Limitations and Perspectives of Optically Switched Interconnects for Large-scale Data Processing and Storage Systems (Invited)

Slavisa Aleksic[1], Gerhard Schmid[1], and Naida Fehratovic[2]
[1]Institute of Telecommunications, Vienna University of Technology, Gusshausstrasse 25-29/389, 1040 Vienna, Austria
[2] Kapsch TrafficCom AG, Am Europlatz 2, 1120 Vienna, Austria

ABSTRACT

The ever-growing Internet data traffic leads to a continuously increasing demand in both capacity and performance of large-scale ICT systems such as high-capacity routers and switches, large data centers, and supercomputers. Complex and spatially distributed multirack systems comprising a large number of data processing and storage modules with high-speed interfaces have already become reality. A consequence of this trend is that internal interconnection systems also become large and complex. Interconnection distances, total required number of cables, and power consumption increase rapidly with the increase in capacity, which can cause limitations in scalability of the whole system. This paper addresses requirements and limitations of intrasystem interconnects for application in large-scale data processing and storage systems. Various point-to-point and optically switched interconnection options are reviewed with regard to their potential to achieve large scalability while reducing power consumption.

INTRODUCTION

Within the last few years, the traffic volume in global communication networks has been increasing by approximately 50% to 60% per year [1]. In order to keep track with this increasing demand for bandwidth, the capacity of underlying network components as well as data storage and managing elements has to increase too. Indeed, most of the state-of-the-art electronic core routers comprise a large number of switching modules and line cards to achieve large capacities in the order of several Terabit per second. For this reason, the most of the high-performance routers currently being developed are multi-rack systems. Since the internal interconnection network becomes more critical when increasing the number of components, line cards, and racks, the high-capacity network elements are often limited by the maximum achievable size of the internal interconnection network. Thus, a scalable interconnection system is a crucial prerequisite for implementing high-capacity, highly efficient, and scalable ICT Systems.

This paper reviews various options for scalable and efficient optical interconnects and presents some results on scalability and energy efficiency of optically switched interconnects. The structure of the paper is as follows. The subsequent section presents some current trends in high-capacity data processing and storage systems. Then, we discuss design considerations for highly scalable and efficient large-scale optical interconnects and review technologies for optical printed circuit boards (PCBs), optical backplanes, and optical rack-to-rack interconnects. In the last section of this paper, we consider various switch architectures and network topologies for the use in internal interconnection networks and present some results on scalability and energy efficiency of optically switched interconnects, which are able to provide both transmission and switching functionality directly in the optical domain.

CURRENT TRENDS

The internal interconnection network is usually limited by the maximum data rate per link and per cable, the required number of links, and the maximum length of a single interconnection link. This is due to the fact that current interconnection networks mostly consist of electronic backplane interconnects and optical point-to-point links that interconnect modules placed in different chassis and racks. In fact, very high data rates over long electronic backplane traces are hardly achievable due to associated high signal losses and inter-symbol interference (ISI). On the other hand, processing electronics show continuous advance in computational bandwidth as well as in reduction of its feature size, thus providing more functionality and higher speed on cards within modules, i.e., system boards. This implies more point-to-point interconnects on boards and between boards, which dancer packing is, however, limited by crosstalk. Similar issues are faced by interconnection networks for the use in data centers and high performance computing (HPC) systems. In fact, data centers are experiencing a heavy increase in the amount of traffic to store and process. Therefore, optical cables have been already used to interconnect racks of equipment within data centers and high performance computer clusters. Due to the fact that both optical transmission and switching technologies are generally able to provide higher data rates over longer transmission distances and faster switching operation than electrical transmission and switching systems, a natural answer to the scalability problem could be to use both optical transmission and switching technologies (i.e. optically switched interconnects) in order to relax the limitations and improve the scalability of internal interconnects.

According to [4], communication networks of today are responsible for about 2% of the global carbon emissions. Therefore, power consumption demands a lot of attention when considering scaling of the network and managing elements to future required aggregate capacities. Figure 1a) shows the estimated total power consumption of electronic routers for capacities up to 100 Tbps with some examples of current high-performance routers [5]. Power consumption has increased linearly with bandwidth and each new generation of high-capacity routers consumes more power than the previous one. State-of-the-art electronic core routers providing Tbps capacity consume about 8-11 kW per rack of equipment. For a total capacity of about 100 Tbps, the power to be supplied to such a system is about 1 MW. The internal interconnection network is responsible for about 19% of the total power consumed in a router [5]. A straightforward scaling of current routers to higher capacities would result in a very high total power consumption of order of several MW and a very large number of internal interconnection links. Additional to the increased number of links also their length increases due to the increased size of the entire system, which consequently leads to an increase in power consumption because more transceivers providing both high-bandwidth and large signal power are needed.

Figure 1b) represents some recent trends and projections for power consumption of high-performance computers (HPCs). It is evident that already today HPC systems consume about 10 MW of electricity. In the figure, two examples of recent high-performance computers are indicated. One of them is the currently most powerful HPC system (K-Computer), which reaches maximum performance of 10.5 PFLOP/s while consuming more than 12 MW of energy. This leads to an energy efficiency of 0.93 GFLOP/s/W. The second example is the most energy-efficient system (BlueGene/Q) providing about 2 GFLOP/s/W and a maximum performance of 0.34 PFLOP/s. As it can be seen from Figure 1b), an increase in performance is only possible

with an increase in both complexity and power consumption. Thus, future Exascale computer systems will probably consume more than 20 MW, which will set very high requirements on power supply and cooling systems. Therefore, large effort has to be put into research and development of more energy efficient structures and technologies in order to make possible further scaling in both capacity and performance.

Several comparisons show strong potential for reduction of the power consumption through replacing electrical lines by optical interconnects. Above a certain length, the break-even length, optical interconnects consume less energy than electrical ones. The break-even length is calculated in [3] to be 50 µm and in [6] to be some mm for on-chip interconnects. The basic physical differences between optical and electronic interconnections lead to potential benefits of optics over electronics in terms of interconnect density, timing, signal integrity, and power consumption. Thus, deeper penetration of optics into high-capacity network and data managing elements such as routers, switches, data centers, and high-performance computers could provide benefits regarding scalability and power consumption. Additionally, future viability of optical interconnects also depends on a reduction of costs per input/output port as well as on the achievable performance of an interconnection network. It is thus important to consider all the different factors when proposing and examining new concepts for highly scalable and efficient optical interconnects.

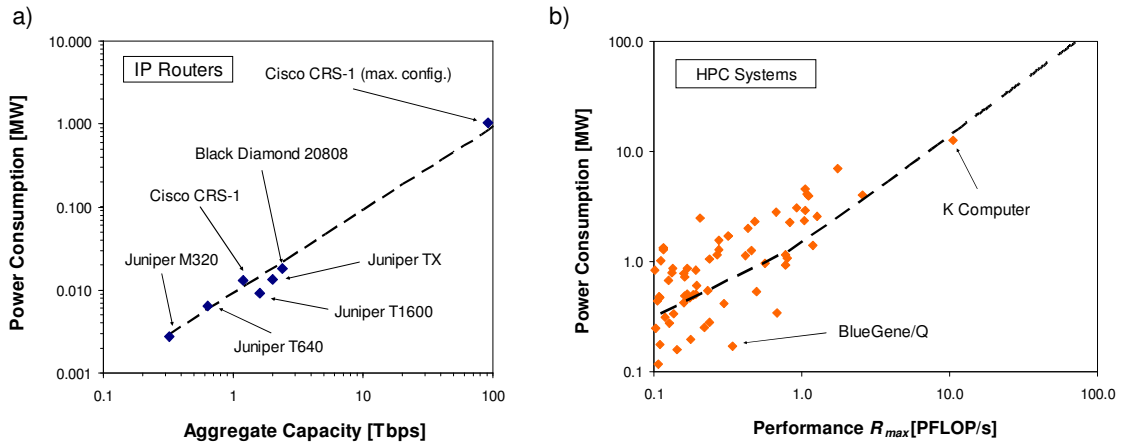

Figure 1. Trends and projections for power consumption of a) high-performance routers [5] and b) high-performance computers according to data taken from [7].

DESIGN CONSIDERATIONS FOR LARGE-SCALE OPTICAL INTERCONNECTS

Extensive research has been conducted on the relieving bandwidth bottlenecks related to current electronic point-to-point interconnects. Electrical interconnects suffer from high crosstalk as well as distance and frequency dependent attenuation due to dielectric losses coupled with high frequencies. Hence, very high data rates over copper-based cables and PCB traces are hardly achievable. In fact, recent high-capacity routers and switches, supercomputers, and data centers are increasingly relying on optical point-to-point interconnection links that typically support up to 10 Gbps data rate per channel. Standard optical point-to-point links are based on directly modulated 850 nm vertical cavity surface emitting lasers (VCSEL) and multimode fibers (MMFs). The capacity of such interconnects is limited both by modulation bandwidth of the laser and by the intermodal dispersion in MMFs. Recently, various methods have been proposed

and investigated in order to enable transmission at higher data rates: i) increasing the modulation rate of the laser [8], ii) using multiple fibers and transceivers in a parallel manner [9], iii) using advanced signal formats such as optical orthogonal frequency division multiplexing (OOFDM) and high-order modulation formats [10].

Current high-capacity routers and switches, data centers, and supercomputers comprise many electronic data processing nodes and electronic switches interconnected by a huge number of point-to-point, single-channel links. When considering future requirements, such architecture will not only lead to poor scalability and large latency, but will also cause low power efficiency and high implementation costs. On the other hand, optically switched interconnects that make use of optical switches and wavelength-division multiplexing (WDM) technology can benefit from inherent parallelism and optical transparency. Several realizations of optically switched interconnects based on different interconnecting arrangements and optical switching devices have been investigated in literature [11-15]. It has been shown that optically switched interconnects have the potential to realize high-performance and energy-efficient interconnects. However, their future viability depends also on a further improvement in scalability and reduction of cost per input/output port.

It is crucial to recognize that large internal interconnection systems are much more than a single point-to-point transmission link. Indeed, additional to the large number of transceivers and cables (waveguides), they also comprise elements that implement different other functions such as synchronization, switching, switch control, arbitration, signaling as well as managing and routing of data units through the internal interconnection network. All these additional elements contribute to an increased complexity and a higher energy consumption of the interconnection system. Also the architecture of the interconnection network influences the achievable performance and efficiency of the entire system. Therefore, new concepts for interconnection systems should be examined by considering additional to the transmission properties of point-to-point links also various other technological aspects and interconnection arrangements under all performance metrics, namely scalability/feasibility, traffic-related performance, power consumption and techno-economics as depicted in Figure 2.

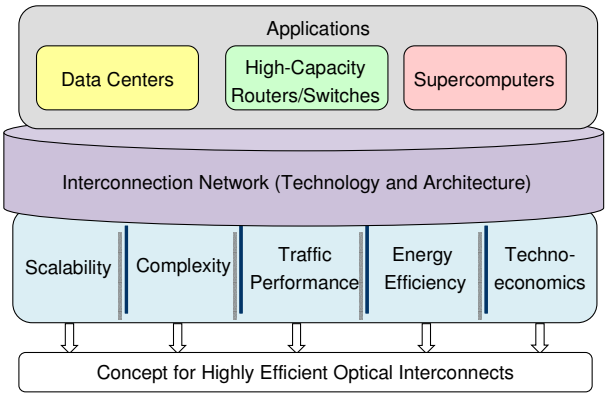

Figure 2. Conception and evaluation of highly efficient and scalable optical interconnects for large-scale systems.

The aim of this paper is to draw attention to the different aspects that have to be considered in evaluating the suitability of new interconnection concepts for large-scale data processing and storage systems.

OPTICAL PCBs AND BACKPLANES

Internal interconnection systems are usually classified into the following four groups: rack-to-rack (from 1 m to several hundreds of meters), backplane (board-to-board, typically between 15 cm and 1 m), chip-to-chip on a board (typically less than 25 cm) and on-chip (< 2 cm) interconnects. In this paper, we focus on interconnects ranging form chip-to-chip links on printed circuit boards (PCBs) to system level interconnects (rack-to-rack).

As the processing power of single silicon chips has constantly increased during the last three decades, the requirements on interconnects between chips and modules have increased too. According to the recent projections of the International Roadmap for Semiconductors (ITRS), single processor chips with 100 TFLOP/s can be expected in 2020. The interconnections between processing nodes should be able to provide capacities of more than 200 Tbps and energy efficiency significantly below 1 pJ per bit. All these requirements can hardly be met by electrical interconnection technologies; while optical interconnects have the potential to provide both high bandwidth density and high energy efficiency. Electrical interconnects suffer from distance and frequency dependant attenuation due to two kinds of losses coupled with high frequencies: i) dielectric loss in PCB substrates and ii) skin-effect in coaxial cables. Optical interconnects exhibit no frequency dependant loss. Thus, optical interconnects are capable of overcoming most of the physical limitations associated with electronic interconnects regarding interconnection density, timing, signal integrity, crosstalk, and energy consumption. In the following, we present and discuss optical interconnects on PCB boards on example of an innovative method based on two-photon absorption (TPA) to rapidly write multi-core optical waveguides within a polymer material, which can be coated on any standard printed circuit board.

Optical Circuit Boards

Optical interconnects within printed circuit boards (PCBs) allow denser waveguides, with potentially lower energy consumption per transmitted bit compared to pure electrical interconnects [3]. Further advantages of optical PCBs such as the robustness against electromagnetic interference [16] and the galvanic isolation make them a promising alternative to microstrip lines. The high number of interconnects in supercomputers and data centers require interconnects with very high reliability. One of the biggest challenges is to develop optical interconnects which satisfy this requirement [17]. Most of recently demonstrated optical PCBs are produced using the well-known photolithographic methods for waveguide production [18], but also embossing technologies are used for structuring the optical waveguides [19].

The two-photon absorption (TPA) process as used by AT&S allows a direct writing of waveguides within a few fabrication steps as illustrated in Figure 3 [20]. A GaAs PIN photodiode and a vertical cavity surface emitting laser (VCSEL) are mounted onto a conventional PCB and embedded into an inorganic-organic hybrid polymer (ORMOCER©). Before the process of waveguide writing is performed, the positions of these electro-optic elements are determined. The waveguides are written according to the actual positions of the light emitting area of the laser and the sensitive area of the photodiode. The transversal refractive index profile of one single core of the written waveguide may be approximated with a Gaussian profile and has a diameter of approximately 20 µm. The example in Figure 3 shows a multi-core waveguide structure, which increases the waveguide diameter and hence the tolerance to transversal misalignment between the waveguide and the electro-optic elements.

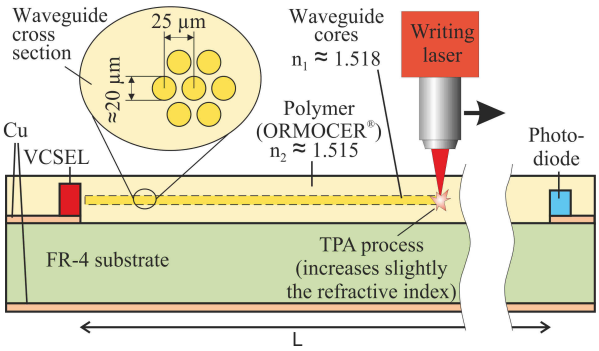

Figure 3. Direct writing of multi-core waveguides by the two-photon absorption (TPA) process.

Figure 4 shows the measured bit error ratio (BER) vs. data rate, R, of a single point-to-point interconnect at an opto-electronic circuit board prototype. The waveguide was written using the TPA process. The measurements were carried out at three ambient temperatures T_{a1} = 20°C, T_{a2} = 50°C and T_{a3} = 80°C. At T_{a1} = 20°C the board allowed transmission of signals with a rate of up to R = 5.5 Gbps with BER $\leq 10^{-9}$. Only at elevated temperatures (T_a = 80°C) the BER increased slightly. The test word was a pseudo-random bit sequence with the length $2^{31} - 1$. Note that the data rates presented in Figure 4 are obtained for a particular pair of VCSEL and PD and over a 12 cm long multi-core optical waveguide. In this experiment, the achievable data rate was mainly limited by the characteristics of the VCSEL , which is specified for data rates up to 5 Gbps. The purpose of this experiment was not to show the maximum possible data rate, but just to demonstrate the principle of operation. Generally, much higher data rates could be possible when using high-data-rate VCSELs and PDs. For more details about the presented approach for rapid fabrication of optical PCBs the reader is referred to [17].

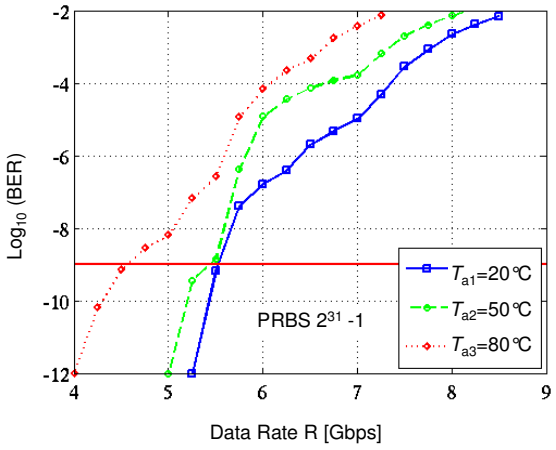

Figure 4. BER vs. date rate R at several ambient temperatures T_a of an opto-electronic circuit board with a TPA-written waveguide.

POINT-TO-POINT INTERCONNECTS

More functionality and processing power on boards leads to a need for high-speed and high-capacity interconnects between boards. Electrical interconnects are more vulnerable to cross-talk and the need for higher bandwidth is usually obstructed by increased dielectric losses.

In order to overcome this electronic bottleneck and meet growing demand for both high speed and bandwidth, various interest groups and standardization bodies have developed enhanced technologies for printed circuit boards as well as standards and protocols with improved signal integrity specifications, line coding formats, and design techniques such as pre-emphasis and equalization. High-speed electrical signaling has also experienced considerable enhancements. These technologies are in the process of being adopted by system and chip vendors.

Each interconnection technology is upper-bounded regarding its transmission distance, channel data rate, and number of assigned channels. Some of the technologies such as CEI-6G, CEI-11G and sRIO (Serial Rapid I/O) are only intended for electrical backplane applications, while others such as those based on Ethernet and InfiniBand (IB) also support board-to-board and rack-to-rack applications through using optical point-to-point interconnects. A lot of effort has been made by standardization bodies such as IEEE 802.3ba Ethernet Task Force, Optical Internetworking Forum (OIF), Fiber Channel (FC), InfiniBand (IB), Rapid IO Trade Association, and PCI Express to achieve an improvement in existing interconnecting technologies regarding data rate and efficiency. The IEEE 802.3ba Ethernet Task Force has already finished the work on the standardization of 40 Gbps (40GbE) and 100 Gbps Ethernet (100GbE). The Physical and Link Layer (PLL) Working Group of the OIF has developed the physical specifications CEI-25G and CEI-28G for achieving lane signalling rate of up to 28 Gbaud/s. They are intended for next generation chip-to-chip and chip-to-module as well as for backplane applications that support transmission up to 100 Gbps. InfiniBand has introduced additionally to single, double, and quadruple data rate (SDR, DDR, and QDR) also enhanced data rate (EDR) systems with 20 Gbps per lane. Similarly, the PCI-SIG group has also released the final specifications for PCI Express 3.0 with improved data rates. Rapid IO focuses on higher data rates which will come with the serial Rapid IO 3.0 specification. Although a lot of progress in signal processing and modulation has been made to realize and standardize high-data-rate electrical and optical interconnects, the most of the effort has been put into design and characterization of simple point-to point links, while the system perspective is mostly ignored. In the following section, we will address optically switched interconnects from both device and system perspective. The optically switched interconnects are able to provide both transmission and switching functionalities directly in the optical domain.

OPTICALY SWITCHED INTERCONNECTS

Combining optical transmission and optical switching in an optimal way to realize high-capacity optically switched interconnects could be a promising approach for high-performance systems. In the following subsection, we review different optical switching devices that can be used as an alternative to widely used electronic switches.

Optical Switching Elements

Optical switching technologies for the use in optical interconnection systems can be classified based upon underlying physical effect that is used for the switching process, i.e., into electro-optic (EO), acusto-optic (AO), thermo-optic (TO), and opto-mechanical (OM) switching. The EO, AO and TO effects rely on refractive index changes of the matter through application of an external physical field or action, while in OM switches, optical beams are reflected by electro-mechanical means.

In the EO effect, an applied electrical field induces change in the index of refraction, which then channels light to the appropriate port. $LiNbO_3$ is a unique crystal that shows large EO effect, AO effect, TO effect and nonlinear effects. EO devices based on this substrate have very fast response, small dielectric constant and relatively low energy dissipation. Other EO switches are based on liquid-crystals and exhibit high extinction ratio, high reliability and low power consumption. Also semiconductor optical amplifiers (SOAs) can be used as an ON-OFF switch by varying its bias current. By reduced bias voltage, no population inversion is achieved and device rather absorbs input signal, thereby building the OFF-state. In contrary, if a sufficiently high bias voltage is applied, input signal will be amplified and, thus, the ON-state is achieved. However, EO switches suffer from high insertion loss and possible polarization dependent loss (PDL). PDL can be combat at the cost of higher driving voltage, and consequently lower switching speed, which is not desirable. The switching speed of EO switches is in the order of several nanoseconds or even hundreds of picoseconds.

The changes of refractive index due to the interaction between acoustic and optical waves in the crystal are utilized in the AO switches. The switching speed of AO switches is in the order of hundreds of nanoseconds and is limited by the propagation speed of acoustic waves. AO switches can also be implemented on Lithium Niobate.

The TO effect utilizes the temperature dependence of the refractive index. The advantage of thermo-optical switches is its generally small size, but high driving voltage and high power dissipation make such switches highly impractical. Crosstalk and insertion loss values are also not very satisfactory. Mostly used materials for TO switching implementations are silica and polymers. Switching time of TO switches lies in the order of milliseconds.

Opto-mechanical (OM) switches are based on mechanics and free-space optics and switching is performed by some electro-mechanical means such as by moving mirrors or directional couplers. Regarding its optical performance parameters, OM switches have low insertion loss, low polarization dependant loss (PDL) and low crosstalk. However, drawbacks of this type of switches are their relatively low switching speed in the order of few milliseconds, which could be unacceptable for some applications requiring dynamic operation. MEMS (Micro Electro Mechanical System) switches form a subcategory of the OM switches. In particular, 3-D MEMS is the most promising option for applications that do not require fast switching, but large port counts. Switches with more than 1,000 ports have already been demonstrated. MEMS devices are scalable, cascadable and consume low power. Challenges regarding MEMS are packaging and time-consuming fabrication.

Arrayed waveguide grating (AWG) based switches have gained a particular attention for large-scale switching fabrics and there are several architectures that base upon these particular elements. Since they are passive elements, they can potentially provide low power operation. However, additional active elements such as wavelength converters (WCs) are needed to implement switching operation. Switching time of AWG-based switches is determined by the tuning speed of the wavelength converter.

According to its switching time, insertion loss, crosstalk and PDL, a switching device is more or less suitable for a particular application. For example, the switching time required for fast packet switching applications should be small in comparison to the average length of data packets. Some examples of realization options of optical switching elements are presented in Figure 5. In Figure 5a), a switching device based on SOAs is shown, another one using AWGs and tunable wavelength converters (TWCs) is presented in Figure 5b) and the third one is based on MEMS (see Figure 5c)).

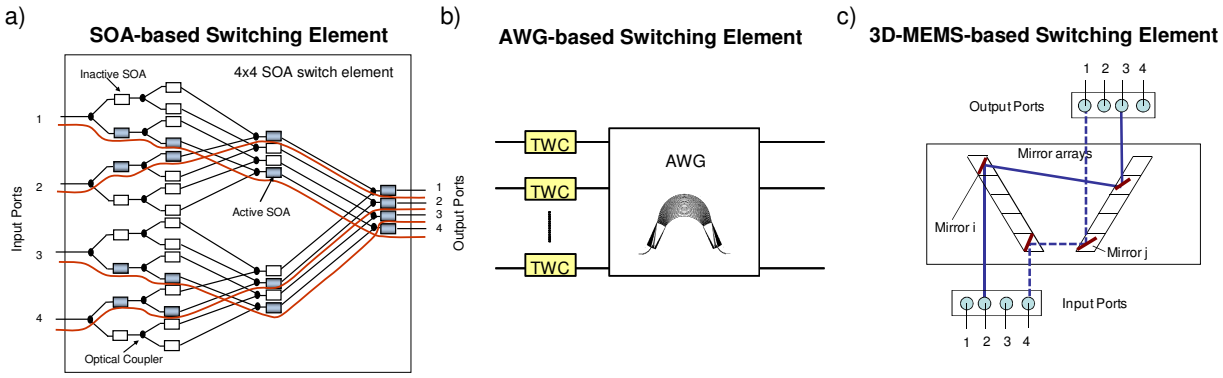

Figure 5. Examples of 4x4 switching elements based on a) semiconductor optical amplifiers, b) arrayed waveguide gratings with tunable wavelength converters, and c) 3D optical micro-electro-mechanical systems (MEMS).

Switch Architecture and Network Topology

There are various interconnecting arrangements, and thus various architectures of internal switching fabrics that can be used in large-scale systems. Table 2 summarizes the most important architecture types and shows their blocking characteristics.

Switch Architectures	Blocking Type	Network Topologies	Blocking Type
Classical log*N*	blocking	Clos (Fat-Tree, Multistage)	strict sense non-blocking if $p^* \geq 2n\text{-}1$
Benes	rearrangeably non-blocking	d-dim Symmetric Mesh	rearrangeable, blocking if $p > 2$
Crossbar	wide sense non-blocking	d-dim Symmetric Torus	rearrangeable, blocking if $p > 2$
Spanke	strict sense non-blocking	d-dym Hypercube	reareangeable
Cantor	strict sense non-blocking		
Banyan	blocking		

* p is the number of edge switches
n is the number of ports of a single switching element

Table 2. Blocking characteristics if some switch architectures and network topologies.

All architectures have several important characteristics such as: the number of stages, the number of feasible connections and the type of switching elements used to construct a large fabric. Multistage interconnection network (MIN) is an important class of interconnection arrangements that consists of multiple stages with a number of switching units in each stage. There are several grades of the connection quality such as generally blocking, wide-sense non-blocking, rearrangeably non-blocking and strict-sense non-blocking. The interconnecting arrangements can be classified in different ways, e.g. regarding its blocking probability, packet loss probability, the number of stages, or with respect to number of paths through the switching fabric.

Scalability and Energy Efficiency

Both switch architecture and network topology have a strong influence on the achievable performance and scalability of the entire interconnection system. As it can be seen from Figure 6a), a full-mesh network provides the worst scalability because the number of links increases

rapidly with increasing the number of nodes in a HPC system or line cards in a network router. However, the full mesh topology is able to provide a strict sense non-blocking operation and the maximum possible line data rates for each connection in the system because dedicated connections are provided between all nodes in the network. The two examples of the Torus topology (2D and 3D Torus) have much better scalability, but they are either blocking or rearrangeably non-blocking networks with links shared by several connections. The 3-stage Clos network can provide both good scalability and high line data rates. However, the main disadvantage of the Clos network is its relatively high switch cost.

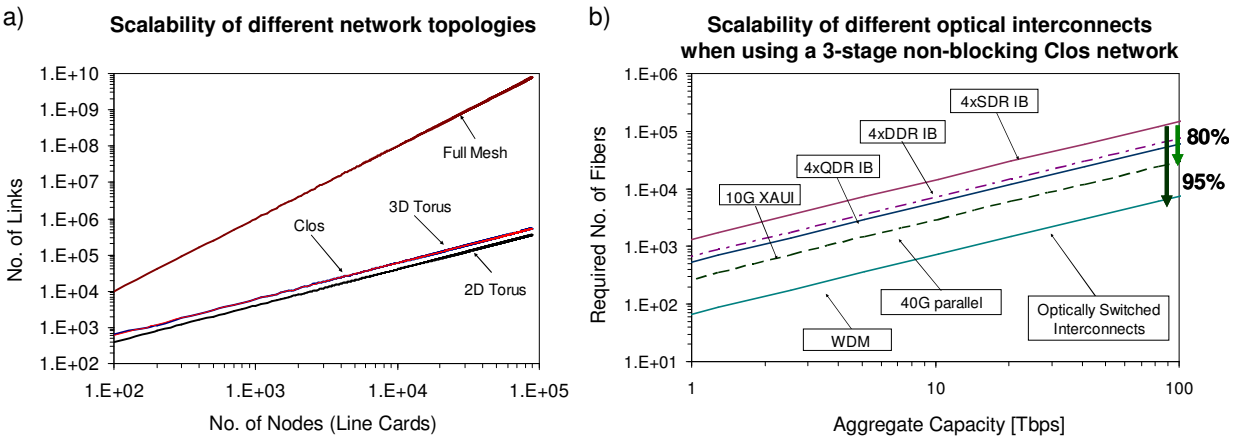

Figure 6. Scalability of large-scale interconnects when using a) different network topologies and b) various transmission and switching technologies in a 3-stage non-blocking Clos network.

Assuming a fixed network topology one can reduce the required number of fibers through increasing the data rate per channel, using wavelength-division multiplexing (WDM), and applying optically switched interconnects as presented on the example of a 3-stage Clos network in Figure 6a). The required number of fiber links can be reduced by up to 80 % through increasing the channel data rate, while using WDM and optically switched interconnects can lead to an even higher reduction of 95% [20]. However, the scalability of optically switched interconnects is mostly limited by physical impairments on optical signals that set limits on the maximum achievable size of switching elements without applied signal regeneration. The scalability of optical switching elements can be obtained by carrying out physical layer simulations and taking into account all relevant effects such as attenuation, dispersion, noise accumulation, nonlinear effects and crosstalk [21]. For instance, Figure 7a) shows the maximum achievable number of ports for the three types of optical switching elements presented in Figure 5 when assuming a target upper bound on eye closure penalty of 4°dB. The results indicate that AWG-based optical switching elements have the worst scalability when inband crosstalk is considered. In contrary, if the inband crosstalk can be eliminated, very large switches are imaginable. SOA-based switches are scalable to about 32 ports without any signal regeneration. The largest switching elements are possible when using 3D-MEMS switches. Large switch fabrics can be implemented by interconnecting the optical switching elements in a multistage network and applying signal regeneration after a number of switch stages.

Energy efficiency is an important parameter to be considered in designing future interconnecting solutions. Recently, three different realization options for optically switched interconnects (based on SOA, AWG and MEMS switching elements and using optical signal

regeneration) have been investigated and compared to the conventional approach based on optical point-to-point links with electronic switches [22,23]. As it can be observed from Figure 7b), optically switched interconnects are, additionally to requiring the lowest number of fiber links, also able to provide superior energy efficiency. The energy efficiency of the electronically switched interconnects lies in the range between 0.8 and 1.18°nJ/bit, while the optically switched interconnects are able to provide an improvement in energy efficiency of at least 50% (260 - 610°pJ/bit). Furthermore, optically switched interconnects base on 3D-MEMS are 50% more energy efficient than those based on SOA. However, although the MEMS-based interconnects are the most energy efficient and scalable solution, they are not very well suitable for dynamically switched interconnection systems because of their inherently large switching times in the order of milliseconds (see Table 1).

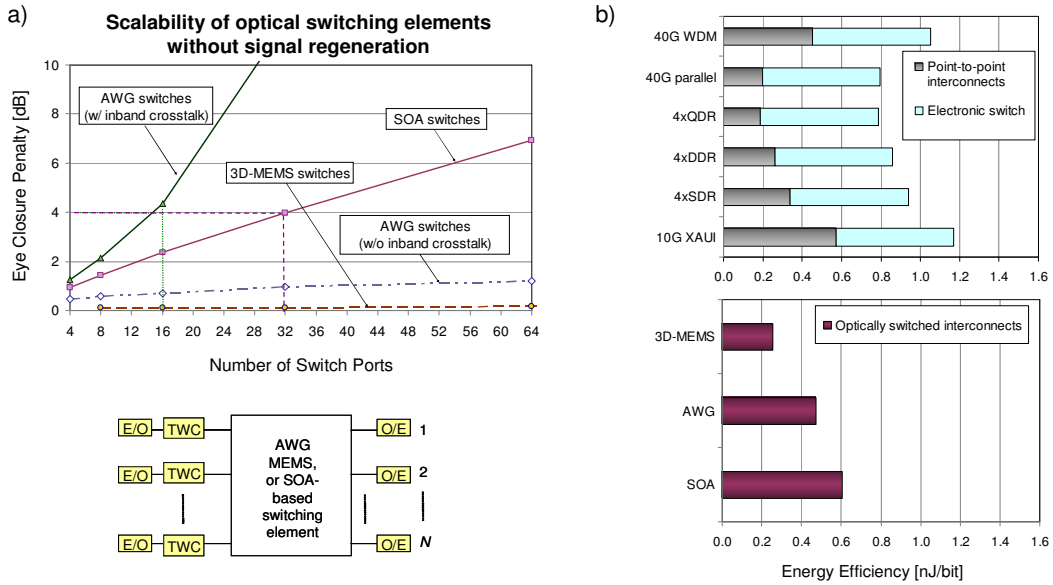

Figure 7. a) Scalability of optical switching elements without signal regeneration and b) energy efficiency of optical point-to-point interconnects with electronic switches and optically switched interconnects with signal regeneration.

CONCLUSIONS

This paper reviews recent developments in technologies and architectures for optical interconnects in large-scale network and data processing elements. It focuses on interconnects ranging from chip-to-chip links on printed circuit boards (PCBs) to system-level interconnects. We propose to design and evaluate new concepts for large-scale interconnects by taking into account different hierarchy levels and performance metrics such as feasibility, traffic-related performance, power consumption and techno-economics. It can be concluded that optically switched interconnects have the potential to significantly improve both performance and energy efficiency of high-capacity systems. However, due to the fact that optically switched interconnects provide switching in an optically transparent manner, the number of cascaded switch stages is limited by various impairments in the optical domain, which can be overcome by performing 3R signal regeneration. Even when optical signal regeneration is applied to improve scalability, the optically switched interconnects still provide at least 50% better energy efficiency than systems comprising optical point-to-point interconnects and electronic switches.

REFERENCES

1. C. Lange, et al., ”Energy Consumption of Telecommunication Networks”, ECOC 2009, Vienna, Austria, Sept. 20- 24, 2009, paper 5.5.3.
2. X. Zhou, et al., ”Transmission of 32-Tb/s Capacity Over 580 km Using RZ-Shaped PDM-8QAM Modulation Format and Cascaded Multimodulus Blind Equalization Algorithm”, Journal of Lightwave Technology, Vol. 28, pp. 456 - 465, 2010.
3. D. A. B. Miller, ”Device Requirements for Optical Interconnects to Silicon Chips”, Proceedings of the IEEE, Vol. 97, No. 7, pp. 1166 - 1185, July 2009.
4. The Climate Group, ”SMART 2020: Enabling the low carbon economy in the information age”, GLobal eSustainability Initiative, Techn. Rep. 2008.
5. S. Aleksic, ”Analysis of Power Consumption in Future High-Capacity Network Nodes”; IEEE/OSA JOCN, Vol. 1, No. 3, pp. 245 - 258, August 2009.
6. M. R. Feldman, et al., ”Comparison Between Optical and Electrical Interconnects Based on Power and Speed Considerations”, Applied Optics, Vol. 27, pp. 1742 - 1751, 1998.
7. Top500 Computer Sites, ”Statistics on high-performance computers”, http://www:top500:org/
8. P. Westbergh, et al., ”32 Gbit/s multimode fiber transmission using high-speed, low current density 850 nm VCSEL”, Electronics Letters, Vol. 45, No. 7, pp. 366-368, March 26th 2009
9. Petar Pepeljugoski et al., ”Low Power and High Density Optical Interconnects for Future Supercomputers”, OFC 2010, San Diego, California, USA, March 2010, paper OThX2.
10. Y. Benlachtar, at al.,”Optical OFDM for the Data Center”, ICTON 2010, Munich, Germany, June 27 -July 1, 2010, paper We.A4.3.
11. X. Ye, et al., ”Assessment of Optical Switching in Data Center Networks”, OFC 2010, San Diego, California, USA, March 21-25, 2010, paper JWA63.
12. G. I. Papadimitriou, C. Papazoglou, and A. S. Pomportsis, ”Optical Switching: Switch Fabrics, Techniques, and Architectures”, IEEE/OSA JLT, Vol. 21, No. 2, pp. 384-405, 2003.
13. K. Vlachos, et al., ”Photonics in switching: enabling technologies and subsystem design”; OSA Journal of Optical Networking, Vol. 8, Issue 5, pp. 404-428, 2009.
14. N. Fehratovic and S. Aleksic, ”Power Consumption and Scalability of Optically Switched Interconnects”, OFC 2011, Los Angeles, California, USA, March 2011, paper JWA84
15. G. Schmid, W. R: Leeb; G. Langer, ”Experimental Demonstration of the Robustness Against Interference of Optical Interconnects on Printed Circuit Boards”, IEEE Photonics Society Winter Topicals Meeting, Mallorca, Spain, 2010, pp. 93 – 94.
16. K. K. Tung; W. H. Wong, E. Y. B. Pun, ”Polymeric Optical Waveguides Using Direct Ultraviolet Photolithography Process”, Applied Physics A 80 (2005), Nr. 3, pp. 621 – 626.
17. G. Langer, V. Satzinger, V. Schmid, G: Schmid, W. R. Leeb, ”PCB with Fully Integrated Optical Interconnects”, SPIE Photonics West 2011, Optoelectronic Interconnects and Component Integration XI Bd. 7944, 2011, pp. 794408-1 – 794408-15.
18. X. Wang, et al., ”Fully Embedded Board-Level Optical Interconnects from Waveguide Fabrication to Device Integration” IEEE/OSA JLT Vol. 26 (2008), No. 2, pp. 243 – 250.
19. M. A. Taubenblatt, ”Challenges and opportunities for integrated optics in computing systems”, SPIE Conference Series Bd. 6124, 2006, pp. 612406-1 – 612406-11.
20. S. Aleksic, N. Fehratovic, ”Requirements and Limitations of Optical Intermodule Interconnects for High-Capacity Network Elements” (invited), ICTON 2010, Munich, Germany, June 27 -July 1, 2010, paper We.A1.2.
21. S. Aleksic, ”Energy Efficiency of Electronic and Optical Network Elements” (invited), IEEE Journal of Selected Topics in Quantum Electronics, Vol. 17 (2011), No. 3; pp. 1 - 13.
22. S. Aleksic, ”Electrical Power Consumption of Large Electronic and Optical Switching Fabrics”; IEEE Winter Topicals 2010, Majorca, Spain; January 2010, pp. 95 - 96.
23. S. Aleksic, N. Fehratovic, ”Analysis of Scalability of Optical Intrasystem Interconnects”, Journal of Networks (JNW), Academy Publisher, scheduled for publication in June 2012.