# Photonic integration enabling new multiplexing concepts in optical board-to-board and rack-to-rack interconnects

Dimitrios Apostolopoulos<sup>\*a</sup>, Paraskevas Bakopoulos<sup>a</sup>, Dimitrios Kalavrouziotis<sup>a</sup>, Giannis Giannoulis<sup>a</sup>, Giannis Kanakis<sup>a</sup>, Nikos Iliadis<sup>a</sup>, Christos Spatharakis<sup>a</sup>, Johan Bauwelinck<sup>b</sup>, Hercules Avramopoulos<sup>a</sup>

<sup>a</sup>National Technical Univ. of Athens, 9 Iroon Polytechniou Str., Zografou, Athens, Greece, GR15780; <sup>b</sup>INTEC/IMEC, Univ. of Ghent, Sint-Pietersnieuwstraat 41, 9000 Ghent, Belgium

## ABSTRACT

New broadband applications are causing the datacenters to proliferate, raising the bar for higher interconnection speeds. So far, optical board-to-board and rack-to-rack interconnects relied primarily on low-cost commodity optical components assembled in a single package. Although this concept proved successful in the first generations of optical-interconnect modules, scalability is a daunting issue as signaling rates extend beyond 25 Gb/s. In this paper we present our work towards the development of two technology platforms for migration beyond Infiniband enhanced data rate (EDR), introducing new concepts in board-to-board and rack-to-rack interconnects.

The first platform is developed in the framework of MIRAGE European project and relies on proven VCSEL technology, exploiting the inherent cost, yield, reliability and power consumption advantages of VCSELs. Wavelength multiplexing, PAM-4 modulation and multi-core fiber (MCF) multiplexing are introduced by combining VCSELs with integrated Si and glass photonics as well as BiCMOS electronics. An in-plane MCF-to-SOI interface is demonstrated, allowing coupling from the MCF cores to 340x400 nm Si waveguides. Development of a low-power VCSEL driver with integrated feed-forward equalizer is reported, allowing PAM-4 modulation of a bandwidth-limited VCSEL beyond 25 Gbaud.

The second platform, developed within the frames of the European project PHOXTROT, considers the use of modulation formats of increased complexity in the context of optical interconnects. Powered by the evolution of DSP technology and towards an integration path between inter and intra datacenter traffic, this platform investigates optical interconnection system concepts capable to support 16QAM 40GBd data traffic, exploiting the advancements of silicon and polymer technologies.

**Keywords:** optical interconnects, silicon photonics, multi-level modulation, WDM, multi-core fiber, Active Optical Cable, datacenter, High Performance Computing.

# 1. INTRODUCTION

Optical interconnects (OI) have spearheaded the paradigm shift towards a content-centric Internet architecture, ushering the age of network affluence. In this network evolution end-users seek instant, ubiquitous access to information content that is stored in datacenters. With online content reaching an enormous capacity of more than 500 billion gigabytes, most Internet traffic currently originates or terminates in a datacenter [1]. The sustained growth of Internet, fostered by the broad uptake of cloud applications gives rise to exorbitant interconnection requirements inside the datacenter, which are further exacerbated by additional traffic demands between different intra-datacenter units. The situation is vividly illustrated through recent traffic forecasts, estimating datacenter traffic to reach 7.7 zettabytes per year by 2017, as compared to an annual traffic of 1.4 zettabytes expected to cross the Internet and IP WAN networks within the same time frame [1][2]. These projections add up to a hectic growth rate for datacenter traffic reaching 25% per year, which poses an outright challenge to the technological approaches currently used in optical interconnects.

So far, optical interconnects relied primarily on low-cost commodity optical components assembled in a simple package. Although this concept proved substantially successful in the first generations of optical-interconnect modules, scalability is a daunting issue to meet the requirements of future datacenters. With per-chip density and performance of electronics

<sup>\*</sup>apostold@mail.ntua.gr; phone +30 210 772 2871; fax +30 210 772 2077; www.photonics.ntua.gr

continuing to improve along the trajectory of Moore's law, optical interconnection needs to keep pace and offer the capability to handle the increasing amounts of traffic. This is vividly reflected in the latest standardization activities in the field, either concluded or under discussion (100 G Ethernet, Infiniband EDR, 32G Fibre Channel), which have raised the bar in signal rate to 25 Gb/s levels. As component manufacturers grapple to develop products meeting these performance targets with reasonable cost, it is well understood that new technological paradigms have to be pursued for the next upgrade steps, to accommodate this bandwidth explosion.

Optical interconnects have established themselves through Active Optical Cables (AOCs), performing rack-to-rack interconnection at distances of a few tens -or even hundreds- of meters. AOC components currently rely on two main technological approaches, which are also heavily researched for emerging board-to-board applications, namely Vertical Cavity Surface-Emitting Lasers (VCSELs) and silicon photonics. VCSEL technology has been the main workhorse of optical interconnects enjoying support of all leading companies in the field due to their inherent low cost, low power consumption and reliability. Another salient advantage of VCSELs is their high yield which is associated to their small size (20,000 VCSELs can be grown on a 3-inch wafer) and capability for testing at the wafer level before dicing and packaging. Multi-mode VCSELs emitting at 850 nm find broad commercial application, representing the most mature and ubiquitous technology in the sector [3][4]. Meanwhile long-wavelength VCSELs have been identified as a reasonable migration step, offering improved energy efficiency and higher modulation speeds with the same reliability [3][6]. One of the challenges associated to VCSELs lies in their speed scaling; although there have been demonstrations of VCSEL operation up to 40 Gb/s, performance is highly dependent on temperature. Current state-of-the-art is limited to 25 Gb/s at 55° [7] which prevents commercial application of VCSELs beyond this speed. In addition, although engineering a VCSEL's emission wavelength is feasible, fabricating WDM VCSELs on the same chip retaining high speed capability remains a challenge [8]. To scale speed despite these limitations, VCSELs are typically fabricated in large 2D arrays as for example in typical parallel optics arrangements with up to 12 channels at 12.5 Gb/s per channel, however requiring a bulky 24-line optical fiber array and the respective driver electronics. System integration of VCSELs involves different options like mounting in simple SNAP12 packages [9] or bonding to PCBs [10].

The second technology that is attracting increasing attention in the context of optical interconnects is silicon photonics. Silicon is an ideal ecosystem for low-cost mass market applications due to the vast availability of silicon foundries and the huge market of silicon electronic products that allows sharing of equipment and development costs. High yield implementation of high-performance passive photonic structures on silicon is now a reality, but it is the active photonic functionalities that hinder widespread adoption of silicon photonics. Recent breakthroughs have proven the feasibility of silicon modulators and Germanium detectors as well as the heterointegration of III-V lasers on silicon and the room temperature operation of Germanium lasers [14][11]-[13]. On the downside, these advanced components often involve customized fabrication processes and specialized equipment that undermine the cost benefits brought by sharing the resources with the electronics industry [15]. In addition, their reliability has not been proven yet, thus discouraging adoption by the end users. Despite these challenges, the potential advantages of silicon photonics are increasing the count of leading optical interconnects suppliers actively investigating the technology (e.g. Luxtera/Molex, Intel, Kotura/Mellanox) and the first commercial products are available [16]. Scaling the speed above 25 Gb/s with silicon photonics is proven to be technologically feasible but several challenges remain to be overcome towards commercial viability. State-of-the-art silicon modulators typically achieve moderate extinction ratios at 40 Gb/s, whereas interferometric designs can sacrifice power efficiency due to the compromise between modulator length and driving voltage [17][19][20]. Migration to hybrid silicon designs can improve speed yet at the expense of fabrication complexity and CMOS incompatibility [21]. Similar tradeoffs between performance and manufacturability are encountered in silicon lasers where further steps still need to be taken towards technological maturity. On the other hand, Germanium photodetectors have already reached 40 Gb/s performance rivalling their III-V counterparts [22]. Hence, although a complete set of active and passive functionalities has been realized with silicon photonics, it is rather premature to consider silicon photonics as a commercially available, self-contained photonic ecosystem. Even more, the combination of photonic with electronic functions poses additional considerations regarding the optimum solution (front-end or backside) as a function of design flexibility and yield [15][18].

Obviously, both technologies that emerged during the proliferation of optical interconnects have valuable strengths but there are also significant hurdles hindering them from further scaling efficiently. As such, to address pressing requirements and avoid an imminent capacity crunch it is necessary to reassess the existing technological baseline and to capitalize on concepts that have been proven in relevant domains, in view of the requirements set by the application and dictated by practical considerations.

# 2. ENABLING CONCEPTS FOR CAPACITY SCALING IN OPTICAL INTERCONNECTS

Optical interconnects started as a niche application, offering a combination of speed and reach unparalleled by their electrical counterparts. The excess cost of optics, rationalized by the alluring performance benefits brought to the overall system, was partly tackled with the use of inexpensive, commodity optical components. Scaling the capacity of parallel optical links has been regularly addressed by either increasing the number of parallel "lanes" or enhancing the line rate at each lane. Fast forward to today and optical interconnects has evolved into a multi-billion market covering a breadth of different applications spanning from High Performance Computing to Data Centers and expanding even to consumer electronics. The phenomenal surge of online data is for the first time pushing "traditional" OI technologies to their speed boundaries. Increasing the number of parallel lanes (and associated I/Os) may be considered in the short-term but cannot cope with the need for higher information density and optimized energy efficiency. For example, the most efficient way to provide optical interconnection is to align the state-of-the-art electrical signal rate with the optical signal data rate, thus avoiding expensive and power hungry serializer/deserializer (SerDes) circuitry. Next generation common electrical interfaces (CEI) have recently defined future signaling rates as high as 28 Gb/s and 25 Gb/s for chip-to-chip and backplane applications [23] and new, faster signaling rates will be adopted in the years to come. As a result, new technologies are being put in the spotlight in the context of optical interconnects, such as single-mode VCSELs, silicon photonics as well as advanced packaging techniques proven in the electronics industry, such as 3D integration. Their common denominator is "industry qualified" photonic integration, holding promise for high speed and massive information density using reliable fabrication processes that are amenable to mass production. The latter is a key ingredient for the "industrial qualification" of a technology as manufacturability and technological maturity are pivotal in order to reap the benefits of photonic integration. Although directly competing with the cost of field-proven commodity components remains a formidable challenge, photonic integration comes with substantial flexibility that enables intriguing design concepts for cost-per-bit optimization. In this context, introduction of new multiplexing schemes in optical interconnects is pursued as a viable solution for substantially upgrading the capacity of a link without scaling its cost proportionally. The following multiplexing dimensions, enabled by photonic integration, are actively pursued in optical interconnects.

- Wavelength-division multiplexing (WDM): Transmission of several wavelength channels over the same physical medium (optical fiber or integrated waveguide) is applied to share its cost and size over multiple connections. WDM is considered hand-in-hand with single-mode operation as the next logical step to support intra-datacenter connectivity in warehouse server farms at line rates beyond 25 Gb/s, overcoming the effective modal bandwidth limitations posed even by state-of-the-art multimode fibers [24]. Besides better transmission characteristics, single-mode fibers are cheaper than multimode whereas single-mode waveguides can boost integration density in integrated components. It is expected that for a 100 Gb/s datacom link at various reaches, the slope (€/meter) of the WDM links is 10 times lower than the slope for single mode fiber without WDM and 20 times lower than the slope for multi-mode fiber [25].
- Space-division multiplexing (SDM) with multi-core fibers (MCF) is finding increasing interest for telecom applications: high-quality glass fibers are making their debut and prices are expected to drop due to increasing uptake. Leveraging these advanced products in AOCs can lead to significant cost savings by reducing the number and size of interfaces to the transceivers. Smaller fiber cross-section compared to fiber bundles can alleviate ventilation and fiber management issues currently plaguing large datacenters, also saving valuable chip real-estate for fiber interfacing. Compared to multi-mode implementations in plastic [26] or glass [27], single mode multicore fibers offer improved transmission characteristics without suffering from differential phase-skew issues that afflict multimode fibers when data streams span multiple data lanes.
- Multi-level modulation has already established itself in telecom applications, allowing high throughput without stressing the line rate beyond the capabilities of commercially available electronics. Multi-level approaches have been researched also in datacom contexts, with gradual complexity regarding the required electronics at the transmitter and receiver [28]-[30]. Much like in telecom, it is expected that multi-level formats and associated digital signal processing (DSP) will be a valuable tool in optical interconnects, which will be adopted progressively depending on the availability of cheap and fast electronic processing circuitry. In the longer run, even coherent detection may become economically viable for particular application cases.

Each of the above concepts comes with its own implications in terms of system design and performance, a prominent example being WDM's requirement for laser sources emitting at distinct wavelengths. In the same fashion, SDM-MCF

calls for specialized chip-to-MCF interfaces whereas multi-level modulation poses requirements on the linearity of the associated components. Optimum choice of a multiplexing dimension -or combination thereof- is a function of the specific OI application targeted and its particular requirements, versus technology-imposed implications. In an optical interconnects landscape currently extending over a broad variety of diverse application cases there can be no one-sizefits-all solution; instead our approach is to generate a quiver of concepts and their enabling technologies that can be used, combined and tailored to the specific application requirements. Considering the main application domains for highspeed interconnects, namely HPC and data centers, it is possible to extract some rather simplistic yet indicative observations. In HPC systems the main emphasis is on throughput and latency, favoring the application of new multiplexing concepts provided that the latency overhead of any potentially required DSP is minimal. Data centers on the other hand give highest priority to the cost; the cheapest technology that achieves reliable operation is preferred, and capacity requirements are typically less "exotic" than in HPC. Although at first consideration this discourages application of any unconventional disciplines or technologies, there is still room for the concepts described earlier in this section. In the emerging paradigm of mega-datacenters it is often necessary to deploy connections up to 2 km and current technologies based on 850 nm VCSELs face fundamental limitations to achieve this reach. Migration to singlemode designs comes as the only viable solution, making an upgrade to WDM more of an evolutionary step rather than a disruptive technological leap.

As discussed at the beginning of this section, photonic integration can be enabled by -and at the same time be the enabler of- new multiplexing schemes in optical interconnects. Nevertheless, it could be argued that despite the cost savings achieved by multiplexing, the current maturity of photonic integration would still require a learning curve until the provided cost/bit outmatches that of established "commodity" optics, thus hindering the commercial take-up of these new concepts. This hurdle can be overcome by the unique performance attributes of the integrated/multiplexing approach. As elaborated in the previous paragraph there are "sweet spots" in the application space where no rival can compete in terms of performance, which can serve as opportunities for early market adoption. In such application cases cost is not necessarily relevant – history has proven this with the introduction of optical interconnects per se over their cheaper electrical counterparts.

In this background the pivotal challenge is to develop the enabling photonic integration technology that can be tailored to the application requirements, is upgradable so as to follow the current levels of technology maturity (future-proof) and relies on industrial qualified processes. In the following sections we present our recent efforts towards this goal in the framework of two collaborative research projects co-funded by the European Commission, MIRAGE and PhoxTrot.

# 3. TAILORING PHOTONIC INTEGRATION FOR MULTIPLEXING IN OPTICAL INTERCONNECTS: USE CASES

## 3.1 The MIRAGE project: a 3D optical engine combining VCSELs with silicon photonics

MIRAGE is a three-year collaborative project focusing on photonic integration for cost-optimized board-to-board and rack-to-rack interconnects. The project aims to disrupt current development efforts and present a technology capable of achieving up to 1 Tb/s capacity with a single optical interconnect. To meet this ambitious goal, MIRAGE pursues to combine key integration technologies on a flexible optical engine that can be configured according to the particular application requirements and contemporary technological maturity. MIRAGE blends the most prominent technologies in optical interconnects, VCSELs and silicon photonics, using innovative integration techniques. Flexibility of the MIRAGE optical engine permits different circuit configurations with increasing complexity and functionality and allows the technology to support a multitude of multiplexing schemes in the wavelength, space and modulation-format domains. Advanced integration processes enable the cost-efficient assembly of diverse components irrespective of their functionality (electrical, optical, electro-optic), material system (silicon, compound semiconductor) and supplier. MIRAGE pursues a combination of 2.5D and 3D integration to assemble the OI components at chip or wafer scale, optimizing fabrication yield and overall cost.

The optical engine of MIRAGE is shown schematically in Figure 1. At the core of the 3D integrated opto-electronic stack lies a silicon board consisting of two layers. The upper layer provides silicon photonic functionality using a SOI  $400 \times 340$  nm strip waveguide platform. At present the silicon photonic layer performs passive functionalities such as wavelength multiplexing and demultiplexing and hosts the optical interfaces to the on-chip interconnected photonic components. The silicon photonic layer can be gradually upgraded to provide active functionalities (e.g. modulation, photodetection) as soon as the underlying processes are industrially qualified and economically advantageous, leveraging ongoing efforts in the field. The lower silicon layer implements a high speed electrical interposer that

facilitates interconnection between the on-chip active electrical and optoelectronic components at the horizontal plane. Electrical connectivity at the vertical direction is enabled by the use of proprietary through silicon vias (TSVs). The silicon photonic and electronic boards are provided by different foundries and are bonded at wafer scale using CMOS compatible processes. Active electro-optic functionalities are performed by compound semiconductor components, as they currently represent the most mature solution. Data generation is accomplished by short-cavity (SC) InP VCSEL arrays operating at the C-band, flip-chipped on the silicon photonic board. Direct modulation of the SC VCSEL has been recently shown with NRZ format at 40 Gb/s as well as with PAM-4 format at 25 Gbaud [7][31]. Photodetection is performed either with a back-illuminated InP photodiode attached to the silicon photonic board or with a waveguide photodiode butt-coupled to the silicon waveguide layer. All electrical driving circuitry such as VCSEL drivers and TIAs are attached to the electrical interposer layer at chip scale, as shown in Figure 1. Electrical driving circuitry in MIRAGE is based on 0.13  $\mu$ m SiGe BiCMOS technology that provides optimum combination of power consumption and performance (considering e.g. eye opening, jitter, linearity and noise) at speeds up to 40 Gbaud. Interfacing to single-mode, multi-core fiber is facilitated by grating couplers or in-plane mode-converters followed by 3D glass waveguide fan-outs that translate the linear waveguide array of the silicon chip to a circular arrangement interfacing with the MCF cores [36].



Figure 1. Schematic diagram of the MIRAGE optical engine. All supported optical interfaces and configurations are depicted.

One of the key targets set in the development of the MIRAGE optical engine is the flexibility to implement different circuit designs using a reconfigurable technological toolbox that can adapt the provided functionality (and corresponding system complexity) to the requirements of the particular application. To tackle with this challenge, MIRAGE has defined a generic integration concept based on the combination of discrete modules with its 2.5D and 3D integration scheme (as shown in Figure 2 and described in the previous paragraph) and is working towards establishing the entire integration process. The modular architecture of MIRAGE allows straightforward modification of the platform simply by omitting unnecessary modules along with the corresponding processes. Availability of multiple types of optical and electrical interfaces in the MIRAGE toolbox facilitates different configurations of the MIRAGE modules. An example of how the MIRAGE optical engine can be tailored to meet different requirements and circuit designs is illustrated in Figure 2 showing two of the possible configurations, both pursued within the project. The first configuration (type I), shown in Figure 2, is an optical interconnect implementing SDM-MCF multiplexing and multi-level modulation. At the transmitter side SC-VCSEL arrays are flip-chipped on the SOI board and the generated light is coupled in the silicon waveguide layer through grating couplers, optimized for vertical coupling with reduced back-reflections [32]. This

approach alleviates the requirement for micro-optic lenses and associated ferrule attachments owing to the proximity of the grating coupler to the VCSEL. Adaptation of the SC-VCSEL to operate at single-polarization is under final optimization to account for the polarization sensitivity of the grating coupler. The SC-VCSELs are connected electrically to the VCSEL drivers attached on the other side of the chip, through the silicon interposer layer and TSVs. The silicon waveguides terminate to grating couplers following the same physical arrangement with the MCF, to allow vertical coupling from the grating coupler directly to the assigned fiber core. At the receiver side, the MCF is connected to a 3D glass waveguide fan-out that converts the MCF core arrangement to a linear arrangement matching the pitch and configuration of the back-illuminated photodiode array. Electrical interconnection to the photodiode TIAs is facilitated through the interposer and TSVs in the same fashion as with the transmitter. The use of linear VCSEL drivers and photodiode TIAs enables generation and detection of multi-level formats such as PAM-4. Using the type I configuration MIRAGE aims to demonstrate total capacity of 208 Gb/s using four parallel lanes operating at 26 Gbaud with PAM-4 modulation.

#### **MIRAGE TYPE I interconnect**



Figure 2. Schematic of supported configurations with the MIRAGE optical engine. Top: Type I interconnect 4-lane transmitter (left) and receiver (right). Bottom: Type II interconnect 12-lane transceiver (only 4-lanes depicted for clarity).

A more complex configuration possible with the MIRAGE optical engine is shown on the bottom row of Figure 2, showing an optical interconnect supporting multiplexing in the wavelength, space and modulation-format domain (type II). Bonding and coupling of the VCSELs follows the same approach with the configuration of interconnect type I. VCSELs emitting at different WDM wavelengths and fabricated on the same monolithic dye are used in this case and wavelength pairs are multiplexed at the silicon photonic layer. Coupling from the silicon chip to the MCF follows a different method, following an in-plane arrangement. In-plane chip-to-MCF coupling is performed in two stages; an initial beam expansion takes place at the SOI board by means of an SU-8 overcladding followed by further beam expansion and adjustment of the waveguide arrangement in a 3D glass waveguide fan-out that interfaces with the MCF. This approach is further described in section 3.3 and allows polarization insensitive operation of the receiver without the requirement of complex polarization diversity techniques. Photodetection is achieved by a waveguide photodetector array butt-coupled to the silicon chip, using suitable mode converters at both chips. To improve receiver sensitivity, an optical pre-amplified receiver design is adopted employing semiconductor optical amplifiers (SOAs) before each photodetector. Using the type II configuration MIRAGE aims to demonstrate up to 960 Gb/s total capacity using six parallel optical lanes with two WDM channels in each lane, operating at 40 Gbaud with PAM-4 modulation.

Table 1. Left: simulated insertion losses of MIRAGE optical engine and total link loss for each interconnect type. Right: Compatible loss budget and margin for a 100 m AOC. NRZ and PAM-4 modulation formats are considered for interconnect type I and type II at 26 and 40 Gbaud respectively.

Type I MIRAGE optical interconnect					
component loss		compatible loss budget (26 Gbaud)			
grating coupler (VCSEL-to-SOI)	3.8 dB		without FEC	with RS (255,223) FEC	
SOI waveguide	0.5 dB		(threshold $1 \times 10^{-12}$ )	(threshold $1 \times 10^{-3}$ )	
grating coupler (SOI-to-fiber)	3.8 dB	NRZ	15.6 dB	19.5 dB	
fiber loss	0.1 dB	PAM-4	10.8 dB	14.5 dB	
3D glass fanout	0.7 dB		margin		
TOTAL	8.9 dB	NRZ	6.7 dB	10.6 dB	
		PAM-4	1.9 dB	5.6 dB	
<u>Type II MIRAGE optical interconnect</u>					
component loss		compatible loss budget (40 Gbaud)		loss budget (40 Gbaud)	
grating coupler (VCSEL-to-SOI)	4 dB		without FEC	with RS (255,223) FEC	
SOI waveguide & MUX	2.5 dB		(threshold $1 \times 10^{-12}$ )	(threshold $1 \times 10^{-3}$ )	
in-plane coupler	3 dB	NRZ	20.3 dB	23.9 dB	
fiber loss	0.1 dB	PAM-4	15.5 dB	18.9 dB	
in-plane coupler	3 dB				
SOI waveguide & DEMUX	2.5 dB		margin		
SOI inverse taper	$1 \ dB$	NRZ	4.2 dB	7.8 dB	
TOTAL	16.1 dB	PAM-4	-	2.8 dB	

From the system perspective, a compelling challenge with the MIRAGE optical engine comes with the number of optical interfaces and the associated loss introduced to the link. Tables 1 summarizes the optical losses currently achievable at each type of interface and the compatible loss budget for each configuration, for both NRZ and PAM-4 modulation formats. In the case of a type I interconnect, the compatible loss budgets allow error-free detection without FEC for both NRZ and PAM-4 at 26 Gbaud (current state-of-the art signaling rate) including sufficient margin for end-of-life performance degradation. This type of interconnect could find application in mega data-centers due to the enhanced reach provided by the single-mode operation and the cost advantages over an all-silicon implementation that require non-standardized and thus expensive processes. In the type II interconnect, the SOA amplifier gain only partially compensates the losses induced by the increased number of optical interfaces. As a result, only NRZ operation can be

achieved without FEC, whereas PAM-4 operation necessitates the use of "standard" RS (255,223) FEC. This type of interconnect can find application in HPC systems where throughput is the main target. Selection of NRZ or PAM-4 modulation is dependent on the combination of throughput, latency and energy/bit required and may be statically or even dynamically reconfigured for each optical data lane in a given interconnect, offering more flexibility to the system.

#### 3.2 The PhoxTrot project: Terabit/s Optical Interconnect Technologies for On-Board, Board-to-Board and Rack-to-Rack data links

PhoxTroT is a four-year, large-scale research effort, focusing on high-performance, low-energy and cost and small-size optical interconnects across the different hierarchy levels in Data Center and High-Performance Computing Systems: onboard, board-to-board and rack-to-rack (Figure 3). Optics at board- and chip-level have still not convinced about their cost-efficient character, the main reason for that being the plethora of photonic technology platforms required in the realization of optical links at these communication distance levels. On the other hand, the "one size fits all" does certainly not apply in this short-range datacom area, necessitating the use of a wide technology portfolio for realistic optical interconnection implementations: VCSELs and, in general, III/V materials are mainstream in optical Tx/Rx functions, polymeric or glass material platforms offer the optimal board-level waveguiding characteristics, mirrors and micromechanics apply in the chip-to-board and board-to-board coupling problems, while CMOS electronics remains without doubt the "intelligent" technology layer. The recent revolution of Silicon photonics brings an additional candidate technology into the game, since they have revealed a tremendous potential for CMOS-compatible optical chipscale circuitry. Reducing the cost when such a variety of technologies is incorporated into the fabrication line is certainly a non-trivial task. Following this rational, PhoxTroT tackles optical interconnects in a holistic way, synergizing the different fabrication platforms in order to deploy the optimal "mix&match" technology and tailor this to each interconnect layer targeting cost efficiency that can only be achieved when all different technology platform requirements are aligned to a well-established process-level framework and only if the proper balance on the use of the different on-board technologies is achieved.



Figure 3. PhoxTrot's optical interconnect platform prototype across the different hierarchy levels, on-board, board-to-board and rack-to-rack.

Moreover, as commercial datacom environments have only recently started to "absorb" optical interconnect technologies, there is still plenty of room available for defining their photonic interconnect roadmap. This path includes a necessary "experience transfer" through the establishment of Optical Packet Switch Transport and Advanced Modulation Format transmission features as datacom device deployment guidelines. OPST technologies can definitely

expedite bandwidth availability and resource/energy consumption optimization, since bandwidth is occupied only for the duration required releasing bandwidth and hardware resources when no traffic exchange is needed. On the same line, Advanced Modulation Formats have shown to provide data-rate enhancements without requiring additional impetus on the bandwidth of optical technologies and, consequently, on their energy consumption envelope. This roadmap takes into account that Active Optical Cables should be replaced by passive fiber ribbons at several stages in the beyond 2020 optical interconnect area, increasing the perspective to avoid o/e/o conversion at specific rack-level routing applications. Moreover, Advanced Modulation formats at inter-rack level will be certainly utilized as a means to increase aggregate rack-to-rack transmission rates without enhanced optical component and device bandwidth requirements. To this end, a specific level of routing procedures performed directly in the optical domain should support multi-Gbaud packet traffic formats.

Within this frame, PhoxTrot intends to deploy technology building block and components complying with OPST and Advanced Modulation Format operational criteria:

- Active Optical Cable technologies involving 16QAM modulation formats towards breaking the 1Tb/s aggregate data-rate capabilities by employing just 40Gb/s optical component technology. This will bring advanced intensity/phase modulation schemes into rack-to-rack communication level.
- Convergence of both OPST and Advanced Modulation Format characteristics in optical PCB- and rack-level router deployments, evaluating the perspectives of PhoxTrot's on-board and board-to-board routing modules towards operation with advanced intensity/phase modulation formats.

PhoxTrot's AOC, shown in Figure 4 is designed to will deploy an 8-lane cable with 1.28 Tb/s aggregate rate and lower than 5mW/Gb/sec relying entirely on SOI-based Tx/Rx chips and offering in this way energy reductions close to 60% compared to state-of-the-art AOCs. This scheme will be compatible with existing packaging solutions, so that a 4-lane cable (640Gb/s) with QSFP+ or an extended 12-lane (1.92Tb/s) with CXP form factor can be produced and enter the market. Volume manufacturing processes to be established within PhoxTrot for the SOI-based Tx/Rx chips are expected, according to provisional estimates, to bring the cost of this 16QAM Active Optical Cable Technology down to less than 3\$/Gb/sec even for the QSFP+ form factor, leading to a cost reduction of more than 75% compared to conventional QSFP+ modules.



Figure 4. PhoxTrot's Active Optical Cable concept.

## 3.3 Development of in-plane chip-to-MCF coupling interfaces

One of the key objectives of the MIRAGE project is the implementation of Space-Division Multiplexing (SDM) through the utilization of Multi-Core Fibers (MCFs) so as to increase the bandwidth of the Active Optical Cables (AOC). MCFs

will lead to smaller cable crossection compared to fiber bundles alleviating ventilation and cable management issues in datacenters, hence, reducing the overall operating cost. Design and fabrication of several types of multicore fibers have been reported so far exhibiting low loss and low crosstalk [34]. However their commercialization is currently prevented due to lack of low-loss and low-cost reliable interfaces for signal in and out coupling, hindering the efficient coupling to optical transceivers. Only limited number of solutions has been proposed so far, including tapered multicore fiber connectors (TMC) or fan-out devices able to interrogate each core of the 3-core fiber independently. Nevertheless, these approaches exhibit either non-uniform insertion losses between cores or low scaling capabilities, whereas as none of those developments report on the interfacing to integrated waveguide platforms or components.



Figure 5. a) Multi-Core Fiber crossection b), MCF "Fan-out Glass Interface and c), SU8-loaded Silicon Spot-Size Converter

Key advancements in the MIRAGE project focus on developing in-plane chip-to-MCF interfaces with increased corecounts, enabling the seamless connection between 7-core MCFs (Figure 5 a)) and Silicon-on-Insulator (SOI) Photonic Integrated Circuits (PICs). To this end, a recent technology that has been successfully commercialized by Optoscribe [35], will be employed utilizing embedded waveguides fabricated using ultrashort pulse laser direct writing [36]. Using this approach 3D waveguide topologies can be produced within glass substrates in order to 'fan-out' from the cylindrical geometry of MCF to planar waveguide configurations, suitable for coupling to SOI PICs, as shown in Figure 5 b). This approach is robust, readily manufacturable and easily integrated with other platforms.



Figure 6. a) Crossectional view of the SSC, b) 1-Dimensional linear SU8 overcladding taper, c) linear and exponential inverse silicon taper profiles and d) overall tapered exponential SSC concept

On the SOI PIC side, a technology that was published by AMO-GmbH [37], will be employed, ensuring efficient and polarization independent coupling between a Single-Mode Fiber (SMF) and an SOI nanowire through the development of a Spot-Size Converter (SSC) (Figure 5 c)). The SSC structure, as depicted in Figure 6 a), consists of an inversely 1D tapered silicon waveguide cladded in a polymer (SU8) overcladding. The optical signal is initially coupled and guided in the SU8 overcladding that equally supports both polarizations, due to the refractive index contrast and the crossection geometry. As the silicon waveguide widens, a phase matching condition, between the SU8 mode and the Si core mode, starts to appear, leading to the gradual shift of the propagating light towards the Si core. The conversion efficiency highly depends on the geometrical characteristics of the SU8 cladding as well as on the widening profile of the Si inverse taper. Up to now, coupling losses of less than 3 dB for both TE and TM polarization have been reported, based, however, on the use of lensed fibers at the input, so as to keep the beam size as low as possible, allowing for polymer overcladding crossectional dimensions in the order of few microns.

Within the MIRAGE project, the SSC design will be modified so as to enable the efficient coupling between the aforementioned MCF "Fan-out" Glass Interface and the SOI PIC, avoiding at the same time the need for micro-lenses, in order to ensure low fabrication complexity and cost. To this end, the SU8 crossection will have to be increased so as to match the original Mode Field Diameter (MFD) of the glass interface that is somewhat less than 10  $\mu$ m. Extensive simulations, though, have revealed that increasing the SU8 crossection has a rather detrimental effect on the performance of the SSC as conversion efficiency drops logarithmically, counteracting the improved coupling losses stemming from the spatial mode matching between the glass interface and the fan-out. Moreover, from the fabrication perspective, developing quality SU8 polymer waveguides thicker than 5  $\mu$ m is a quite challenging task that will severely affect the fabrication yield. In order to overcome these issues, an 1D SU8 taper (Figure 6 b)) will be incorporated in front of the SSC so as to increase coupling efficiency by matching the MFDs in the horizontal axis while keeping the SU8 thickness below 5  $\mu$ m. This configuration will lead to decreased coupling losses without affecting the conversion efficiency. In fact, it will allow for further optimization of the SSC design that can be achieved by changing the profile of the Si inverse taper to exponential instead of linear, as shown in Figure 6 c). This way, the phase matching condition will be enhanced increasing the graduality of the mode conversion and, consequently, the efficiency. Figure 6 d) illustrates an overall 3D representation of the tapered SSC with the exponential silicon inverse taper.



Figure 7. a) Propagation profile of the TE and TM fundamental mode in the 1D-SU8 taper and b) Sideview of the conversion process (norm of the electrical component) in the SSC for linear and exponential profiles for both TE and TM polarization

Figure 7 demonstrates preliminary simulation results obtained on the design of the chip-to-MCF interface using the 2D FDTD and 3D FDTD software solution provided by Phoenix BV and Lumerical respectively. Figure 7 a) shows the propagation profile of the fundamental TE and TM modes in the 1D SU8 taper. The width of the taper on the left side is 10.5  $\mu$ m and it is linearly decreased to 4.5  $\mu$ m while the SU8 thickness is 3  $\mu$ m. The simulation results revealed losses less than 2% for both polarization modes, proving the feasibility of the proposed 1D tapering structure. Figure 7 b) depicts a sideview impression of the conversion process (norm of the electrical component) in the SSC for linear and

exponential Si taper profiles and both TE and TM polarizations at 1.55  $\mu$ m center wavelength. The SU8 cladding is, in both cases, 4.5x3  $\mu$ m2, the height of the Si taper is 340 nm while the width starts from 60 nm and widens up to 400 nm. The length of the overall device is 200  $\mu$ m. It is clear that the mode conversion for the TM polarization is smoother, resulting to less residual power guided in the SU8 overcladding after the conversion, while eliminating the higher order modes that appear when the Si taper is linear. This is also evident by the 65% improvement of the TM conversion efficiency. It should be noted that the TE polarization exhibits only a marginal improvement of 5%, however, this is quite expected as the conversion efficiency of the TE mode is adequately high (87%) also for the linear Si taper.



Figure 8. a) Conversion efficiency (%) of the SSC for exponential Si taper and 3µm SU8 thickness and b) improvement in the performance of the overall coupling scheme achieved by the design modifications

Figure 8 a) shows the exponential profile SSC conversion efficiency plotted against the SU8 overcladding width for values ranging from 1.5  $\mu$ m to 6  $\mu$ m, revealing that the optimum width for obtaining both high efficiency and low Polarization Dependent Loss (PDL) is 4.5  $\mu$ m. Figure 8 b) on the other hand, presents the improvement on the performance of the overall chip-to-MCF structure achieved through the introduction of the 1D SU8 taper and the exponential Si taper profile in the initial (regular) SSC design shown in Figure 6 a) [37]. It can be seen that the 1D SU8 taper improves the performance of both TE and TM modes by almost 3dB while the exponential Si taper reduces the PDL to almost zero.

To conclude, it is clear that the proposed modifications have profoundly advanced the performance of the overall coupling scheme, constituting an important step towards the realization of the chip-to-MCF interfaces of the MIRAGE project. Further optimization steps will focus on enhancing the coupling efficiency between the MCF "Fan-out" glass interface and the tapered SSC, targeting the improvement of the spatial mode matching, between the two devices, also in the perpendicular direction.

#### 3.4 Development of 40 Gbaud linear VCSEL drivers and TIAs

High-speed electronic circuits are crucial to the success of optical interconnections, not only to generate, process and store huge amounts of data, but also to interface between two very different worlds: the purely digital CMOS and the rather analog opto-electronic (E/O and O/E) devices, such as VCSELs and their respective drivers as well as photodiodes and receivers.

Dedicated analog and mixed-signal circuits are needed to obtain the best cost-performance-power trade-off, taking into account the system requirements, modulation scheme and photonic component parameters. The most critical point is the interconnection between driver and modulator/laser and between photodiode and transimpedance amplifier (TIA). First of all, bringing electronics and photonics close together is important to minimize electrical parasitics (e.g. inductance and capacitance) that would otherwise reduce the bandwidth or require transmission line interconnects (with considerable power dissipation in their termination resistors). Excessive parasitics can also strongly impact the stability and, consequently, the sensitivity of the TIA. Additionally, compact, low-power and low-cost multi-channel transmitter and receiver modules are an absolute necessity for high-speed short-links. The close integration of several driver and TIA circuits on a single chip poses several design challenges e.g. low power consumption to alleviate thermal issues,

small footprint, low RF crosstalk and power supply integrity. Following a co-simulation and co-optimization process of electronics, photonics and interconnects, there is no need for a pre-defined impedance level (typ. 50 Ohm) for the RF interfaces because high-speed electronics and photonics can be co-optimized with interconnect parasitics [31].

In the context of this paper, optical board-to-board and rack-to-rack interconnects, several driver and TIA arrays are currently under development, with a focus on multi-level modulation. Multi-level signaling obviously increases the transmission capacity, but brings a number of additional design challenges compared to the design of NRZ drivers and receivers. First of all, nearly all signal sources or chips work with two-level NRZ signals. To generate the 4-PAM, system experiments often use a power combiner to add two binary scaled bit streams together giving four signal levels. Such a solution is, however, not practical or power efficient when integrating multiple drivers on a single chip. For this reason, the VCSEL driver array under development includes an NRZ to 4-PAM converter on chip, together with a retiming function to regenerate the NRZ streams on chip to remove unwanted signal distortions resulting from electrical signal impairments in connectors and PCB traces. The signal conversion in the VCSELs or interconnects. Finally, the output stage is optimized to drive single-mode VCSEL arrays, with a common anode contact. At the receiver side, 4-PAM requires linear signal processing in order to maintain the spacing between the different signal levels, but also a higher sensitivity is required compared to NRZ. Figure 9 shows a detailed layout view of some core circuits in the recently finalized VCSEL driver and TIA chips in 130nm SiGe BiCMOS.





Figure 9. Layout view of some core VCSEL driver (left) and TIA circuits (right)

## 4. CONCLUSION

We have present our work towards the development of two technology platforms for migration beyond Infiniband enhanced data rate (EDR), introducing new concepts in board-to-board and rack-to-rack interconnects. The first platform, developed in the framework of ICT MIRAGE project, relies on proven VCSEL technology, wavelength multiplexing, PAM-4 modulation and multi-core fiber (MCF) multiplexing, combining VCSELs with integrated Si and glass photonics as well as BiCMOS electronics. An in-plane MCF-to-SOI interface is demonstrated, allowing coupling from the MCF cores to 340x400 nm Si waveguides. Development of a low-power VCSEL driver with integrated feed-forward equalizer is reported, allowing PAM-4 modulation of a bandwidth-limited VCSEL beyond 25 Gbaud. The second platform, developed within the frames of the European project PHOXTROT, considers the use of modulation formats of increased complexity in the context of optical interconnects. Powered by the evolution of DSP technology and towards an integration path between inter and intra datacenter traffic, this platform investigates optical interconnection system concepts capable to support 16QAM 40GBd data traffic, exploiting the advancements of silicon and polymer technologies.

## ACKNOWLEDGEMENTS

This work was supported by the European Commission through FP7 project MIRAGE (318228) and PhoxTrot (318240). We would like to thank the consortium partners of both initiatives.

#### REFERENCES

- [1] Cisco, "Cisco Global Cloud Index: Forecast and Methodology, 2012–2017", 2013, <<u>http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns1175/Cloud\_Index\_White\_Paper.pdf</u> >.
- [2] Cisco, "Cisco Visual Networking Index: Forecast and Methodology, 2012–2017", 29 May 2013, <<u>http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/white paper c11-481360.pdf</u>>.
- [3] Blokhin, S., A., Lott, J., A., Mutig, A., Fiol, G., Ledentsov, N., N., Maximov, M., V., Nadtochiy, A., M., Shchukin, V., A., and Bimberg, D., "Oxide-confined 850 nm VCSELs operating at bit rates up to 40 Gbit/s," Electronics Letters, 45(10), (2009).
- [4] Rylyakov, A., Schow, C., Proesel, J., Kuchta, D., M., Baks, C., Li, N., Y., Xie, C. and Jackson, K., "A 40-Gb/s, 850-nm, VCSEL-Based Full Optical Link," Proc. OFC, OTh1E.1. (2001).
- [5] Kapon, E. and Sirbu, A., "Long-wavelength VCSELs: Power-efficient answer", Nature Photonics 3, 27-29 (2009).
- [6] Hoffman, W., "Monolithic 2D high-power arrays of long-wavelength VCSELs", Proc. SPIE 6908, (2008).
- [7] Muller, M. et. al., "1550-nm High-Speed Short-Cavity VCSELs", IEEE Journal Sel. Top. Quant. Electr., 17(5) (2011).
- [8] Hoffmann, W. et. al., "1.55- μm VCSEL Arrays for High-Bandwidth WDM-PONs", IEEE Photonics Technology Letters, 20(4), 291-293 (2008).
- [9] Benner A., F. et. al., "Exploitation of optical interconnects in future server architectures", IBM J. Res. & Dev., 49(4), (2005).
- [10] Schow, C., L. et. al., "225 Gb/s Bi-Directional Integrated Optical PCB Link", Proc. OFC, PDPA2 (2011).
- [11] Keyvaninia, I., S., Verstuyft, S., Van Landschoot, L., Van Thourhout, D., Roelkens, G., Duan, G., Lelarge, F., Fedeli, J., M., Messaoudene, S., De Vries, T., Smalbrugge, B., Geluk, E., J., Bolk, J. and Smit, M., "Heterogeneously integrated III-V/silicon distributed feedback lasers," Optics Letters, 38(24), 5434-5437 (2013).
- [12] Liang, D., Srinivasan, S., Fattal, D., A., Fiorentino, M., Zhihong, H., Spencer, D., T., Bowers, J., E. and Beausoleil, R., G., "Teardrop Reflector-Assisted Unidirectional Hybrid Silicon Microring Lasers," IEEE Photonics Technology Letters, 24(22), 1988-1990 (2012).
- [13] Jain, S., R., Tang, Y., Chen, H., W., Sysak, M., N. and Bowers, J., E., "Integrated Hybrid Silicon Transmitter," Journal of Lightwave technology, 30(1), (2012).
- [14] Camacho-Aguilera, R., E., Cai, Y., Patel, N., Bessette, J., T., Romagnoli, M., Kimerling, L., C. and Michel, J., "An electrically pumped germanium laser," Opt. Express, 20, 11316-11320 (2012).
- [15] Baehr-Jones, T., Pinguet, T., Guo-Qiang, P., L., Danziger, S., Prather, D. and Hochberg, M., "Myths and rumours of silicon photonics", Nature Photonics, 6, 206–208 (2012).
- [16] Molex, "Molex Demonstrates World's First 100 Gbps Integrated Optical Transceiver Products", <<u>http://www.molex.com/molex/news/display\_news.jsp?channel=News+In+Brief&channelId=-8&oid=978</u>>.
- [17] Alduino, A. et. al., "Demonstration of a High Speed 4-Channel Integrated Silicon Photonics WDM Link with Hybrid Silicon Lasers", Proc. OSA Integrated Photonics Research, Silicon and Nano Photonics, PDIWI5 (2010).
- [18] Kopp, C., Bernabé, S., Bakir, B., B., Fedeli, J., Orobtchouk, R., Schrank, F., Porte, H., Zimmermann, L. and Tekin, T., "Silicon Photonic Circuits: On-CMOS Integration, Fiber Optical Coupling, and Packaging," Selected Topics in Quantum Electronics, IEEE Journal of, 17(3), 498-509 (2011).
- [19] Gardes, F., Y., Thomson, D., J., Emerson, N., G. and Reed, G., T., "40 Gb/s silicon photonics modulator for TE and TM polarisations," Opt. Express, 19, 11804-11814 (2011).
- [20] Tu, X., Liow, T.,Y., Song, J., Luo, X., Fang, Q., Yu, M. and Lo, G., Q., "50-Gb/s silicon optical modulator with traveling-wave electrodes," Opt. Express, 21, 12776-12782 (2013).
- [21] Chen, H., W., Peters, J., D. and Bowers, J., E., "Forty Gb/s hybrid silicon Mach-Zehnder modulator with low chirp," Opt. Express, 19, 1455-1460 (2011).
- [22] Vivien, L., Osmond, J., Fédéli, J., M., Morini, D., M., Crozat, P., Damlencourt, J., F., Cassan, E., Lecunff, Y. and Laval, S., "42 GHz p.i.n Germanium photodetector integrated in a silicon-on-insulator waveguide," Opt. Express, 17, 6252-6257 (2009).
- [23] Fields, M., "Optical interconnects for chip-to-chip communications", Proc. ECOC 2010, 19-23 (2010).
- [24] Coleman, D., "10G and Beyond in the Data Center with OM3 Fiber", Cabling Business Magazine, 32-35 (2008).
- [25] Liu, H., Lam, C., F. and Johnson, C., "Scaling Optical Interconnects in Datacenter Networks: Opportunities and Challenges for WDM", Proc. 18th IEEE Symposium on High Performance Interconnects, 113-116 (2010).

- [26] Loquai, S. et. al., "10.7Gb/s discrete multitone transmission over 25m bend-insensitive multicore Polymer Optical Fiber", IEEE Photonics Technology Letters, 22(21), 1604-1606 (2010).
- [27] Lee B., G. et. al., "120-Gb/s 100-m Transmission in a Single Multicore Multimode Fiber Containing Six Cores Interfaced with a Matching VCSEL Array", Proc. IEEE Photonics Society Summer Topical Meeting Series, Playa del Carmen, 223-224 (2010).
- [28] Cunningham, J., E., Beckman, D., Zheng, X., Huang, D., Sze, T. and Krishnamoorthy, A., V., "PAM-4 Signaling over VCSELs with 0.13μm CMOS Chip Technology," Opt. Express, 14, 12028-12038 (2006).
- [29] Ingham, J., D. et. al., "32 Gb/s Multilevel Modulation of an 850 nm VCSEL for Next-Generation Datacommunication Standards", in Proc. OSA/ CLEO, CWJ2 (2011).
- [30] Szczerba, K., et. al.. "30 Gbps 4-PAM transmission over 200m of MMF using an 850 nm VCSEL", Proc. ECOC (2011).
- [31] Rodes, R., Mueller, M., Li, B., Estaran, J., Jensen, J.B., Gruendl, T., Ortsiefer, M., Neumeyr, C., Rosskopf, J., Larsen, K.J., Amann, M., Monroy, I.T., "High-Speed 1550 nm VCSEL Data Transmission Link Employing 25 GBd 4-PAM Modulation and Hard Decision Forward Error Correction," Journal of Lightwave Technology, 31(4), 689-695 (2013).
- [32] Chen, X., Li, C., Tsang, H., "Fabrication-Tolerant Waveguide Chirped Grating Coupler for Coupling to a Perfectly Vertical Optical Fiber," IEEE Photonics Technology Letters, 20(23), 1914-1916 (2008).
- [33] Bauwelinck, J., Vaernewyck, R., Verbrugghe, J., Soenen, W., Moeneclaey, B., Van Praet, C., Vyncke, A., G., Torfs, Yin, X., Qiu, X., Z., Vandewege, J., Sotiropoulos, N., De Waardt, H., Cronin, R., Maxwell, G., Tekin, T., Bakopoulos, P., Lai, C.,P. and Townsend, P., D., "High-speed electronics or short-link communication [invited]", Proc. ECOC, (2013).
- [34] Takenaga, K. et. al, "Reduction of Crosstalk by Quasi-Homogeneous Solid Multi-Core Fiber," Proc. OFC, (2010).
- [35] Optoscribe, "Optoscribe 3D Optofan Series", <<u>http://optoscribe.com/wp-content/uploads/2013/09/OPTOSCRIBE-</u> <u>4-and-7-Multicore-fiber-fanout.pdf</u>>.
- [36] Thomson, R., R. et. al., "Ultrafast-laser inscription of a three dimensional fan-out device for multicore fiber coupling applications", Opt. Exp., 15(18), 11691-11697 (2007).
- [37] Wahlbrink et al., "Fabrication of high efficiency SOI taper structures", Mic. Eng., 86, 1117–1119 (2009).