short-distance communications and interconnects

The transition from serial telecommunications to the computer world (dominated by parallel interconnects) occurs at short-distance extending over a few meters. A large bandwidth is needed for computer clusters and multiprocessor interconnects, as for instance in the Cray model T3E [8], the IBM model SP2 [9], the Intel model Paragon [10], the Silicon Graphics Model Origin system [11]. Electronic parallel interconnects dominate because they allow, across a few meters, the extension of the global bandwidth without increasing the operation frequency. These interconnects are generally a cheap solution. However, several networks that were implemented with paralleled electrical links now offer serial (or parallel) optical alternatives for increasing the bandwidth and the transmission distances. Some of these networks are HIPPI (high-performance parallel interface) at 6.4Gb/s [12]; SCI (scalable coherent interface) at 1.6 Gb/s [13]; Myrinet at 1.28 Gb/s [14] and so on. These networks makes possible the building of multicomputers for supporting cluster computing, which is an area in which there is a lot of experimentation at the moment. Cluster computing is partly motivated by the preoccupation of developing more modular, low-cost hardware that would simplify maintenance and compatibility issues for manufacturers. However, it must be stressed that cluster computing is suitable for some, but not all applications because the latency of internode communications becomes extremely long (with respect to the processor cycle, it is currently 1ns) when the inter-computer distance attains a few meters (one meter translate to 5 ns). Therefore, a distributed system will execute numerous applications much more slowly (especially applications with a distributed memory and those that require numerous internode exchanges) than does a tightly bound multiprocessor enclosed in a single cabinet.

D. Ultra-short distance interconnects

At present, ultrashort-distance interconnects ranging from a few centimeters to a few tens of centimeters are in the electronic domain. The machines under consideration here are monoprocessors (PC and workstations) or SMPs such as the Silicon Graphics modelPower challenge [11], the Sun model Enterprise [15]. In these machines, the communication latency is never controlled by the propagation (internode or inter-units) but by electronic terms (memory latency, routing time, etc…). This control is an essential difference from the distributed systems described in the previous section. The bandwidth extension has been achieved to date by by the increasing of the transmission parallelism and by the replacement of shared busses (which are multipoint electrical lines) by dedicated point-to-point parallel interconnections. For instance in the Pentium architecture data is transmitted between the memory controller and the chip set through a 64-bit-wide bus at 100-MHz, and possibly in the future with 128 bits to attain 12.8 Gb/s [16].

E. Intra-chip and MCM interconnects

By far, most of the communications at this level are in the electronic domain. However, an optical clock distribution exists at the inter-board level in the Cray model T-3D [17] with an research is being carried out to extend this technology at intraboard level for the Cray model T-90 [18]. On the research side studies are in progress for the construction of optical backplanes [19-20] and on micro-optical components for interchip and intra-hip communications (Vrje Universiteit Brussel [21, 22]).

F. Optics in the logic-level processing

Most of the all-optical computing studies launched in the middle of the eighties have been reoriented towards special-purpose systems because: 1) The dramatic increase in electronic processing power has progressively eliminated many arguments that favored optical binary processor (switching time, switching energy) and 2) The performance of optical numerical demonstrators comparatively has stagnated. Today, interest in optical processing seems limited to dedicated processors.

Parallel optical I/O may be exploited for information transport in smart pixels, early vision processing, artificial retina applications [23], as well as database and symbolic applications [24]. The low-level processing in the field of pattern and image recognition might take advantage of the potential of optics for the extremely fast implementation of simple Boolean functions that were carried out in the field of ultra-fast optics [25].

III. Optical technologies and the memory issue

A. General remarks

In Section II, we showed that OI's have ruled the telecommunication world by the replacement of electrical links with optical serial links, without affecting the communication protocols or the network architecture. They will likely invade local-area networks similarly. The penetration of OI's *into* computers is much more problematic despite numerous advantages that favor optical technologies [26]. Some of these arguments are listed below:

- Optical technology can provide extremely high bandwidth, almost independent of interconnect length at the length scales considered.
- OI's have advantages in terms of weight and volume. The interconnect density is potentially higher because of 1) the much lower interference of optical signals, either in free space (electron beams vs. light beams) or when guided (mutually interfering conductors vs. optical fibres), and 2) the fact that each optical communication channel requires a cross-section of the order of the wavelength of the light used. Hence, per given total cross section and interconnect length, a larger number of interconnects is possible optically. Connection of a high-density OI's is much less bulky than that of electrical interconnects.
- Broadcasting is feasible because of the capability of high fan-out. However, high fan-out requires high power, which introduces some limitations such as an increase in the latency.
- Wavelength Division Multiplexing (WDM) can be used to increase bandwidth and achieve special advantages. Though requiring tuneable light sources, there is no interference between wavelengths; in terms of networking, add-drop capabilities and wavelength routing are possible.
- Optics is at least competitive in terms of power and control supply requirements: the speed-power product of advanced centimetre-range OI's is becoming better than that of electrical ones [27,28,29].
- Optics may be capable of the rapid reconfiguration of static interconnection patterns. Light polarization also makes possible switching and fast configuration routing. Apart from power losses, the means needed for optical reconfiguration do not impair signal quality or even latency, as is often the case with electrical switching or reconfiguration.
- OI's provide galvanic isolation between the interconnected subsystems. This leads to improved noise immunity and security for applications monitoring high-voltage systems.

With all these arguments, how to explain the paradox that optical technologies have so far hardly gained application in computers? We feel that the underlying reason is that optics generally does not directly shorten the access time to the memory that is the most crucial issue of stored-program machines. Therefore it is extremely difficult to translate the numerous technological arguments that claim optical advantages into an optical architecture (or a demonstrator) that might effectively overcome the electronic computers (except for some dedicated applications related to vision or image processing). We briefly review the memory problem in the next paragraph.

B. Memory issues and architectural evolution of machines

The evolution of stored-program computers has repeatedly been influenced by the fact that the performance of microprocessors has increased much faster than that of memory systems. In early microprocessor systems (i.e., in the 1970's), processors operated at about the same rate that memory could be cycled, and the processor was connected directly to the memory system (dynamic RAM, DRAM) [30]. This is no longer the case. Whereas processor speeds have been doubling every few years, dynamic memory has increased in speed only marginally over the last two decades, although its size has also doubled every few years. This lack of improvement in speed means that the time needed by the processor to fetch instructions or data from the main memory has increased permanently compared with the processor cycle time making direct exchanges with the main memory more and more penalising for the *global performance* of the computer.

Latency is the key parameter of memory-processor interactions (much more than the bandwidth as in telecommunications) because the processor exchanges very short bursts of information (usually at least one word, i.e., four bytes, more often a cache block at a time, i.e., 32 or 64 bytes). The processor never establishes a steady communication stream with the memory. The major limitation in the speed of DRAM is in the circuits used for detecting the stored charge on a memory cell. There is a trade off between the size of the memory and the rate at which this tiny stored charge can be sensed. Static RAM (SRAM), on the other hand, is optimised to be significantly faster than DRAM, although this is achieved at a cost of a larger memory cell and therefore, a significantly reduced memory size (or an increased memory cost). For this reason most desktop and server systems use DRAM memories to maintain large memory size at modest cost. The architecture of chips and machines has evolved permanently to maximise global performance despite the degradation of the MAL. A typical example is shown in Fig. 1 for the PC architecture. The evolution in PCs and multiprocessor machines has consisted primarily of *hiding* the MAL with hardware or software solutions because it has been impossible to change the memory technology. Therefore:

- Modern computer architectures use a complex hierarchy of memories. Two on-chip level 1(L1) caches are provided in pipelined architectures (one for data and one for instructions) as data and instructions must be read concurrently. These caches are typically 32–64 Kbytes in size (128 KB in the next processor model K7 from AMD). L1 caches may then be connected to a second-level cache (L2), also on-chip, or a much larger off-chip cache. The latter, being typically around 1 Mbytes in size. If there is a second level on-chip cache, there may also be a third level cache off-chip (L3). The off-chip cache will then be connected to main memory. Caches work by exploiting locality of references in access to data, either spatially, where data adjacent to that recently used is likely to be reused, or temporally, where data recently used is likely to be used again. The aim of a cache is to make the memory system appear to be as large as the largest component and to appear as fast as the fastest component. Unfortunately when the cache system does not work well, through lack of locality, the slowdown is severe, as the DRAM memory access time is at least an order of magnitude larger than the processor in its cycle time.

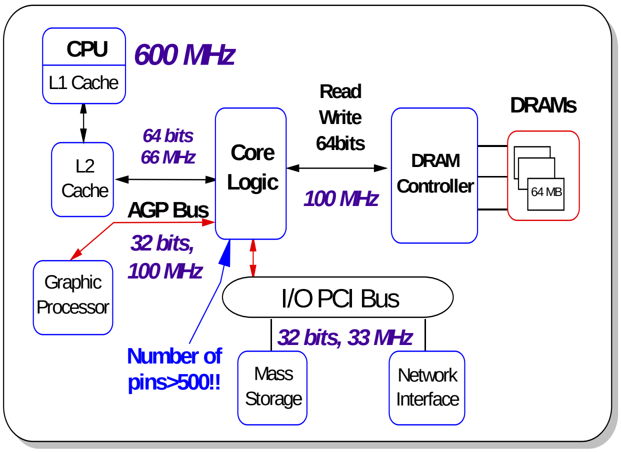

***Figure 1**: Current PC architecture. Notice the two levels of caches (L1, L2) and the absence of the shared memory bus replaced by dedicated point to point connections between the logic core, the graphic processor, the DRAM controller and the L2 cache. Note also that the number of pins in the core logic is larger than 500!*

- Current microprocessors are designed to *hide* the latency associated with a memory fetch. Techniques used to tolerate high-latency memory include speculative execution in which the results of branches and even data values are predicted. When a miss-prediction occurs, data generated along wrong branch paths or based on mispredicted values must be cleaned up and the original state restored. Other techniques used include out-of-order execution in which instructions start and even terminate before previous instructions in the instruction stream. Operands of instructions that have been completed out of order must be held in renaming buffers prior to being retired. This means that operands to dependent instruction must then be retrieved from these registers and not the registers indicated by the instruction. This process requires tables for register renaming for results that have not been retired as well as memory for data-flow matching in order to determine which instructions can be executed. The prediction and the clean up logic and the additional registers and tables used in out of order execution mean that modern microprocessors are very complex. Less complex mechanisms exist for tolerating high latency into main memory, such as micro- or multi-threading [31,32].

In addition to these latency tolerant techniques, most processors attempt to issue more than one instruction in a single cycle by use of multiple execution units. This process is meant to increase throughput for a given clock cycle, again at the expense of complexity. Most recent microprocessors have at least 4-way issue but seldom achieve an effective instruction per cycle count (IPC) greater than 2. These general considerations hold for any computer with specific constraints depending on the number of processors and on the communication network of the machine.

C. Evolution or revolution of the architecture?

Two approaches prevail (with some possible intermediate points of view):

- The first one, which we call the evolutionary approach, consists of trying to integrate optical communication systems in forthcoming machines. This approach requires the analysis of communication bottlenecks in existing or future computers and the capability of Optical Communications to solve these problems 1) with much more effectiveness than electronic solutions and 2) with a good chance to reach a cost-effective mass production. Thus, this approach tries to predict the role of optical communications in the next 5-10 years starting with the present state.
- The second approach, which we tentatively call the mutational approach, considers that optics might induce new computer architectures with outperforming specifications that will justify abandoning (or at least dramatically modifying) existing electronic solutions. This is a more speculative approach about the possible long-term evolution of computer architectures. Notice that it does not release designers from having to know quite well the state of the art of existing electronic architectures that cannot be reduced to the pure communication aspects if they are to propose and demonstrate the advantages of the new optical solutions.

In the rest of the paper we follow the evolutionary approach as it is much less risky than the mutational approach and can be used as a reference for appreciating the relevance of more advanced proposals. Architectures are analyzed in the ascending order of the number N of connectable processors. Although hundreds of network topologies have been proposed, the number of commercial implementations is handful and mostly reduces to busses, rings, meshes, tori and central switches. Thus, we begin with the mono-processor. Then we consider SMP's (say typically $2 < N < 32–64$), rings (say, potentially $10 < N < 100$), and supercomputers that could connect up to several thousands of processors.

IV. Optical interconnects in monoprocessors

What could be the role of optical interconnects in monoprocessor machines? The registers are linked to L1 caches, the L1 to the L2 cache, the L2 to the memory or in some machines to L3 cache, and the L3 to the memory. The L1 and L2 caches, which are often integrated in the processor chip, are built with SRAMs and can operate at the processor frequency. Inserting OI's at this level (i.e., between the register and L1, or between L1 and L2) increases the transfer latency (due to the opto-electronic conversion time) and degrades the processor performance [33].

Perhaps, the integration of optical communications ought to be considered for the longest distances in the memory hierarchy, namely, between the last cache level (considered as L2 in the following) and the main memory. Two terms contribute to the MAL, namely:

- The *intrinsic* MAL, which is the leading term to the MAL, and which depends on the internal architecture and on the technology of DRAMs.
- The communication latency between L2 and the memory.

The need for memory bandwidth in future machines will grow dramatically owing to the increase of the processor operation frequency and to foreseeable architectural evolutions such as the extension of instruction-level parallelism of processors, the use of higher-order non-blocking caches, etc. Today, processors run around 700 MHz, and issue as many as 4 instructions per cycle. In the years 2005–2010, operation at nearly 4 GHz is expected with the execution of 32 instructions per cycle, corresponding to a sole instruction bandwidth B_0 (between the processor and L1) of the order of 400 Gbytes/s. The two levels of caches will substantially reduce the main memory accesses to a few percents of B_0, but all in all, a bandwidth in the range 10–20 Gytes/s is expected between L2 and the memory! Introducing OI's (operating around 1–2 GHz) might exhibit some advantages here (see the arguments listed in paragraph IIIA), but ultimately, the role of OI's will depend on the evolution of chip packaging and motherboard technology. The point is that no major problems is envisioned by electronic engineers in moving out to 1-GHz processors using electronic signalling in monoprocessor and tightly coupled architectures as explained below.

Neither bus frequency of the order of 500 MHz nor buswidth in the 500-pin range are seen as an insurmountable obstacle. The memory bus at this end of the market is not a significant issue as its cycle time is limited by DRAM speed. The effects of high latency can be mitigated by the provision of very wide electrical buses carrying up to 512 bits (plus error correction bits) of data simultaneously. Possibly, this broad bus might be split in several independent narrower busses (say 64 bits wide) to make access parallel to different memory banks. These solutions might require new chips with several thousands of pins but current mass-produced devices including from 2,500 to 10,000 pins are already available in research [34] with possibly a role for optical interconnects.

In conclusion, the possible introduction of optical interconnects (between the registers and L1, L1 and L2, or L2 and the main memory) seems hypothetical because: 1) The memory access time is dominated by the intrinsic memory latency and optical communications (whatever their bandwidth is) do not change this issue and 2) The bandwidth challenge between L2 and the memory, which is in the range 10 Gbyte/s, seems accessible to the forthcoming electronic packaging and to the motherboard technology. Electronic solutions will likely suffice for building cheap and efficient mono-processor machines in the next ten years.

V. Optical interconnects in multiprocessor architectures

A. General remarks

The logical way to reach a performance level not accessible with a mono-processor consists of connecting several processors through an interconnection network, thus building a multiprocessor system. As stressed in Section IV, the MAL is a critical parameter for the performance of the machine. In multiprocessors, it can increase drastically and can be as much as three orders of magnitude larger than the processor cycle time in the worst case, in particular when the memory is distributed in a number of processors (or clusters of processors) and when the execution of the application necessitates numerous internode transfers. It is possible to distinguish (at least) five contributions to the latency, namely:

1) The *intrinsic memory latency* already mentioned in section IV.

2) The *software latency* (communication overhead) associated with formatting, sending, and receiving messages. It is not clear at this time how optics can reduce software latency. A comprehensive study of the effect of the large communication bandwidth capability on the overall communication latency has not been done. Are there possible architectural innovations in inter-processor communications with optical interconnects that will eliminate or largely diminish the effect of software latency?

3) The (network) *propagation* latency, which depends on the network topology and the processing overhead for routing and solving contention problems. This latency is extensively discussed below.

4) The (network) *contention* latency, which critically depends on network saturation. The network latency is the sum of the propagation and the contention latencies.

5) The *coherence latency*. in which maintaining the coherence of the caches in tightly bound machines requires broadcasting (or multicasting) coherence messages through the communication network. This coherence maintenance contributes to slowing down the memory access. This factor strongly depends on the network topology. Snooping protocols, usually implemented with buses, are much simpler and faster than directory-based protocols that have to be implemented in distributed networks [30].

Which type of OI's (topologies, technologies, packaging schemes, etc.) is the most suitable in the short-term as well as in the long-term? First, it is clear that the network latency is the dominant term in existing multiprocessors (in a mono-processor, this is the intrinsic memory latency). Second, the network latency in strongly coupled multiprocessors is dominated by the bypass time of electronic nodes and *not* by the inter-node-propagation time (IPT). This is a key difference from the telecommunications networks because of the short inter-node distance in the systems under consideration (i.e., a few tens of cms, see paragraph IID) with an IPT in the range of a few ns. Optics can provide high connectivity, but increasing the connectivity generates new issues and must be used parsimoniously, as discussed below:

- When the connectivity is low (1 for the unidirectional ring shown in figure 2a), the node processing is simple (only add/drop of information in the network, possibly error detection and correction, priority treatment) but still longer than the IPT. The drawback of unidirectional rings is that the number of nodes a message must pass through in its round trip to remote memory (RTRM) equals the total number of nodes.
- Increasing the connectivity is therefore particularly attractive (using meshes, tori, hypercubes, etc.., see figure 2b) because the average internode distance decreases. Unfortunately, two new issues arise:
  1) The average internode distance generally decreases sub-linearly versus the connectivity whereas the network complexity increases linearly. Thus, increasing the connectivity becomes soon or late prohibitively expensive. Let us consider an example for clarity, namely, the connection of N=256 nodes with k-dimension bidirectional tori. The average internode distance is $D(k) \approx \frac{k}{4} N^{1/k}$ and the number of connection is $N_C(k) = N.k$. For k=2 (2D mesh), k=3 (3D torus) and k=4, we get *D(2)≈8*, *D(3)≈4.8* and *D(4)≈*4 respectively. Therefore, from k=2 to k=3, the average distance D is divided by 1.66, and from k=3 to k=4 by 1.2. Now, simultaneously, the number of connections is multiplied by 1.5 and 1.33, showing that increasing the connectivity becomes more and more costly.

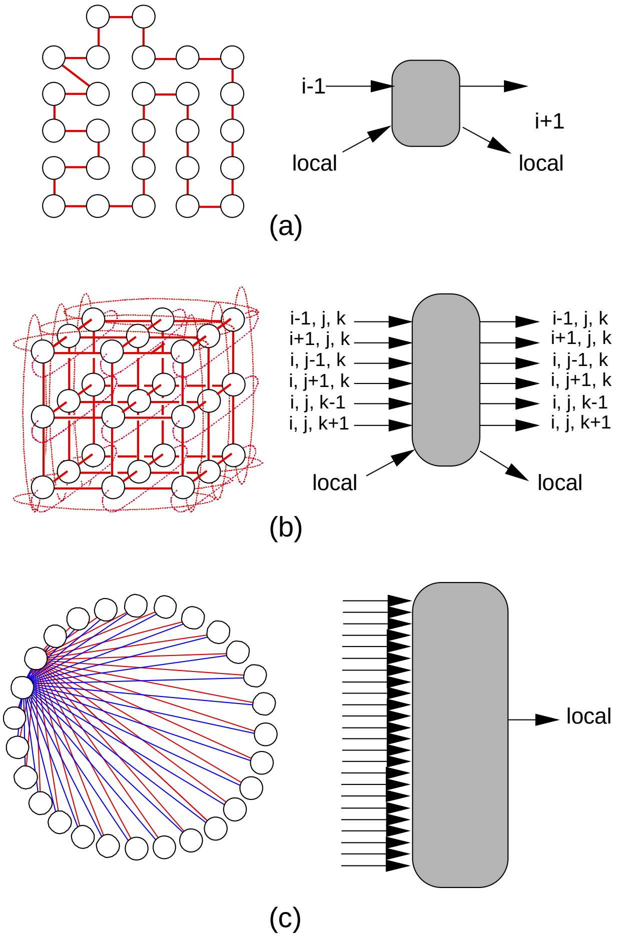

***Figure 2:** Three networks for connecting N = 27 nodes. (a) Unidirectional ring that requires to bypass all the nodes (here 27) in a Round Trip to read a Remote Memory (RTRM). The node is a 2x2 switch; (b) 3D torus with an average RTRM of about 4. The second network is about 7 times faster in terms of hop number but the node becomes a 7x7 switch. The increase of the node bypass time critically depends on the switch design; (c) Fully interconnected network. For clarity, the connections of only two nodes are drawn. The RTRM reduces to 1 hop, but the input structure of each node becomes a N-to-1 multiplexer that must operate in the asynchronous mode to reduce the MAL.*

2) The node bypass latency *T(k)* increases with the connectivity due to the increase of the node complexity. The increase of the *T(k)* critically depends on how the implementation of the node switch can be, depending on the operation frequency, the parallelism of transmissions, the communication protocol, the routing algorithm, the admissible cost, etc. A simple solution consists of decomposing a k-dimension switch in k successive 1-dimension switches optimized for straight traffic. In that case, the average internode latency *L(k)* of bidirectional tori scales as $L(k) \approx D(k) + (k-1) = \frac{k}{4} N^{1/k} + (k-1)$, where *D(k)* is the average number of bypassed nodes. The second term accounts for the increase of the note latency. Again with N=256, and k=2, 3, 4 we deduce *L(2)≈9*, *L(3)≈6.8* and *L(4)≈7*, showing that increasing the connectivity from 3 to 4 does not shorten the average internode latency. The conclusion here is that the topological arguments on the benefice of increasing the connectivity in strongly coupled multiprocessor networks cannot be separated from analyzing the complexity of the node routing implementation.

– If the network is fully interconnected (figure 2c), inter-node distances are minimal. However, this network topology transfers most of the latency reduction challenge to the design of the node-input circuits that must multiplex N asynchronous inputs and serialize the access to the node memory. Preserving the memory ordering constraints also seems a particularly complicated issue for a fully connected shared-memory machine. Perhaps, the input block of each node of a fully-connected network might be seen as common bus shared by the N inputs with a snooping protocol to maintain the coherence.

Therefore, finding the connectivity that provides the lowest network latency (i.e., the propagation latency in our classification at the beginning of this subsection) for a given number of N nodes is an open and complicated problem [35]. When N is very large (say several thousands as in a supercomputer), the hypercube topology may be attractive as the connectivity scales as $\log_2 N$ (see Refs. [36, 37] for comparisons with other topologies). But this solution is obviously expensive. For instance, with N=1024, twelve links per node chip are necessary, i.e., $\log_2(P+2)$ links/chip, requiring approximately 800 pins for a transmission parallelism of 64. For smaller machines, say for N less that 64 or 128, simplicity may be favored with the use of shared buses or rings. The latencies of the memory or the network are purely electronic. Optical interconnects can contribute to annihilate the contention latency by providing a huge bandwidth. This can be decisive for the machines based on 1-D interconnection networks such as SMP's or rings (see next Subsections V.B and CV.).

The scalability of the network parameters is hardly predictable in the long term. It is clear that the operation frequency of electronic nodes will increase whereas the IPT is incompressible. Thus, the IPT might become the leading contributor to the remote access latency, with a corresponding increase in the memory access time (in terms of electronic cycle) with strictly no hope of reduction. This long-term evolution would transform any strongly coupled multiprocessor into a weakly coupled machine. However, this situation is not inexorable as it is based on the sole scalability of the operation frequency of processors and nodes. It is likely that it will be accompanied by a reduction of dimensions (a possible metric being for instance the processor power divided by the processor volume) so that the processor cycle and the inter-node distance would diminish simultaneously, maintaining the pre-eminence of the node bypass latency over the IPT.

B. Symmetric Multiprocessors (SMP)

A SMP is a physically-shared-memory machine with a uniform memory access time. Early machines comprised a small number of processors, e.g. not exceeding 32, connected to a memory system via a single multiplexed shared bus [38]. The architecture has evolved and the number of nodes is larger today than expected only a few years ago with as many as 64 processors for the Sun Microsystems Model Enterprise 10,000 [39].

Interconnect links for data and addresses have been separated in modern machines (see Fig. 3). Today, the solution, for increasing the *data* bandwidth, consists of connecting each processor by a private link to a crossbar, and then connecting the crossbar to the memory. But the necessity of preserving the coherence of cache makes this method unsuitable for the addresses. Therefore, the shared address bus has become a critical communication bottleneck of the SMP architecture, because it serialises the accesses to the memory and adds an important contention latency to the MAL. The greater the number of processors, the longer the contention latency. The solution, which would consist of increasing the bus operation frequency, is particularly complicated, as the bus is a multi-point electronic line.

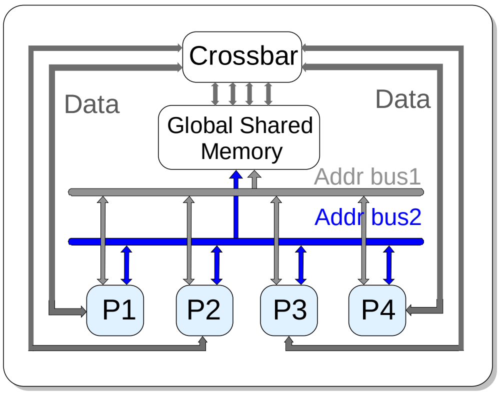

Figure 3: *Modern symmetric multiprocessor architecture. In this example, two address busses access the memory while preserving the coherence of caches.*

The only palliative solution has so far consisted of duplicating the number of address busses to reach the needed bandwidth, each of them being attached to an address memory range. In addition, bandwidth can be increased by the use of the notion that a logical bus can be implemented with a tree structure. This approach seems impractical for large SMPs (i.e., for SMPs with 64, 128 or more processors) and makes attractive optical solutions. The simplest technological change might consist of replacing each electrical *address* bus by an optical bus connecting the processors to several interleaved memory banks. This approach is attractive for three major reasons:

- Bus operation to as high as 1 GHz (or higher) would become possible (by replacement of electrical bus operating around a few tens of MHz) because the transmission of optical pulses in guides is not penalised by capacity effects and critical load adaptations encountered for electrical transmissions in a multipoint line. As a result, the SMP architecture (i.e., the processor, the bus and the memory) would become more scalable.
- Parallel transmission through optical lines is almost skew free in the GHz domain for transmission over a few tens of centimeters. This simplifies the data recovery in case of parallel transmission.
- The introduction of such an optical connection basically would not change the bus operation that would always rely on the access serialisation and a snooping protocol to maintain the coherence of caches.

However, several severe limitations cannot be ignored, namely

- The top transaction efficiency of a shared bus is close to 1 transaction/bus cycle with pipelined arbitration. This limit becomes a bottleneck for large SMP's with 32 or more processors [40]. Speeding the bus will surely improve the machines operation but will not solve all the contention problems of large multiprocessors, as it seems unrealistic to assume that the bus might operate faster than the processor. Large SMPs with several optical busses seems inevitable.
- The bus operation frequency cannot be increased without ensuring that each cache controller is able to check and update its directory at the bus operation frequency.
- Speeding up the bus will sooner or later generate an integration issue, as the bus length ought to be limited to make sure that the optical signals can be stationary within a single bus period. The light propagation velocity (c = 20 cm/ns) requires the bus to be shorter than 10 cm at 2 GHz, shorter than 5 cm at 4 GHz, etc. This size constraint disappears if more than one transaction can be simultaneously in progress in the bus. In that case, the bus architecture is akin to that of rings described in the next paragraph.

C. Rings and hierarchical rings

Ring multiprocessors are distributed-memory machines with non-uniform memory access time [41]. The ring is a *multiaccess* interconnection topology attractive for the following reasons:

(1) It allows usage of simple interfaces, because the ring connects to a given node by means of only one input and one output port. The node-ring interface is basically a 2x2 switch. This simplicity reflects itself in a relatively low requirement for the number of connecting wires, which often corresponds directly to the number of pins on physical connectors. The number of connections is considerably smaller than in 2D or 3D network topologies (torus, mesh, hypercube…) where

there exists more than one direction for incoming and outgoing signals inducing a larger overhead for processing routing.

(2) It provides a natural broadcast and multicast mechanism. This feature can be exploited in the implementation of producer-driven prefetching of data, which can improve performance significantly. Unlike with the bus, it is not possible to order parallel messages between different pairs of node and for most implementations, flow control can violate the ordering constraints. Thus, the ring structure preserves a *partial* ordering of transmitted packets, which can be exploited for implementation of a cache coherence scheme [42].

(3) There are point-to-point connections between successive nodes that do not suffer from undesirable effects such as loading and signal reflections from multiple connectors, which plague electrical bus-based schemes and effectively reduce their feasibility to small sizes (see Subsection V.B). Therefore, signals can be transmitted on such links at high clock rates. Operation at 3-4 GHz with a parallelism of 256 will provide a huge bandwidth close to the terabit-per-second range.

The effective bandwidth is essentially determined by the transfer rates attainable at individual nodes, and it can be improved by increasing the clock frequency or by increasing the width of the transfer path. The bandwidth of a multiprocessor interconnection network can also be increased by means of a hierarchical structure, whereby a number of localised transfers can take place concurrently on several rings. For example, if several local rings are connected by means of a central ring, then the number of concurrent transfers that can be supported is much higher if the transfers are only between sources and destinations on the same local ring. Transfers that pass through the central ring take more time than local ring transfers, but they are in general shorter than they would be if all nodes were connected to a single long ring.

D. Supercomputers and distributed shared-memory systems

Top of the range supercomputers use more than 1000 processors. Although the memory may be distributed, each processor can access all memory in the system. The distributed-shared memory (DSM) architecture attempts to provide a single addressing space for the distributed-memory to enable the user to get transparent access to computational resources in scalable systems. One achieves this by hiding the remote-communication mechanisms from the application writer, thus preserving the programming ease and portability of shared-memory systems. Additionally, the scalability and cost-effectiveness of underlying distributed-memory systems are also inherited.

However, local memory is accessed very much faster than remote memory in which messages have to be exchanged across the network in order to fetch data. The nonuniformity is due not only to the network topology but also due to packaging technologies and will be substantially degraded by heavy traffic loads and congestion. Additionally, the reliance on locality and memory allocation requires heavy caching of memory in order to reduce remote references. Unfortunately, caching shared data introduces the problem of cache coherence, the solution of which relies significantly on the efficiency of the interconnection network. As was stressed in Section 5, designing low latency nodes is also a critical issue. The problem will get worse with the advances in the speed of current microprocessors in which the price/performance advantage of microprocessors is increasing. In order to build a scalable DSM multiprocessor that utilises state-of-the-art off-the-shelf microprocessors with GB/s connections to local memory, one must utilise technology that supports inter-processor connections at least in the GB/s range and average access times to shared data in the nanosecond range. Optical interconnects could be the only cost-effective technology for the internode communications.

VI. Optical Communications in Reconfigurable Architectures

In nonconventional architectures (such as custom computing), in the reconfigurable computing domain, increasing use is being made of arrays of field programmable gate arrays (FPGA). High interconnect density is critical because of the difficulty in finding (or the non-existence of) natural, weakly interconnected partitioning of the functions. Furthermore, the on-chip interconnect facilities are relatively slow due to their configurability. This is not a passing phase depending on integration level, but will persist as the density of chips grows continuously. Optical interconnect may have a role in FPGA arrays, essentially to speed them up, by the introduction of a new routing layer. The density of optical chip-to-chip I/O may be applicable here, even for very short distances, such as adjacent chips. If I/Os can be justified for this purpose, they may even be viable to replace the relatively slow electrical communications within a single chip. Implementing wide busses from the FPGA chip to reconfiguration memory to achieve rapid reconfiguration could be of benefit in non-conventional architectures.

VII. Dedicated Optoelectronic Processors

Dedicated processors are very different from general-purpose monoprocessors or multiprocessors because they generally do not execute stored programs. They are designed for a specific task, which is often related to the processing of optical information. The MAL extensively discussed in Sections III-VI is no longer a problem (as there is no memory or almost no exchange with the memory). Dedicated optoelectronic processors traditionally have an optoelectronic front-end and back-end interface. The optical data-streams impinge on photodetectors that convert the light intensity of beams into electronic signals amplified and processed electronically. The resulting processed data can be converted back into optical signals for further processing if necessary. The number of optical channels can range from tens to around 10,000 over a 1 cm ×1 cm chip area.

The communication bandwidth becomes a critical issue. For example, in a vision machine, a 1024×1024 correlation on a 1024×1024 image requires around 170 $\times 10^6$ multiply-and-accumulate operations, which correspond to 10^{10} operations/s at a video frame-rate of 30 frames/s. In the same way, a matrix multiplication of 1024×1024 requires around 1 billion multiply-and-accumulate operations, which correspond to $6 \cdot 10^{10}$ operations/s for the same frame rate. General-purpose electronic machines cannot cope with the input/output needs of such computationally intensive applications. 0ptically interconnected electronic chips have been shown to be the only technology so far capable of providing a match between computational intense applications [43]. There is a case therefore for dedicated optically-connected electronic information processing systems for applications which require large data bandwidth capability. Such applications range from image processing primitive operations (FIR, IIR, Fourier transformation, 2-D convolution and correlation, dot-product and dot-matrix multiplication), to switching fabrics in telecommunications.

Demonstrators, based on optically interconnected electronics deal with such functions. See for example, the FFT machine built at the University of North Carolina, Charlotte (USA) which has an I/O bandwidth of 29 Gbytes/s and calculates a 1024 FFT in a few microseconds [44], the bitonic sorter at Herriot-Watt University which sorts 1024 15-bit deep words within 10 microseconds [45].

VIII. Conclusion

The most critical issue in computer architectures (from the monoprocessor to large multiprocessor systems) is the access time to the main memory. This is the key problem that architectures must live with, regardless on whether they are optoelectronic. Although memory chips have become much denser (and therefore, much larger), they have not become significantly faster. Furthermore, the techniques that are currently used to realise large memory systems suitable for multiprocessors create coherency problems for which no simple, well-scaling electrical solutions are known. This situation further aggravates the latency properties of complex memory systems. It also makes the processor architecture complex as many techniques in the processor are specifically targeted at the problems of memory latency and bandwidth limitations, as well as the unpredictability of hierarchical memory systems. Thus, designing new *low-latency* memory chips is a critical challenge. The changes are open, possibly at the physical level (with the introduction of new materials, for instance superconductors) or at the architectural level regarding the organisation of the memory (for instance, with design of multiported memories), or alternatively by cooling down current memory chips around liquid Nitrogen temperature.

With respect to future mono-processor machines, the possible introduction of OI's in the memory hierarchy seems hypothetical as it is likely that the evolution of the electronic packaging and the motherboard technology will suffice to build efficient machines in the next ten years.

With respect to multiprocessor machines, the situation is more favorable but a major problem is the economic risk factor in introducing a new technology. This is a strategic rather than a technological issue, but unless optics can solve a major problem and provide a significantly better solution at a cheaper cost, no one is going to take the risk of an untried technology.

Introducing optics is conceivable in several types of multiprocessor machines. For instance, the supercomputer line has traditionally been the first to experiment with novel techniques because the economic risk may be acceptable in the manufacture of a supercomputer, for which performance is the primary issue. But one may also consider developing systems that exploit the affordable technology with the basic idea that the commercial success of multiprocessors will depend not only on their computational capability, but also on their cost/performance ratio. Successful products might be those that could allow configuration of a viable entry-level machine at a correspondingly low cost, which could then be expanded into a larger system merely by acquiring additional hardware modules of essentially the same kind. The multiprocessors related to the different strategies are reviewed below.

Symmetric Multiprocessor Machines

Perhaps the most significant area where optical technology may be applicable is in the upper limits of SMP's. Rings and buses are one-dimension network whose performance critically depends on the traffic bandwidth of the communication system. The huge bandwidth aids in reducing the traffic latency caused by contention access. In SMP, the number of nodes is much larger today than expected only a few years ago, and there would be great benefit in increasing it further, though the techniques required for bus-based implementation are already heroic. If optical technology could help extend SMP, even by a factor of two beyond the current maximum (64 processors for the Sun Microsystems Model Enterprise 10000), it would readily find application.

Rings

Optical implementation of a ring-structured backplane makes it possible to achieve highly parallel links with very large bandwidth (of the order of the terabits per second). Complete parallel transmissions (requiring the implementation of a parallelism as high a 600 to insert simultaneously several transactions in the ring) would enable to reach the huge bandwidth needed in forthcoming processors [40]. A massively parallel optical ring (i.e., several thousands of channels) could be divided in several concentric sub-rings to increase the bandwidth further. However, It must be stressed that the key to success will be 1) the efficiency of the optical/electronic interface, 2) the capacity to carry out node operations (add, drop, bypass, and possibly on-the -fly error corrections) in one or two ring cycles to compress as much as possible the different node latencies, and 3) the capacity to maintain the coherence of the local cache at the ring operation frequency. To be successful, the OI will have to show considerably enhanced performance in comparison with its electrical counterpart.

Supercomputers

0ne may consider a demonstrator of a highly parallel computer connected by a hypercube interconnection network, using wide fast busses. The one area where optics can have a major advantage over electronics is in the interconnection busses in a network based parallel computer. Here a high pin-out and high data rate are both required in order to minimise the latency of a network based memory access. One can imagine a router chip with 10 busses of O(500) bits, which would be impossible for electronic communication. This would require the collaboration of a parallel computer manufacturer who is experienced in the use of commodity microprocessor parts. The major issue would be to obtain microprocessor die in which the pin-out could be taken to flip-chip bonding sites for emitters and where optical inputs were also provided.

Reconfigurable-Network Machines

Architectures able to exploit the capability of optics to support static networks that can be reconfigured very quickly (in one cycle?) may find optics attractive. This is largely incompatible with modern, shared-memory systems, which require much asynchronous, variable-sized packet switching. One application suggested is the use of cache consistency protocols that use write-update, with special multicast configurations to exploit knowledge about sharing patterns. However, write-update protocols aggravate the memory ordering problems, and this application may depend on future directions in this as-yet unsolved problem.

Dedicated Optoelectronic Processors

There are still issues concerning the design, fabrication and test of such systems. On the design front the smartness of the pixel (i.e., the degree of complexity in relation with the number of electronic gates) that each optically interconnected element should possess in order to maximize the overall aggregate data-throughput rate is still under study. Useful figures of merit can include the power-weight product of the overall system, the power consumption density per Gbit/s and the throughput rate itself. The performance of such demonstrators is limited by either the available laser source power, by the electronic chip area or the heat removal capability of the system. The interfacing technology between the optoelectronic components and the VLSI circuitry is still immature with several options still under consideration: flip-chip bonding, epoxy glue, anisotropic bonding, monolithic integration etc. The optical hardware needs further miniaturisation and to be industrially compliant, especially in terms of alignment accuracy and connections to the outside world. The yield and test of such systems have not been studied in great detail and deserve more fundamental work.

Acknowledgements:

This work partly summarises the conclusions of the Workshop on Optical Communications and Computer Sciences (WOCCS) that was held in Toulouse (France) in March 1999. This workshop was sponsored by the European Commission in the framework of the MicroElectronic Advanced Research Initiative (Mel ARI). It is our pleasure to thank Pierpaolo Malinverni for his permanent support. We also wish to thank the participants to WOCCS for numerous fruitful and informal discussions, namely, Howard L. Davidson, Alfred Forchel, Gilles Jacquemond, Graham Jenkin, John D. Lambkin, Dominique Lavernier, Henk Neefs, Rami Melhem, Matthias Pez, and Roger Vounckx, Zvonko Vranesic. The position taken by the authors of this paper does not necessarily reflect the opinion all the members who attended the workshop.

References and Notes

[1] Proc. of Optics in Computing '98
Editors: Pierre Chavel, David.A.B. Miller, and Hugo Thienpont, SPIE Vol. 3490, (1998)

[2] Proc. of the 4th International Conference on Massively Parallel Processing Using Optical Interconnects, Montreal, June 22-24, (1997); edited by IEEE Computer Society Press (1998)

[3] D.A.B. Miller, and H.M. Ozasktas
Limit to the bit-rate capacity of electrical interconnects from the aspect ratio of System Architecture; Journal of Parallel and Distributed Computing. Special issue on Parallel Computing with Optical Interconnects, vol 41, pp 42-52, (1997)

[4] See the URL: http://www.profibus.com/

[5] See the URL: http://www.fieldbus.org/

[6] See the URL: http://www.can-cia.de/

[7] See the URL: http://www.interbus.com/

[8] S. Scott
*Synchronization and Communication in the T3E Multiprocessor;*Proc. of 7th International Conference on Architectural Support for Programming Languages and Operating Systems. October 1996, pp 26-36; Edited by ACM Press, NY, 1996
See also the URL: http://www.sgi.com/t3e/

[9] C.B. Stunkel, D.G. Shea, D.G. Grice, P.H. Hochschild, M. Tsao
The SP-1 High Performance Switch; Proc. Scalable High Performance Computing Conference May 1994, Knoxville, TN. pp 150-157; Edited by the IEEE Computer Society Press, 1995

[10] See the URL: http://www.ssd.intel.com/

[11] See the URL: http://www.sgi.com/origin/

[12] For a complete document on HIPPI 6400, see the URLs:
http://www.noc.lanl.gov/~jamesh/hippi64/ or
http://www.scizzl.com , or
http://www1.cern.ch/HSI/sci/sci.html;

[13] S. Scott, M. Vernon, J.R. Goodman.
Performance of the SCI Ring
Proc. 19th International Symposium On Computer Architecture May 1992, pp. 403-414; Edited by ACM Press, NY, (1992)

[14] N. Boden, D. Cohen, R.E. Felderman, A.E. Kulawik, C.L. Seitz, J.N. Seizovic, and Wen-King Su
Myrinet: A Gigabit-per-Second Local Area Network
IEEE Micro 15, vol 1, pp. 29-38, (1995)

[15] Technical Report: http//www.sun.com/datacenter/products

[16] See URL: http://www.rambus.com/html/documentation.html

[17] S. Tang, Ting Li, Feiming Li, L. Wu, M. Dubinovski, R. Wickman, and R.T. Chen
An 1-GHz clock signal distribution for multiprocessor super computers; Proc. of Inter. Conf. on Massively Parallel Processing using Optical Interconnects (MPPIO96), pp 186-191, (1996), Edited by IEEE Computer Society Press

[18] R.T. Chen, L. Wu, F. Li, S. Tang, M. Dubinovsky, J. Qi, C.L. Schow, J.C. Campbell, R. Wickman, B. Picor, M. Hibbs-Brenner, J. Bristow, Y.L. Liu, S.Rattan, and C. Nodding
Si CMOS process compatible guided-wave multi-Gbit/sec optical clock signal distribution system for Cray T-90 supercomputer
proc of Inter. Conf on Massively Parallel Processing Using Optical Interconnects (MPPOI97), pp. 10-24, 1997; Edited by IEEE Computer Society Press

[19] T. Szymanski, and H. Scott
Design of a Terabit free-space photonic Backplane for parallel computing; Proc. of Massively Parallel Processing using Optical Interconnects (MPPOI95); Edited by IEEE Computer Society Press, pp 16-27, (1995)

[20] YS. Liu, B. Robertson, G.C. Boisset, M.H. Ayliffe, R.Iyer, D.V. Plant
Design, implementation, and characterization of a hybrid optical interconnect for a four-stage free-space optical backplane demonstrator
Applied Optics, Vol.37, No.14, pp.2895-2914, (1998)

[21] G. Verschaffelt, R. Buczynski, P.Tuteleers, P.Vynck, V.Baukens, H.Ottevaere, C.Debaes, S.Kufner, M.Kufner, A.Hermanne, J.Genoe, D. Coppée, R. Vounckx, S. Borghs, I. Veretennicoff and H.Thienpont,
*Demonstration of a monolithic multi-channel module for multi-Gb/s intra-MCM optical interconnects;*Photonics Technology Letters, vol. 10, nr. 11, pp. 1629-1631, November 1998

[22] V. Baukens, G. Verschaffelt, P. Tuteleers, P. Vynck, H. Ottevaere, M. Kufner, S. Kufner, I. Veretennicoff, R. Bockstaele, A. Van Hove, B. Dhoedt, R. Baets and H. Thienpont,
Performance of optical multi-chip-module interconnects: Comparing guided-wave and free-space pathways ; Journal of the European Optical Society A – Special issue on optics in computing, vol. 1, nr. 2, pp.255-261, (1999)

[23] J.C. Rodier, P. Chavel, A. Dupret, E. Belhaire, P. Garda, D. Prevost, P. Lalanne,
Video-rate simulated annealing for stochastic artificial retinas;
Optics Communications, Vol.132, No.5-6, pp.427-431, (1996)

[24] J. Tanida, Y. Ichioka,
Programming of optical array logic : Image data processing
Applied Optics, Vol. 27, N. 14, pp. 2926-2939 (1988).

[25] A.M. Weiner
Femtosecond optical pulse shaping and processing
Prog. Quant. Electr., Vol. 19, pp. 161-237, (1995).

[26] D.A.B. Miller
Physical reasons for optical interconnection
International Journal of Optoelectronics, Vol 11, No 3, p155-168, (1997)

[27] G. Yayla, P. Marchand, and S. Esener
Energy and speed analysis of digital electrical and free-space optical interconnections, in O*ptical Interconnections and Parallel Processing: The Interface:* Edited by A. Ferreira and P. Berthome eds., Kluwer, chap. 3, (1997)

[28] H.Scott Hinton
*An introduction to Photonic Switching Fabrics,*1993, Plenum press

[29] O. Kibar, D.A. Van Blerkom, Chi Fan, and S. Esener,
Power Minimization and Technology Comparisons for Digital Free-Space Optoelectronic Interconnections
Journal of Lightwave Technology, Vol. 17, No 4, p.546-555, (1999)

[30] J.L. Hennesy et D.A. Patterson
Computer Architecture, a Quantitative Approach
editor: Morgan Kauffmann Publishers, second edition, (1996)

[31] A. Bolychevsky, C.R. Jesshope, and V.B. Muchnick
Dynamic scheduling in RISC architectures
IEE Proc . Comput. Digit. Tech., vol 143, pp309-317, September (1996)

[32] A. Iannucci
Multithreaded Computer Architecture - a summary of the state of the art; Kluwer Academic (Boston/London/Dordrecht), 400pp, (1994)

[33] H. Neefs, Pim Van Heuven, and J. Van Campenhout
Latency requirements of optical Interconnects at different memory hierarchy levels of a computer system
Proc. Optics in Computing Brugge 1998, Edited by SPIE vol 3490, pp 552-555, (1998)

[34] H. Davidson , private communication, March 11th 1999
SUN Microsytems, 901 San Antonio Road, MTV29-235, Palo Alto, California 94303

[35] J.H. Collet, L. Fesquet
Comparison of the Latency for an Optical bus and Several 2D electronic Topologies.
CD-ROM of 11th Int. Parallel Processing Symphosium (IPPS), Geneva April 1997; Address in the CD ROM: X:\workshps\wocs\collet.pdf or X:\workshps\wocs\collet.ps
Edited by IEEE Computer Society Press, published 1997

[36] K. Hwang
Advanced Computer Architecture: Parallelism, Scalability, Programmability
New York: Mc. Graw-Hill, 1993

[37] A. Louri, B. Weech, and C. Neocleous
A spanning Multichannel Linked Hypercube: A gradually scalable Optical Interconnection Network for Massively Parallel Processing
IEEE Trans. On Parallel and Distributed Systems, vol 9, pp 497-512, (1998)

[38] P. Sindhu, J.M. Frailong, J. Gastinel, M. Cekleov, L. Yuan, B. Gunning, and D. Curry.
XDBus: A High-Performance, consistent, Packet-Switched VLSI Bus
Digest of papers of Computer Conferences (CompCon), Spring Edited by IEEE Computer Society Press, pp 338-344, (1993)

[39] A. Charlesworth
Starfire: Extending the SMP Envelope
IEEE Micro, vol 1, pp 39-49, Jan/Fev (1998)

[40] W. Hlayel, D. Litaize, L. Fesquet, and J.H. Collet
Optical versus electronic bus for address-transactions in future SMP architectures
Proc. of Parallel Architecture and Compilation Techniques (PACT) 1998, Paris October 1998, page 22-29; Edited by IEEE Comp Society , (1998)

[41] Z.G. Vranesic, M. Stumm, D.M. Lewis and R. White,
Hector: A Hierarchically Structured Shared-Memory Multi processor
Computer, vol 24, pp 72-79. January (1991)

[42] La Barroso, M. Dubois
The Performance of Cache Coherent Ring-Based Multiprocessors
Technical Report: CENG-92-19. Department of Electrical Engineering-Systems, University of Southern California, November 1992.

[43] A.V. Krishnamoorthy, D.A.B. Miller
Scaling optoelectronic-VSLI circuits into the 21st century : a technology roadmap
IEEE J. of Selected Topics in Quantum Electronics, Vol. 2, N.1, pp. 55-76 (1996)

[44] R.G. Rozier, F.E. Kiamilev, A.V. Krishnamoorthy
Design and evaluation of a photonic FFT processor
J. Parallel and Distributed Computing, Vol. 41, pp. 131-136 (1997).

[45] M.P.Y. Desmulliez, F.A.P. Tooley, J.A.B. Dines, N.L. Grant, D.J. Goodwill, D. Baillie, B.S. Wherrett, P.M. Foulk, S. Ashcroft, and P. Black
,Perfect-shuffle interconnected bitonic sorter: optoelectronic design
Applied optics, Vol. 34, pp. 5077-5090 (1996).