Computing structures and optical interconnect: friends or foes?

Jan M. Van Campenhout

Department of Electronics and Information Systems, Ghent University
Sint-Pietersnieuwstraat 41, B-9000 Ghent, Belgium

ABSTRACT

Optical interconnect is claimed to have significant advantages over electrical interconnect due to the superior physical properties of photon propagation over electron propagation. These advantages translate to the well-known fundamentally different bandwidth/distance/cross-section relationships of optical interconnects as compared to their electrical counterparts.

Despite the validity of these basic results, too naive or straightforward an application of optical interconnects in electrical systems may not bring about the expected performance gains. In fact, the scientific literature contains several examples of proposed applications of optical interconnect where the real, quantifiable benefits are at best doubtful. Often, these examples are only intended to demonstrate the feasibility of a technological approach, not its great potential at the systems or application level. Typical flaws one encounters are (i) addressing the wrong problem: one tries to introduce optical interconnect at a location where the electrical interconnect is not the limiting factor; (ii) failing to realize the differences between interconnect and communications; and (iii) singling out one property of optical interconnect, while disregarding less beneficial properties such as latency, area, power dissipation or optical pathway cost.

Capitalizing on the intrinsic potential of optical interconnects in electrical systems requires a holistic approach, that should address the real issues in future electronic systems. To assess the true benefits of replacing an electrical interconnect by an optical one, it is imperative to take the systems context in which the interconnect is to be used into account. By means of examples taken from the computing area, we shall illustrate the profound impact of this context. We then propose a possible parameter space that is intended to capture part of the relevant context information at the link level, thus providing a badly required common domain of discourse for systems designers and optical component designers.

Keywords: Optical Interconnect, Interconnect Parameter Space, Interconnect Context

1. INTRODUCTION

Using optics in information processing has been a long-standing research goal. In successive waves or generations, attempts have been made at replacing electronic data processing and transport by a photonic or optical counterpart. Several authors[1–3] describe early attempts to build optical processing systems. Free-space Fourier-plane optical computers, ultra-high speed correlators, massively parallel intelligent optical pixel plane architectures are just a sample of the approaches that have been studied at great length.

Yet, today virtually all *processing* is still done electronically, using CMOS processors with internal electrical connections. Optical processing, if any, is still limited to niche applications, and there are no signs that this might change in the years to come. In contrast, quite a different story is to be told about telecommunications, the long-haul *transport* of information. Spurred by the vastly growing demands of wide-area computer networks, and steady evolution towards the integration of communications networks and computer networks, optical-fiber communications in the Gigabit-per-second range are commonplace. The use of wavelength division multiplexing techniques (WDM) are further pushing the bandwidth of a single glass-fiber interconnect into the Terabit-per-second range.

The main reason why there has been such breakthrough in communications but not in processing is related to the performance of the mainstream (electronic) technology. In the case of processing, the evolution of both CMOS technology and computer architecture have been no less than phenomenal. Moore's empirical Law, predicting a performance doubling every 18 months, has been validated over and over. Except perhaps for the most demanding niche applications, traditional technology has been able to keep up with the requirements of the applications. In fact,

Further author information: E-mail: Jan.VanCampenhout@rug.ac.be

roles often have been reversed: more than once, the technological evolution has *created* new demands. This situation was quite different in the telecommunications arena. The mainstream technology of copper interconnect rapidly became incapable of providing the enormous bandwidth increase required by long-distance computer communications. Deploying the very much superior bandwidth-distance product provided by photon propagation in single-mode glass fiber *was the only way* to keep up with demands. This situation has provided the market push for the strongly intensified development and deployment of long-haul optical interconnect technology we have been witnessing.

Technology projections for CMOS keep on forecasting a sustained growth in the capabilities and densities of transistors on the chip surface.[4] Very short-range interconnections, whose lengths scale with transistor scaling, will roughly keep up with the transistors, provided better conductors (using copper) and dielectrics are used (low-k dielectrics). However, as more transistors are put on a chip, systems are expected to become larger and more complex. Relatively more long interconnections will interconnect major system parts. In absolute terms, these interconnections will still be short, but their speed or bandwidth will not be able to keep up with the transistor evolution, essentially for the same reasons that plagued interconnections in telecommunications years before. The fundamental properties of electrical signal propagation in conductors with a finite conductivity impose an upper limit on the cross-sectional bandwidth density that goes down with the *square* of the distance covered.[5] The joint effect of the shrinkage of the interconnect cross-section and the steep increase in speed of operation put ever tighter upper bounds on the lengths of electrical interconnections. Whereas, in the recent past, such an upper bound was *not* the limiting factor in most systems, the projected evolution indicates that this situation will change. Initially (and there are already several examples of this), the limit will show up in inter-PCB or even inter-chip interconnects; in time also the centimeter-range *intra-chip* interconnects may be affected.

Consequently one expects a paradigm shift to take place similar to the one in telecommunications. In many laboratories all over the world, teams of skilled and creative physicists, optical, mechanical and electrical engineers are incessantly improving the technology essential to implement short-range, massively parallel interconnects. Optical devices, driver and receiver circuits, hybridization technologies, and optical pathways are being improved at a high pace, to the extent that centimeter-range optical links can now be demonstrated that feature *better* properties—at least on some counts—than those of comparable, electrical interconnects. The main advantage of optical interconnect is the high areal density achievable without running into problems caused by dissipation, high-frequency losses or signal interference that are so typical of high-performance electrical interconnect. The density advantage cross-cuts various levels of the traditional interconnect hierarchy. At the cabinet or inter-cabinet level, parallel optical interconnects have very small form factors, but at the same time provide high aggregate bandwidths and low signal skew. Essentially the same densities allow direct access into individual chips through optical area I/O. As was observed by several authors, the use of optical interconnect effectively eliminates several levels from the electrical interconnect hierarchy.[6]

Yet, the great breakthrough predicted and expected by so many during past years has not taken place so far. Detailed application studies in a real and relevant systems context that clearly and objectively demonstrate a significant advantage in the use of optical interconnect are scarce. And industry exhibits a remarkable inertia to effectively explore the usefulness of optical interconnects in their products.

This note attempts to provide some understanding, at least in part, of why the effective deployment of short-range optical interconnects in advanced electronic products is not catching on as quickly as one might expect. An explanation is sought in the complexity of matching the properties of optical interconnect with the demands of the systems in which they are going to replace electrical interconnects. In the special case of the long-haul interconnect typical of telecommunications systems, it turns out that there is an almost perfect match between the favorable properties of optical interconnect and the systems requirements. The favorable properties of optical interconnect far outweigh their potential drawbacks. In contrast, when looking at massively parallel, short-haul interconnect in a more general-purpose systems setting, a much more delicate balance appears to exist between advantages and drawbacks. Making this balance requires a clear insight in the requirements imposed on the optical interconnect; these requirements may strongly depend on the particular place in the system where the interconnect is located.

2. EVOLUTION OR REVOLUTION?

To narrow down the scope of our discussion, it is good to dwell a little on two vastly different approaches in using optics in computing. This difference is well captured by the paraphrase 'building a new system' versus 'building a system anew'. The first approach is a radical departure from existing practice. It starts from the particular *strengths*

of optics, and attempts at directly realizing a desired functionality using these strengths. This often leads to physical systems whose structure, appearance and properties are dictated by their optical nature. Such systems are vastly different from an electronic realization with the same functionality. Typical of this approach are the early attempts to build 'all-optical computers', and to a lesser extent the various generations of 'smart pixels' or even the so-called 'fire hose architectures'. In a way, even free-space approaches to optical interconnect are of the same variety, at least as far as their physical realization is concerned.

The second approach is of a much more evolutionary nature. It starts from *existing* (or, preferably, near-future) electrical realizations of a complex functionality, and it raises the question as to whether the introduction of optical interconnect could somehow improve its properties (performance, dissipation, size, cost), or even make it at all realizable.

It should be clear to the reader that these approaches are quite distinct. In particular, in terms of the boundary conditions a designer has to adhere to, the first approach provides significantly more freedom. The net result is that systems designed according to the revolutionary scenario promise huge performance gains, but will typically be limited to very narrow fields of application or functionality. The evolutionary approach, however, is subject to much tighter constraints. Naively replacing existing interconnects by optical ones may not provide a clear-cut advantage—the replacement may even be counterproductive. However, if done successfully, its impact is likely to be much greater than that of the first variety.

In this note we shall concentrate on the evolutionary approach. In particular we shall identify a set of relevant properties of short-range, parallel optical interconnects. These properties are not independent; in fact some of them are conflicting. Hence, a systems designer who wants to introduce an optical interconnect should make context-dependent choices. To that end, we make an attempt to identify a number of typical interconnect contexts one encounters in high-performance digital electronic systems. These contexts will be characterized using a relatively small number of parameters. It are these parameters that will aid a designer in deciding to introduce an optical interconnect and in determining its key properties.

3. PROPERTIES AND REQUIREMENTS OF PARALLEL OPTICAL INTERCONNECTS

The kind of optical interconnect we are considering here aims at covering distances ranging from a few centimeters up to a few meters. These interconnects are intended to interconnect chips (on an MCM or a PCB), boards, or even crates or cabinets. Generally, they consist of a 2-D array of many individual channels, and have a point-to-point topology, but single-channel or multi-point interconnects exist as well. All optical interconnects have conversion circuitry that converts signals at logic levels into the analog signal driving the optical source component (the *drivers*), and converting the weak electrical signal, provided by the optical sink back to standard logic levels (the *receivers*). The optical source component can be a modulator, using externally generated laser light, a vertically emitting laser diode (VCSEL), or a micro-cavity light emitting diode (MCLED). The first two sources use coherent light, allowing low numerical aperture *optical pathways* in free space or in single-mode optical fiber. MCLEDs produce incoherent light with a Lambertian radiation pattern, thus requiring optical pathways with a very high numerical aperture. This limits the choice to multi-mode optical fibers or imaging fiber bundles (in plastic or glass).

Even from this very crude enumeration of possible embodiments, it is clear that an enormous range of possibilities is available to the systems designer. Making a reasonable choice out of the blue is neither sensible nor possible, and a careful analysis is required. To make this possible, it is required to identify the major parameters relevant in a systems context. The following list of—often conflicting—key parameters has been identified in the course of the EC funded MEL-ARI OPTO long-term research action on micro-electronics and optical interconnect.[7,8] This list is not exhaustive, but we believe it to capture the most important properties.

Density. The areal density of the interconnect is an extremely important parameter in short-range inter-chip interconnect. Presently, typical pitches between channels are 250 μm or integer fractions thereof, although systems with pitches of 100 μm have also been reported. From a system's point of view, as high as possible a density is generally desirable: the density limitation is the one electrical interconnect is facing. However, from a technological point of view, the density can be limited by a number of factors: the size of driver or receiver circuitry, channel density in the optical pathway (limited by waveguide size or diffraction effects), heat sinking capabilities, etc. Furthermore

one has to realize that somehow light source arrays and detector arrays need to be aligned with an accuracy which, in any case, is a fraction of the pitch. Hence a higher density can lead to much more stringent requirements for the thermo-mechanical design which can have a strong impact on cost.

Interconnect topology. The possibility to realize *multipoint interconnections* could be a very valuable asset for short-range optical interconnect. Most electrical interconnections at the crate or cabinet level have a point-to-point topology; direct replacement of these by optical interconnects is feasible. This situation changes abruptly at the backplane, PCB, MCM, and chip levels. Studies of wiring statistics in real electrical systems indicate that the number of points on a net (the so-called *net degree*[9]) is predominantly 2, but higher values do occur in non-negligible fractions. Furthermore, in most systems a limited number of nets have a very high net degree. These nets typically belong to buses or clock distribution networks.

It is also desirable that an optical pathway be capable of realizing arbitrary, highly non-planar interconnection graphs at the board level, with as few restrictions as possible on the precise locations of the interconnection end points. Electrical interconnect is capable of doing this by means of multilayer PCB or MCM technology, however with known weaknesses such as density and bandwidth limitations, signal cross-talk, and dissipation.

Severals studies have addressed the clock distribution problem, both using free-space[10] and guide-wave approaches.[11] However, the realization of large numbers of multipoint nets with an *arbitrary* interconnect pattern within an optical interconnect does not yet seem to have an elegant and mature solution.

Aggregate signaling rate. From a system's point of view, the aggregate signaling rate (the sum of the signaling rates of all channels) indicates the maximal number of signal transitions that can be transferred through the interconnect in a given time span. A high aggregate signaling rate can be achieved by using a few ultra-high-speed channels, or by using many moderate-speed channels (as in massively parallel optical interconnect systems). It is clear that the signaling rate of an individual channel, considered in isolation from the other channels in the link, is not the parameter of utmost importance. It is more relevant to talk about the areal signaling rate density (given in baud/cm^2), which takes the channel pitch into account.

Latency and skew. The latency of an interconnection is the time it takes for a signaling transition to arrive at the receiving end. In most telecommunications applications, the interconnect latency is essentially due to the signal propagation in the optical fiber. As this time is incompressible, latency is irrelevant in long-haul interconnects. A totally different situation exists in meter- or centimeter-range interconnects. In particular in the shortest interconnections, the time-of-flight of the optical signal is definitely not the dominant factor in the total link latency. For example, in the total latency of a 1-m optical interconnect in the RWC-1 Massively Parallel Computer[12] is 20.2 ns, of which only 5 ns is due to the time of flight. The total latency is a complex convolution of several small contributions by the driver circuitry, the light emitting or modulating device, the actual time of flight, the detector, and the receiver. In particular the receiver delay is often rather dependent on the power level of the received optical signal, hence on the power dissipation of the entire link.

The interconnect latency of most short interconnections has to be low from a system's point of view. In synchronous circuits, the tolerable interconnect latency is only a fraction of the clock period, which leads to sub-nanosecond values in circuits operated at one Gigahertz. Only sophisticated techniques such as wave pipelining enable one to escape from this limitation.[13] As other factors than the pure time of flight affect optical link latency, it is amenable to optimization. The optimization of interconnect latency leads to other solutions than does the optimization of the signaling rate, e.g. in terms of buffers, or driver and receiver topology. Latency-sensitive systems such as synchronous systems do not allow for the use of complex encoding/decoding techniques, which are frequently used in long-distance interconnect to enhance its BER or insensitivity to cross-talk.

In parallel interconnects, differential channel latency leads to skew. In particular in electrical interconnects, individual channels of a link may affect each other's latency by mutual interference, hence leading to data-dependent skew. Although one of the great assets of optical interconnect is the virtual absence of interaction of the optical channels, both optical and electrical cross-talk could lead to data-dependent skew in the electrical parts of the link.

Clocked or unclocked links. The use of a clock signal (at least at the sender's side) allows the use of coded signals to carry information. Such coding significantly simplifies the receiver design, as signals can be made free of DC components. Some encodings (such as Manchester coding) allow the clock frequency and phase to be sent along with the information, at the expense of requiring a peak signaling rate which is *double* the information rate.

Reliable detection of a coded signal at the receiver's side requires the knowledge of the clock phase, which in the case of Manchester coding can be obtained from the signal itself. Decoding typically results in an additional latency of at least half a clock period. If the clock signal is available at the receiving end, more bandwidth-efficient encodings can be used, and receivers can be of the clocked sense amplifier variety which may have advantages over transimpedance amplifiers in terms of dissipation. Not using a clock and using the simplest encoding of all—NRZ—may lead to lower-latency links, however at the cost of a significantly more complex receiver and higher power dissipation.

These considerations can be applied either to individual channels, or to synchronously operated channels making up one link. In the latter case, if the clock signal has to be reconstructed locally, the data channels of the link can share the cost of doing so using a common clock channel. In synchronous electrical systems, clock signals are available at both ends, and data is transmitted synchronously across many interconnections. Electrical interconnections at that level do not take the presence of the clock into account; optical interconnects could do so to their advantage.

In modulator based systems, one can use very short, externally generated light pulses as a clock. This technique allows to eliminate most of the transmitter-related skew and jitter.[14]

Electrical noise immunity. Unlike telecommunications systems, parallel optical interconnects are to be *embedded* in an operational digital environment, where electrical disturbances such as power and ground noise can take on values that are quite alarming to *any* analog circuit—including sensitive receiving amplifiers. Receivers (and perhaps entire links) should be designed to be immune to this environmental noise, either by using differential circuit techniques with high common mode rejection, or by using clock phase information in synchronous environments.

Bit error rate (BER). The degree of parallelism and the absence of any error-correcting circuits in low-latency links implies that the raw BER requirements of the individual channels are extremely stringent. Typically, BERs smaller than 10^{-15} are expected.

Power dissipation and chip area. The desired high areal density results in severe power dissipation constraints, hence the dissipation of all link components has to be kept as low as possible. The maximal permissible power dissipation density is limited by the heat removal capacity of the system. Hence, the per-channel power dissipation sets a minimal channel pitch and therefore a minimal available chip area per channel. This pitch could eventually be limited by the optical pathway when the link dissipation is strongly reduced. On the other side, a maximal pitch is limited by economical constraints (cost of chip area).

Uniformity and temperature dependency. The limited space per channel does not allow for the integration of feedback systems to monitor and control the optical power levels. Therefore, the component properties (of optical devices and amplifier circuits alike) must be quite uniform over time (reliability) and space (uniformity), and under different operation conditions, in particular over a broad temperature range. The minimal uniformity level depends on the type of transmission: at the detection side, a decision is made by comparing the incoming signal with a reference value. Depending on the design of the link, this reference value could be set for the whole array, or for individuals links; and it could be fixed or adaptive. The first case with a fixed array-wide reference obviously results in more severe uniformity conditions. Links using differential transmission do not depend on external reference levels, and such links are therefore more robust to non-uniformity.

The above list of properties and requirements is obviously not covering every aspect of parallel optical links. Several other important properties, such as the wave length, the coherency of the used light, the nature of the optical pathway were not discussed, as they are more or less hidden from the point of view of the systems designer.

4. SYSTEMS ARCHITECTURAL CONSIDERATIONS

Let us now turn to the systems in which the introduction of parallel optical interconnects is being considered. As argued above, the properties required of the optical interconnect depend on the particular context it will be used in.

4.1. Some terminology

Such contexts can best be described by locating them in a system's *design space*. A design space is a framework that allows to describe systems from various points of view, at various levels of detail, and during various phases of their construction. There exist several possibilities to set up such a design space; the one used here is the well-known Gajski-Kuhn Y-chart.[15]

In this design space, an existing system can be viewed from three different viewpoints or *dimensions*: its *behavior*, its *logical structure*, and its *physical structure*. The behavior describes the relationship that the system implements between its inputs and outputs. Of course this involves the functional relationships between inputs and outputs; but also performance-related aspects such as speed, dissipation and reliability show up here. The logical structure shows how the behavior is realized. Block diagrams, schematics, net lists, interconnect topology show up here. And finally, the physical structure shows how the logical structure has been cast in a concrete shape and technology. Materials, geometrical dimensions, packaging are visible here. It should be noted immediately that properties acquired in different dimensions are *not* independent. The physical dimensions of an interconnect obviously affect electrical behavioral properties like capacitance, bandwidth, time of flight, and dissipation.

To allow the description of complex systems, the design space is endowed with a hierarchy. The levels conventionally used are the *systems level*, the *algorithmic level* (also called Processor/Memory/Switch level), the *register transfer (R/T) level*, the *logic level*, and finally the *circuit level*. Table 1 shows a representation of the design space, and indicates how commonly used terminology can be given a place.

Like any other representation, this design space is not perfect. One weakness is associated with the notion of hierarchy. As suggested in Table 1, corresponding hierarchical levels are used throughout the three dimensions. This works quite well between the first two dimensions, but the correspondence is much less clear with the third. The natural hierarchy in the physical-structure dimension is the *packaging hierarchy*. Any reader with some familiarity with systems design will readily observe that a critical level in this hierarchy, namely the chip boundary, can appear at various places in the behavioral hierarchy. This level obviously depends on the size and complexity of the system. If the system is a hand-held calculator, the chip boundary is located at the systems level. However, if the system is a network of workstations operated as a distributed computing facility, the chip boundary will probably be located at the R/T and algorithmic levels. Furthermore, the technological evolution in CMOS technology increases the integration densities, and thus in time raises the level in the behavioral hierarchy where the chip boundary is located.

Table 1. The design space and its hierarchy

Hierarchical level	Behavior	Logical Structure	Physical structure
Systems level	Communicating processes	Multiprocessor, computer network	LAN, cabinets, ASICs
Algorithmic level, PMS level	Sequential algorithms	Processor, memory, peripherals	Boards, processor chip, back-planes, ASICs
R/T level	Machine code, micro-code, register transfer specification	ALU, controller, data path, registers	Macro-cells, FPGA macros, MSI/LSI logic
Logic level	Finite state machines, Boolean functions,	Gate level diagram, flip-flops	Standard cells, FPGA primitives, SSI logic
Circuit level	Differential equations	Transistor-level diagrams	Rectangles, discrete devices

The usual viewpoint taken when discussing interconnect issues is the physical-structure view. It is the natural domain to consider the geometrical properties (area, density, volume, length) of an interconnect, in particular for free-space optics. However, important behavioral parameters of the interconnect, such as latency, do not show up in the

physical dimension, but rather in the behavioral dimension. The rather fuzzy relationship between the hierarchical level in both dimensions often obscures the impact of introducing parallel optical interconnect, as will be illustrated in the following sections.

4.2. A key property: latency

Latency is an important behavioral parameter in digital systems. At all hierarchical levels, this term denotes the time elapsed between the presence of information at a given point, and the availability of the result of an operation on this information, possibly at another point. If the operation just involves transporting the information from one place to another, we are talking about *communication latency*. Depending on the hierarchical level, this quantity may contain several components that are not directly related to the *interconnect latency* defined in the previous sections.

At the systems level, the domain of asynchronous communicating processes, communication between system parts typically involves synchronization protocols, and may use a shared communications infrastructure (a network, a crossbar, a bus, ...). When the distances covered remain of the order of meters or less, the total time involved is mainly due to protocol overhead, network contention, or format conversion (e.g. series-to-parallel conversion). Although the aggregate *bandwidth* of the interconnect is important, the time scales involved are such that short-range interconnect latency is not. Furthermore, at that level systems are designed to be latency-tolerant to a large extent, e.g. by overlapping communication and computation, by using pipelining techniques, etc. The introduction of parallel optical interconnects may reduce this latency by virtue of their parallel nature, but this could be only of little help in those systems that are already very latency tolerant. The main benefit of introducing optical interconnect should then not be sought in pure performance. However, in systems where the systems-level communication latency is high and not due to pure interconnect latency, the introduction of parallel optical interconnects most likely will increase performance.

At the much lower R/T and logic levels, systems are invariably conceived of as locally or globally synchronous state machines, running at very high clock rates. Interconnects at those levels communicate either words or individual bits between locations that are millimeters to centimeters apart, without any intervening transformations or even synchronization protocols. The traditional notion of bandwidth of the interconnect loses most of its usual meaning, and the critical parameter is the interconnect latency, which is now to be compared to the clock period. So long as the introduction of parallel optical interconnect does not lead to an increase of the clock period, its parallel nature may increase the parallelism in the system and hence its performance.

Although the packaging and the behavioral hierarchies are only loosely related, in any given system the lower the level we are at in the packaging hierarchy, the lower the level we will be in the behavioral hierarchy, and the more important the latency parameter becomes. The deployment of massive numbers of short-range inter-chip interconnects in systems as they are built today, unavoidably will lead to these links emerging at relatively low behavioral levels.

4.3. A few examples from computers and other systems

To illustrate the relevance of the above line of thought, let us first consider interconnections in computer systems, starting with monoprocessor systems.

At the PMS level, the building blocks are coarse, namely the processor, the main memory, etc. Interconnections visible at that level are the external connections between processor and memory. The unit of communication is formed by cache lines consisting of a few tens of bytes (hence a bundle of wires), and the time scale involved is of the order of 50-100 ns. This interaction is not very frequent but rather bursty, requiring a high peak bandwidth. A clock is available at the processor side (clocked communication is possible). These characteristics of the interconnect logically lead to a clocked bundle (because the unit of exchange is several words) of wires where peak aggregate bandwidth is important (because of burstiness). The interconnect latency of a few nanoseconds (on a time scale of 100 nanoseconds) is clearly unimportant. The possibility of a clocked bundle of wires creates many possibilities in choosing the kind of optical interconnect. Furthermore, the allowable latency of a few nanoseconds allows the use of error correction and synchronization, and the use of synchronization in turn allows additional flexibility during the design, in that the latency of the interconnect (including the time-of-flight) should not be known at the time of design. Simulations[16] indicate the some performance may be gained by the introduction of parallel optical interconnect, but

that latency-fighting memory fetching techniques are far more effective. The bandwidth requirements between level-two cache and main memory, which will be in the range of several tens of Gbytes/s in ten years from now, appear to be in reach of the forthcoming electronic packaging and the motherboard technology. Electronic solutions will likely suffice for building cheap and efficient mono-processor machines in the next ten years,[17] and the introduction of optical interconnect may not be warranted solely on the basis of performance.

If, on the other hand, an optical interconnect is introduced deeply embedded in the processor, e.g., inside the logic circuitry of a functional unit, things look totally different. We now are at the logic level, where the unit of communication is a bit, no clock may be available at the interconnect and time scales of the communication involved are of the order of a nanosecond or even smaller. So, an unclocked asynchronous optical interconnect with an extremely low latency and high signaling rate is required. The required low latency prevents the introduction of error correction and synchronization techniques. The lack of the former requires an ultra-low BER (10^{-16}), while the lack of the latter requires knowledge of the worst-case interconnect latency at design time. The insertion of an extra clock period of latency at such low levels in advanced super scalar processor architectures may lead to a performance decrease of up to 30%.[16]

The case for optical interconnect is more clear in the context of multiprocessor machines. The total access latency to remote data can increase drastically with respect to the mono-processor case, and can be as much as three orders of magnitude larger than the processor cycle time in the worst case. This remote access latency consists of two contributions: the inter-node (packet) latency (the time between the start of transmission and the end of reception of a data unit, e.g., a cache line) and the communications-related intra-node latency resulting from routing, resolving contentions, controlling flow, resolving deadlocks, detecting and correcting errors, etc. The actual interconnect latency often is negligible. Increasing the parallelism in the interconnect (e.g., to 512 bits, a typical cache line) could contribute to reducing both the packet transport and contention latencies in multiprocessor networks, *provided the electrical node processing time keeps at par*. This increase in interconnect parallelism can be realized by means of parallel optical interconnect. It would require VLSI chips with a very high pin-out (typically from 5,000 to 10,000), requiring a high integration density of input/outputs already proven with the matrix of emitters or CMOS compatible detectors.[17]

An interesting optical interconnect application at the logic level concerns FPGAs or Field-Programmable Gate Arrays. FPGAs are generic components that are used in a variety of designs. In many cases, FPGAs are gaining in popularity when compared to ASICs (Application Specific Integrated Circuits), because of their programmability and the resulting shorter time to market. FPGA-based implementations of complex problems consist of multiple FPGA chips, hence the interconnection bottleneck of inter-chip interconnections appears. Pin limitations and electronic bandwidth limitations impede the performance of the application.[18] Even in modern mono-FPGA applications, the electrical interconnection of the FPGA with the rest of the system is becoming a point of concern. Here optical interconnects may come to our help; they provide the required interconnect densities at the required signaling rates. Despite the low architectural level of application, which is often as low as the logic level, interconnect latency is not a major problem because the programmable electronic interconnections inside an FPGA have similar latencies.[19] Detailed simulations[20] based on timing-driven implementations of synchronous benchmarks on such *optoelectronic FPGAs* prove that real and sizable performance gains are possible, both in terms of achievable performance and routability of complex applications.

5. IDENTIFYING INTERCONNECT CONTEXTS: A PROPOSAL

The above discussion shows how important it is for the optical interconnect designer to recognize the systems context in which an optical interconnect is going to be introduced. The context determines the overall requirements to be imposed on the optical interconnect. Once these requirements are known, the optical interconnect designer still has plenty of design parameters to consider. For example, the light sources (e.g. LEDs or VCSELs) or modulators have to be chosen, the receiver circuit has to be chosen and designed, as well as the optical pathway (e.g. plastic optical fiber, image fiber, optical pathway block, free-space). At this stage other important parameters come into play like power dissipation, silicon area, robustness (including optical dynamic range, power noise generation and rejection), data encoding, yield, etc. Note that all the different possibilities in choosing and designing an optical link span an enormous optical interconnect design space to be explored.

The problem is that recognizing interconnect contexts and identifying their essential parameters is not at all easy for persons who are not experienced systems designers. In the current state of development of optical interconnects,

the designer of an optical link will typically be a specialist in optics, in electro-optical conversion, and possibly in the design of driver and receiver circuitry. But it unlikely that he or she will be an experienced systems designer. Likewise, very rarely will a systems designer be acquainted with the details of optical interconnect and the subtle trade-offs to be made when designing a link. A common domain of discourse is needed, and this we think can be realized through a systematic definition of the interconnect context. In what follows, a proposal is made of how this could be done.

If it is our sole purpose to characterize the context in which an interconnect is to be realized, the complete design space is far too rich; we should resort to a much leaner formalism. In fact, what we are looking for is a small set of attributes that is rich enough to distinguish essentially different contexts. Inspired by Gajski's formalism, we propose the set of attributes shown in Table 2.

Table 2. Attributes characterizing interconnect context

Attribute name		Attribute values
Structural attributes – *what the link looks like*		
Topological attributes		
	interconnect topology	point-to-point, multipoint
	multiplicity	single-wire, bundle (number of channels)
	dimensionality	2-D, 3-D
Metrical attributes		
	distance covered	length of interconnect
	areal or linear density	density number
Behavioral attributes – *how the link is used*		
Communication attributes		
	kind of synchronism	none (asynchronous)– local – global; clock rate
	directionality	half-duplex – full-duplex
	modulation	NRZ, Manchester, 4B/5B, ...
Reconfiguration attributes		
	reconfigurability	none – slow – dynamic

Once given a value, these attributes characterize an interconnect context. Such contexts (at least those pertaining to meaningful parameter combinations) are mostly very recognizable to system architects, as shown by a few examples in Table 3. In this table, a globally synchronous system was assumed, and the distance scale was arbitrarily set at 30 cm (the PCB level). Similar tables for locally synchronous or asynchronous situations are just variations on this theme. The examples discussed in the previous section can be seen to belong to quite different contexts.

A soon as such a recognizable context is given, values can be specified for individual link properties in a much more restricted and meaningful way, as the relative importance of key parameters of the link is largely determined by the context. The designer of the optical link now has much better defined boundary conditions for the trade-offs he inevitably has to make, without requiring him to be fully acquainted with the details of complex systems design.

A similar systematic approach will be needed to eventually integrate the use of optical interconnects into existing design scripts, as ECAD software is even more vulnerable to the combinatorial explosion of a parameter search space than is a human designer.

6. CONCLUSION

In this paper we have addressed the question why we think parallel, short-range optical interconnect has not yet been taken up on the expected scale by systems manufacturers. The similarities and differences with the telecom area, where such a take-up indeed has taken place with great success have been pointed out. Short-range parallel optical interconnects are complex subsystems and their proper design needs considerable multidisciplinary proficiencies.

Based on several published examples of proposed applications of optical interconnect, and based on the experience acquired in the context of the MEL-ARI OPTO program set up by the European Community, we suspect that the introduction of parallel short-range interconnect in traditional electronic systems needs considerable care in order to

Table 3. A set of interconnect contexts in globally synchronous systems

Topology, configuration	Distance	Multiplicity	Brief context description
Point-to-point, fixed	< 30 cm	single	Short interchip interconnect, single-wire, operating synchronously. Receiver is synchronized with transmitter by external means. Example: *Logic-level synchronous interconnect.*
		multiple	Short interchip interconnect, wire bundle, operating synchronously. Receiver is synchronized with transmitter by external means. R/T or PMS level synchronous interconnect. Example: *Processor-memory link. Processor-I/O adaptor link.*
	> 30 cm	single	Serial, single-wire intercabinet link. Clock signal sent along separately.
		multiple	Inter-cabinet interconnect, wire bundle, operating synchronously. Clock signal sent along separately in additional channel. Example: *Interprocessor interconnect.*
Multipoint, fixed	< 30 cm	single	Short-range single-wire interconnect with fanout. Receiver is synchronized with transmitter by external means. Example: *Logic-level synchronous interconnect.*
		multiple	Short-range parallel interconnect with fanout. Receiver is synchronized with transmitter by external means. R/T or PMS level synchronous interconnect. Example: *Address bus distribution. Memory data broadcast bus.*
	> 30 cm	single	Serial, single-wire intercabinet link with fanout. Clock signal sent along separately. Example: *Serial interprocessor interconnect.*
		multiple	Parallel inter-cabinet interconnect with fanout, operating synchronously. Clock signal sent along separately in additional channel. Example: *PMS-level Interprocessor parallel broadcasting bus.*
Multipoint, variable	< 30 cm	single	Short-range single-wire interconnect with reconfigurable interconnect pattern. Receiver is synchronized with transmitter by external means. Example: *Reconfigurable bit-serial interconnection network in globally synchronous environment.*
		multiple	Short-range wire bundle interconnect with reconfigurable interconnect pattern. Receiver is synchronized with transmitter by external means. Example: *Reconfigurable bit-parallel interconnection network in globally synchronous environment.*
	> 30 cm	single	Serial, single-wire intercabinet link with reconfigurable interconnect pattern. Clock signal sent along separately. Example: *Serial interprocessor interconnect as in Transputer networks or Connection Machine. Globally synchronous optically reconfigurable bit-serial crossbars.*
		multiple	Parallel intercabinet link with reconfigurable interconnect pattern. Clock signal sent along separately. Example: *Globally synchronous optically reconfigurable bit-parallel crossbars.*

be successful. To provide insight in the often complex trade-offs that have to be made, and to provide a common domain of discourse for systems designers and designers of optical links, a framework has been proposed that is based on the concept of interconnect context. This framework allows the designer to isolate the parameters relevant to optical link design from the far more general systems context.

ACKNOWLEDGMENTS

This paper presents results of the EC funded MEL-ARI OPTO Project 22641, OIIC, and of the interuniversity Attraction Poles Program IUAP IV-13 initiated by the Belgian State, Prime Minister's Service, Science Policy Office. The author is indebted to Roel Baets, Joni Dambre and Jo Depreitere for their valuable remarks.

REFERENCES

1. D. Feitelson, *Optical Computing. A survey for computer scientists*, The MIT Press, Cambridge, MA, 1988.
2. M. Murdocca, *A Digital Design Methodology for Optical Computing*, The MIT Press, Cambridge, MA, 1990.
3. H. Arsenault and Y. Sheng, "An introduction to optics in computers," in *Tutorial Texts in optical engineering*, D. O'Shea, ed., vol. TT-8, Spie Optical Engineering Press, Bellinghem, WA, 1992.
4. *The National Technology Roadmap for Semiconductors*, Semiconductor Industry Association, San Jose, CA, 1997.
5. D. Miller and H. Ozaktas, "Limit to the bit-rate capacity of electrical interconnects from the aspect ratio of system architecture," *Journal of Parallel and Distributed Computing* **4**, pp. 42–52, 1997.
6. M. W. Haney and M. P. Christensen, "Performance scaling comparison of free-space optical and electrical interconnect approaches," *Applied Optics* **37**(14), pp. 2886–2894, 1998.
7. *Technology Roadmap: Optoelectronic interconnects for integrated circuits*, Office for Official Publications of the European Communities, Luxemburg (ISBN 92-828-7762), 1999.
8. H. Neefs, ed., *Optoelectronic interconnects For integrated circuits. Achievements 1996 – 2000*, Office for Official Publications of the European Communities, Luxemburg, 2000.
9. D. Stroobandt and F. J. Kurdahi, "On the characterization of multi-point nets in electronic designs," in *Proceedings of the 8-th Great Lakes Symposium on VLSI*, M. A. Bayoumi and G. Jullien, eds., pp. 344–350, IEEE Computer Society Press, 1998.
10. A. Mule, S. M. Schultz, T. K. Gaylord, and J. D. Meindl, "A 10 GHz hybrid optical/electrical clock distribution network for gigascale integration," *Proc. 1999 IEEE LEOS annual Meeting Conference* **2**, pp. 627–628, IEEE, 1999.
11. Y. Li, J. Popelek, L. Wang, Y. Takiguchi, T. Wang, and K. Sum, "Clock delivery using laminated polymer fibre circuits," *Journal of Optics A: Pure and Applied Optics* **1**, pp. 239–243, 1999.
12. S. Nishimura, H. Inoue, H. Matsuoka, and T. Yokota, "Optical interconnection subsystem used in the RWC-1 massively parallel computer," *IEEE Journal of Selected Topics in Quantum Electronics* **5**(2), pp. 360–367, 1999.
13. S. Q. Zheng, K. Li, Y. Pan, and M. C. Pinotti, "New addressing schemes for pipelined optical buses," in *Proc. 6-th Int. Conf. on Parallel Interconnects (PI'99)*, M. Haney, R. Kostuk, C. Lund, and E. Schenfeld, eds., pp. 230–237, IEEE Computer Society, (Los Alamitos, CA), 1999.
14. G. A. Keeler, B. E. Nelson, D. Agarwal, and D. A. B. Miller, "Optical interconnects using short optical pulses," *Proc. 1999 IEEE LEOS annual Meeting Conference* **2**, pp. 830–831, IEEE, 1999.
15. D. Gajski and R. Kuhn, "New VLSI tools. Guest editor's introduction," *IEEE Computer* **6**(12), pp. 11–14, 1983.
16. H. Neefs, P. Van Heuven, and J. Van Campenhout, "Latency requirements of optical interconnects at different memory hierarchy levels of a computer system," in *Optics in Computing '98*, P. Chavel, D. A. B. Miller, and H. Thienpont, eds., *Proc. SPIE* **3490**, pp. 552–555, 1998.
17. J. Collet, D. Litaize, J. Van Campenhout, C. Jesshope, M. Desmulliez, H. Thienpont, J. Goodman, and A. Louri, "Architectural approach to the role of optics in monoprocessor and multiprocessor machines," *Applied Optics* **39**(5), pp. 671–682, 2000.
18. J. Depreitere, H. Van Marck, and J. Van Campenhout, "A quantitative analysis of the benefits of the use op area-I/O pads in FPGAs," *Microprocessors and Microsystems* **22**(2), pp. 145–154, 1997.
19. J. Van Campenhout, H. Van Marck, J. Depreitere, and J. Dambre, "Optoelectronic FPGA's," *IEEE Journal of Selected Topics in Quantum Electronics* **5**(2), pp. 306–315, 1998.
20. J. Dambre, H. Van Marck, and J. Van Campenhout, "Quantifying the impact of optical interconnect latency on the performance of optoelectronic FPGAs," in *Proc. 6-th Int. Conf. on Parallel Interconnects (PI'99)*, M. Haney, R. Kostuk, C. Lund, and E. Schenfeld, eds., pp. 91–97, IEEE Computer Society, (Los Alamitos, CA), 1999.