Ronny Henker, David Schöniger, Guido Belfiore, Laszlo Szilagyi, Jan Pliva, Mahdi Khafaji, Frank Ellinger, Krzysztof Nieweglowski, Tobias Tiedje, Karlheinz Bock

Tunable broadband integrated circuits for adaptive optical interconnects

Erstveröffentlichung in / First published in:

17th Conference on Optical Fibres and Their Applications. Supraśl, 2017. Bellingham: SPIE, Vol. 10325 *[Zugriff am: 23.05.2019]*.

DOI: https://doi.org/10.1117/12.2271014

Tunable broadband integrated circuits for adaptive optical interconnects

Ronny Henker, David Schöniger, Guido Belfiore, Laszlo Szilagyi, Jan Pliva, et al.

Ronny Henker, David Schöniger, Guido Belfiore, Laszlo Szilagyi, Jan Pliva, Mahdi Khafaji, Frank Ellinger, Krzysztof Nieweglowski, Tobias Tiedje, Karlheinz Bock, "Tunable broadband integrated circuits for adaptive optical interconnects," Proc. SPIE 10325, Optical Fibers and Their Applications 2017, 103250P (10 February 2017); doi: 10.1117/12.2271014

SPIE.

Event: 17th Conference on Optical Fibres and Their Applications, 2017, Supraśl, Poland, Poland

Tunable broadband integrated circuits for adaptive optical interconnects

Ronny Henker[a], David Schoeniger[a], Guido Belfiore[a], Laszlo Szilagyi[a], Jan Pliva[a], Mahdi Khafaji[a], Frank Ellinger[a], Krzysztof Nieweglowski[b], Tobias Tiedje[b], and Karlheinz Bock[b]

[a]Technische Universität Dresden, Chair for Circuit Design and Network Theory, 01062 Dresden, Germany
[b]Technische Universität Dresden, Institute of Electronic Packaging Technology, 01062 Dresden, Germany

ABSTRACT

To accommodate the growing demand on higher speeds, low latencies and low energy consumption, the interconnections within and between data centers are supposed to be implemented as optical fiber and waveguide interconnects in future. Optical fiber interconnects provide several advantages over their electrical counterparts as they enable higher bandwidth densities and lower losses at high frequencies over distances longer than few centimeters. However, nowadays optical fiber interconnects are usually not very energy-efficient. The systems in optical networks are mostly optimized for running at their peak performance to transmit the information with the highest available error-free data rate. But the work load of a processor system and hence of an optical link is not constant and varies over time due to the demand of the running applications and users. Therefore, optical interconnects consume the same high power at all times even if lower performance is required.

In this paper a new method for the tuning of optical interconnects for on-board and board-to-board optical communication is described. In this way the performance of the transceiver systems of the link is adapted to the present transmission workload and link requirements. If for example lower data rates are required, the bandwidth and therefore the power consumption of the systems can be reduced. This tuning is enabled by the integrated circuitry of the optical link. Different methods for such an adaptive tuning are described and several practical examples are reviewed. By using adaptive bandwidth reduction in the circuits, more than 50 % of the consumed power can be saved. These savings can result in tremendous reductions of the carbon footprint and of the operating costs produced by data centers.

Keywords: optical interconnects, integrated circuit design, performance and power adaptivity, on-board and board-to-board optical communication

1. INTRODUCTION

According to the recent Cisco Global Cloud Index,[1] more than 75 % of the Internet traffic remains inside the data centers. Due to increasing demand of Internet applications and services the big data centers hosted by Facebook, Google and Apple for instance will grow to dimensions which exceed multiple times to the area of the largest football stadiums of the world. This trend comes along with several challenges: 1) data center interconnects need to become faster, 2) data center interconnects need to span over larger distances and 3) the network power consumption in a data center will significantly increase to a few MW.[2] To accommodate these requirements high bandwidth and low energy consuming interconnections are required in future. In the past few years optical fiber and waveguide links gained a huge attraction as promising candidates for future data center interconnects.[3] Research is conducted in all directions such as materials, integration technologies, fiber and waveguide structures, network architectures and routing algorithms etc. For example a Tb/s vertical-cavity surface-emitting laser (VCSEL) based holey optochip has been demonstrated with 24 bi-directional transceiver channels each running

Further information: (Send correspondence to R. Henker)
E-mail: ronny.henker@tu-dresden.de
Websites: www.addapt-fp7.eu and https://tu-dresden.de/ing/forschung/sfb912

Optical Fibers and Their Applications 2017, edited by Jan Dorosz, Ryszard S. Romaniuk,
Proc. of SPIE Vol. 10325, 103250P · © 2017 SPIE · CCC code: 0277-786X/17/$18
doi: 10.1117/12.2271014

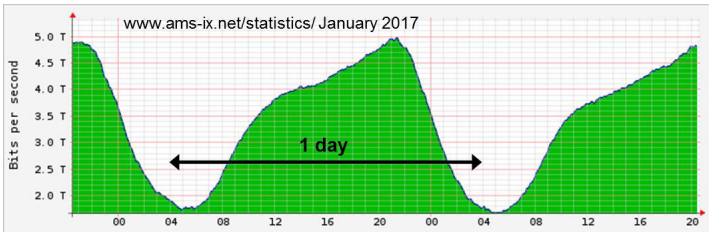

Figure 1. Dynamic network traffic during workday.[15]

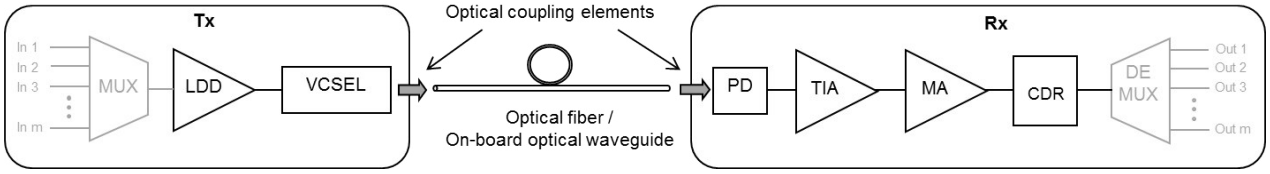

Figure 2. Simplified block diagram of an optical interconnect system.

at a data rate of 20 Gb/s with a link efficiency of 7.3 pJ/bit.[4] Nowadays, single VCSEL-based non-return-to-zero (NRZ) optical links can go up to 71 Gb/s.[5] Even on-chip optical interconnects are proposed.[6] However, such interconnects still show a worse energy efficiency of 7.3 pJ/bit to 25 pJ/bit for data rates not higher than 71 Gb/s. Furthermore, optical interconnection systems are rather static in their performance and are optimized for the peak performance of the link. In contrast the link workload is not constant over time due to user and applications use, as shown in Fig. 1. Therefore, also even optical interconnects do not achieve their maximum efficiency today.

By implementing dynamic transmission systems with tunable performance and power consumption more energy can be saved. Recently some dynamic link concepts have been proposed on network level, e.g. by routing and switching optimization.[7–9] However, no adaptivity on system and component level has been implemented into systems nowadays.

In this paper, a new method for the adaptive tuning of optical communication links is described. In this way the performance and the power consumption of the electrical transceiver circuitry is adapted to the present transmission workload by changing their operating points, i.e. their bias currents. The results of investigations in two research projects are reviewed: 1) Highly adaptive energy-efficient computing (HAEC), a collaborative research center which includes the development of optical on-board NRZ links with polymer optical waveguides and adaptive integrated circuits (ICs) in SiGe BiCMOS technology,[10,11] and 2) Adaptive data and power aware transceivers for optical communication (ADDAPT), a EU FP7 project which focuses on a 4-channel transceiver for optical high-speed energy-efficient NRZ links with multimode fibers (MMF) and adaptive ICs in 14 nm CMOS technology,[12–14] The investigations on IC adaptivity in both projects show that with adaptive tuning of the performance of the transceivers, more than 50 % of the system's power consumption can be reduced.

2. ADAPTIVE OPTICAL LINKS

The main blocks of an optical interconnect are shown in Fig. 2. Mostly a multiplexer (MUX) serializes multiple low data rate electrical interfaces to one high-speed bit stream. This signal is fed to a laser, e.g. a VCSEL, by a laserdiode driver (LDD). The LDD provides the signal current for direct laser modulation. In this way the VCSEL converts the high-speed current signal from the electrical to the optical domain. After transmission over an optical fiber or waveguide the signal is received by a photodetector (PD) which converts the optical signal back to the electrical domain. After amplification by a receiver amplifier the data is restored by a clock and data recovery (CDR) and finally the high-speed data stream is parallelized by a demultiplexer (DEMUX) to lower rate interfaces. The receiver amplifier normally consists of a transimpedance amplifier (TIA) as input stage and a main amplifier (MA) which can be a limiting amplifier (LA) or variable gain (VG) amplifier. The

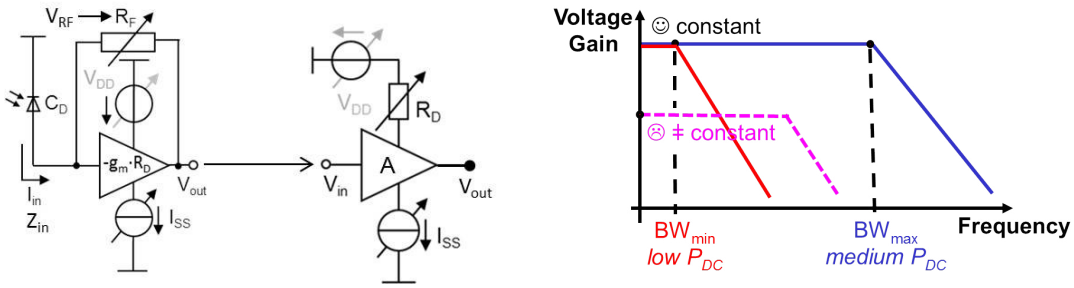

Figure 3. Adaptive tuning of TIA and MA; (left) concept and (right) illustrated frequency response of the MA.

TIA is essential to convert the weak photocurrent to a large voltage while the MA further boosts the voltage signal swings to logic levels.

As it can be seen from Fig. 2, most of the transceiver building blocks are electrical circuitry. Especially, the analog ICs are the ones with high power consumption to reach high speeds. By implementing adaptive tuning methods the performance of the ICs can be adjusted to the link requirements and therefore also the power consumption can be reduced. The basic concept of this tuning is shown in Fig. 3. The adaptation is implemented by changing the operating points (primarily the bias currents) and the load impedances of the circuits at the same time.[16] If lower incoming data rates are present, the supply current I_{ss} of the circuits can be reduced. This obviously lowers the power consumption P_{DC}. At the same time the transconductance g_m and transit frequency f_t of the transistors are decreased. As a consequence both, the bandwidth on the one hand but also gain on the other hand are lowered. Since a constant output level is required for driving a laser or a CDR, the gain drop the MA has to be compensated. This can be achieved either by simultaneously increasing the load impedance R_D of the MA or, for a receiver amplifier as a combination of TIA and MA, by increasing the TIA gain via tuning the feedback resistance R_F which results in a VG-TIA.[17] Furthermore, complete ICs or parts of them can be switched off to save power. In the following section the several adaptivity approaches are described in more detail and the power saving potential is shown.

3. ADAPTIVE BROADBAND CIRCUITS

3.1 Adaptive receiver amplifier

First, an adaptive receiver amplifier with continuous bandwidth and power consumption tuning has been realized in a 0.13 μm SiGe BiCMOS technology.[18] The TIA consists of three main parts: a VG-TIA input stage for the current-to-voltage conversion, a main amplifier (MA) for the amplification and the output driver which is only used for measurement purposes for matching to the impedance of the equipment. By reducing the current of the MA, as it is illustrated in Fig. 3, the performance and also the power consumption can be reduced. However, this reduction is accompanied by a gain drop of the MA. Therefore, a gain drop compensation is implemented in the TIA input stage. The feedback resistance R_F, which defines the transimpedance gain, can be adjusted by using a field-effect transistor (FET), which operates in the ohmic region as a steerable resistor, controlled via V_{RF}. If V_{RF} is decreased, the feedback resistance and thus the input impedance as well as the gain of the VG-TIA input stage become higher. In this way the overall gain can be kept constant and the circuit can be tuned continuously. The frequency response of the continuous adaptive receiver amplifier for several operating points is shown on the left hand side of Fig. 4. A maximum data rate of 88 Gb/s at a power consumption (TIA core without the output driver) of approximate 30 mW was achieved. By reducing the current and therefore the power consumption to approximately 18 mW a data rate of 20 Gb/s can be accommodated while the amplifier gain remains constant at 64 dbΩ. In this way 40 % of power can be saved while the amplification remains the same. Figure 5 shows the measured electrical eye diagrams in low (at 20 Gb/s) and high performance (at 50 Gb/s) states. Thereby, the measured data rate of 50 Gb/s was limited by the available measurement equipment. Currently, only one stage of the MA was equipped with the adaptive tuning feature. By applying the supply current reduction to the other amplifier interstages, an overall current and power consumption decrease of more than 50 % is expected.

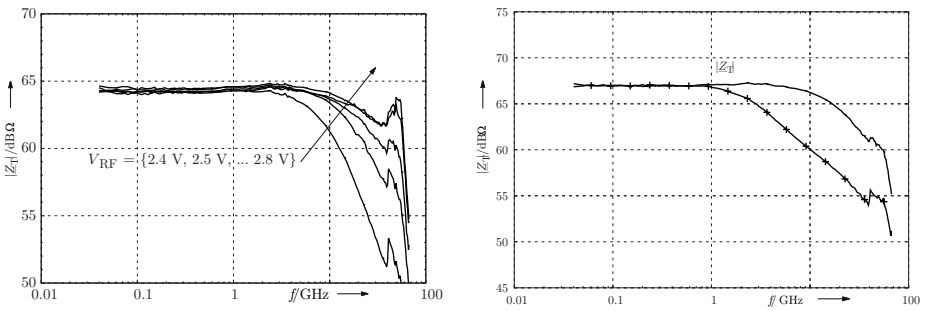

Figure 4. Frequency response of a (left) continuous and (right) discrete power and performance adaptive TIA in BiCMOS technology for different operating points.

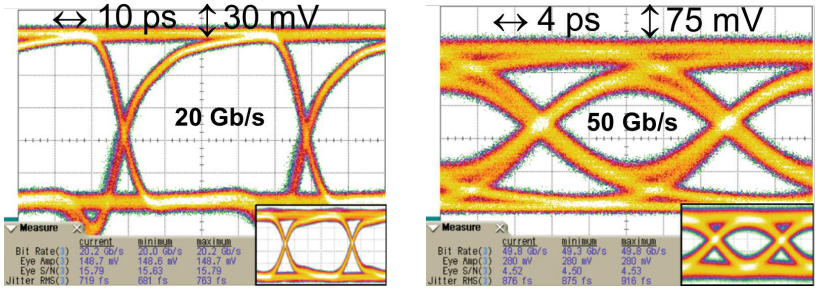

Figure 5. Electrical eye diagrams of a continuous power and performance adaptive TIA in (left) low performance and (right) high performance (limited to 50 Gb/s by measurement equipment) operating modes.

A second design in 0.13 μm SiGe BiCMOS involves a discrete adaptive receiver amplifier. Basically, its structure is similar to the continuous adaptive receiver amplifier, but without the VG option in the TIA input stage. Instead, the compensation of the gain drop due to the reduction of the bias current in the MA is realized by adjusting the load impedance of the MA, as shown on the left hand side of Fig. 3. In this case the load resistances R_D of the MA are replaced with discrete tunable impedance. In a proof of concept a network with two discrete values has been implemented in one of the MA stages to realize two operating modes: a fast-mode at 43 Gb/s and a slow-mode at 6 Gb/s, as shown on the right hand side of Fig. 4. Switching from fast to slow operating mode by reducing the MA supply currents, the overall power consumption of the receiver amplifier decreases from 74 mW to 66 mW. Thereby, the transimpedance gain remains constant at 67 dBΩ due to changing the load impedance network. Also for this discrete adaptive receiver amplifier an overall reduction of the power consumption by 50 % is expected if the currents of all amplifier interstages are decreased as well.

A similar approach is used for a further receiver amplifier realized in a 28 nm super-low power (SLP) CMOS process.[19] The current of the PD is amplified and converted into a voltage by the TIA input stage. Then the voltage signal is further amplified by a LA. In this case six differential amplifiers are used in order to achieve the necessary amplification. The current of the TIA and LA stages can be controlled by their bias currents respectively. By decreasing the currents the bandwidth and also the power consumption are reduced while also the gain drops. As an adaptive receiver must have constant output levels and hence a constant small-signal transimpedance gain regardless of the operating current and the bandwidth, this gain drop in the LA has to be compensated. Therefore, the gain is controlled by an external voltage which increases the load resistance. To show the functionality of the adaptive receiver amplifier, two operating modes, a high performance (at 30 Gb/s) and low performance (at 10 Gb/s) mode, are implemented. The transimpedance gain of the chip is measured in both modes using a vector network analyzer. As it can be seen on the left hand side of Fig. 6, the change of gain in these settings is less than 1 dB. By switching from the high performance to the low performance mode, the power consumption can be reduced by more than 50 % from 23.5 mW to 11.5 mW. The middle and right hand side charts of Fig. 6 show the error-free single-ended electrical eye diagrams at 10 Gb/s and 30 Gb/s which indicate that the gain remains almost constant while switching between the two operating modes. This

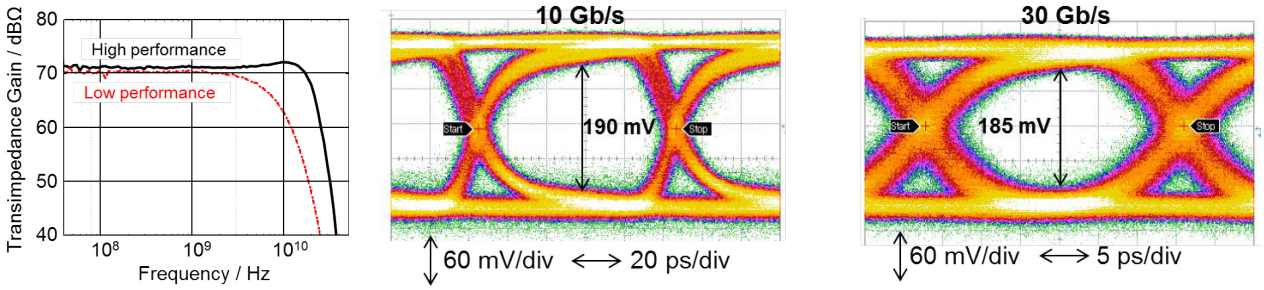

Figure 6. Adaptive TIA in CMOS with low and high performance operating mode; (left) frequency response and (right) corresponding electrical eye diagrams.

is the first power and bandwidth adaptive receiver amplifier for optical communications implemented in CMOS technology.

3.2 Adaptive laserdiode drivers

The same adaptivity approaches as described for the receiver amplifiers can be implemented into the LDDs. If a lower data load is present in the network the performance and hence the power consumption of the circuit can be reduced by decreasing the tail currents of the IC. This also decreases the output voltage and current swings of the driver. Therefore, a compensation is implemented by increasing the loads of the driver. To prove this method also in LDD, a very high-speed adaptive laser driver for direct VCSEL modulation was designed in 0.13 μm SiGe BiCMOS technology. By decreasing the currents and increasing load impedance the performance of the driver is reduced from 50 Gb/s to 16 Gb/s while the output driving capabilities remain constant. This is shown in the electrical eye diagram measurements in Fig. 7. Please note that the highest measurable error-free data rate of 50 Gb/s was limited by the available measurement equipment. As can be seen from Fig. 7 the output amplitudes of the LDD remains constant in both operating modes. However, due to the absence of the VCSEL capacitance for the electrical measurements and since the LDD can be much faster than the 50 Gb/s, no speed change can be observed in this case. By attaching the VCSEL to the driver, the performance reduction can be directly seen, as shown in the chart on the right hand side of Fig. 7. This adaptivity leads to a significant reduction of the power consumption by 46 % from 132 mW to 72 mW.

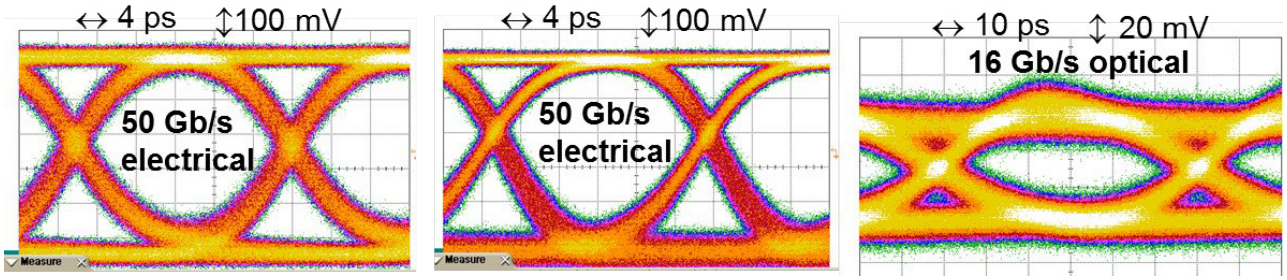

Figure 7. Eye diagrams of adaptive LDD in high and low performance operating mode; (left) high performance mode electrical, (middle) low performance mode electrical and (right) low performance mode optical. The speed change in the electrical low performance measurement cannot be observed due to the absence of the VCSEL capacitance and due to the higher performance capability of the LDD.

Another adaptive LDD was realized in 28 nm super-low power (SLP) bulk CMOS technology. Since the bandwidth bottleneck in direct modulated optical links are the optical components and especially the lasers, different methods to enhance the link bandwidth can be implemented. The 28 nm LDD for example includes an adaptive feed-forward equalizer (FFE). With the FEE the modulation signal of the laser will be pre-distorted in a

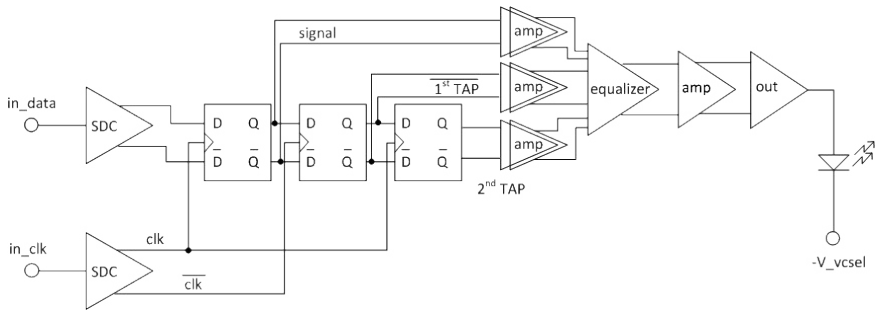

Figure 8. Block diagram of the feed-forward equalizer (FFE) laser diode driver (LDD).

way that it compensates the low-pass behavior of the VCSEL. As a result the overall (optical) link bandwidth is increased. Recently, the world record VCSEL-based NRZ data transmission with a rate of 71 Gb/s was reported using this method.[5] The block diagram of the 28 nm CMOS chip is shown in Fig. 8. The circuit includes two single-to-differential converter (SDC) input blocks used to provide the data and clock input for the three latches. The input SDCs consume 60 mW. The latches create the constant delay of half bit period in the post cursor which is necessary for the FFE. The outputs of the latches are amplified and combined in the equalizer block. The equalized signal is further amplified using two differential amplifiers "amp" and "out". The equalizer adder consumes approximately 9 mW of DC power. The output block fine tunes the bias current in the VCSEL. The DC power consumption of the entire driver is 160 mW at a maximum data rate of 30 Gb/s. When the equalizer is not needed, in case of low link workload, the circuitry of the FFE can be switched off which saves 22.5 mW of power. Neglecting the SDCs, which are not needed in the integrated system, this corresponds to a power consumption reduction by 22 %. For this the equalizer blocks can be controlled by a separate power pad. The blocks, which are going to be switched off, are two latches and four differential amplifiers. Furthermore, the current mirrors of the equalizer can be fine-tuned changing the gain of the taps in order to optimize the adaptive behavior, to investigate the functionality of a 1-tap equalizer and to adapt the equalization to different VCSELs.

4. CONCLUSION

A new method to implement a dynamic performance adaptivity into optical transmission systems has been presented in order to adjust these systems to actual workload requirements of the link and to significantly reduce the power consumption of such systems. The adaptivity is implemented, to the best of the authors' knowledge, for the first time on component and circuit level by changing primarily the bias currents of the electrical circuitry while keeping the gain of the analog stages constant. This is achieved by incorporating VG-TIA input stages and tunable resistive loads in the MAs. Furthermore, the power adaptivity and power saving potential of switching off several taps in a FFE LDD has been studied. It has been shown that the power consumption of the ICs can be reduced by up to 50 % by adjusting their performance. With further extension of the adaptivity methods to all interstages of the circuits it is expected that the overall power savings are even beyond 50 %. This significant reduction in energy consumption paves the way for dynamic ultra energy-efficient optical interconnects.

ACKNOWLEDGMENTS

The research leading to these results has been supported in part by the German Research Foundation in the framework of the Collaborative Research Center 912 "Highly Adaptive Energy-Efficient Computing" and by the European Union's Seventh Framework Programme (FP7/2007-2013) under grant agreement № 619197 in the ADDAPT project.

REFERENCES

[1] Cisco Public, "Cisco global cloud index: Forecast and methodology, 2015-2020." White Paper (2016).

[2] Kachris, C., Kanonakis, K., and Tomkos, I., "Optical interconnection networks in data centers: recent trends and future challenges," *IEEE Communications Magazine* **51**, 39–45 (September 2013).

[3] Tekin, T., Pitwon, R., Hakansson, A., and Pleros, N., eds., [*Optical Interconnects for Data Centers*], Elsevier Science & Technology, 1st ed. (2016).

[4] Doany, F. E., Lee, B. G., Kuchta, D. M., Rylyakov, A. V., Baks, C., Jahnes, C., Libsch, F., and Schow, C. L., "Terabit/sec VCSEL-based 48-channel optical module based on holey CMOS transceiver IC," *Journal of Lightwave Technology* **31**, 672–680 (Feb 2013).

[5] Kuchta, D. M., Rylyakov, A. V., Doany, F. E., Schow, C. L., Proesel, J. E., Baks, C. W., Westbergh, P., Gustavsson, J. S., and Larsson, A., "A 71-Gb/s NRZ modulated 850-nm VCSEL-based optical link," *IEEE Photonics Technology Letters* **27**, 577–580 (March 2015).

[6] Stucchi, M., Cosemans, S., Campenhout, J. V., Tökei, Z., and Beyer, G., "On-chip optical interconnects versus electrical interconnects for high-performance applications," *Microelectronic Engineering* **112**, 84 – 91 (2013).

[7] Idzikowski, F., Orlowski, S., Raack, C., Woesner, H., and Wolisz, A., "Saving energy in IP-over-WDM networks by switching off line cards in low-demand scenarios," in [*14th Conference on Optical Network Design and Modeling (ONDM)*], 1–6 (Feb 2010).

[8] Chen, X., Peh, L.-S., Wei, G.-Y., Huang, Y.-K., and Prucnal, P., "Exploring the design space of power-aware opto-electronic networked systems," in [*11th International Symposium on High-Performance Computer Architecture (HPCA)*], 120–131 (Feb 2005).

[9] Kodi, A. K. and Louri, A., "Energy-efficient and bandwidth-reconfigurable photonic networks for high-performance computing (HPC) systems," *IEEE Journal of Selected Topics in Quantum Electronics* **17**, 384–395 (March 2011).

[10] Nieweglowski, K., Rieske, R., Henker, R., Schoeniger, D., Ellinger, F., and Wolter, K.-J., "Optical interconnects for adaptive high performance computing," in [*2013 International Students and Young Scientists Workshop 'Photonics and Microsystems'*], (12-14 July 2013).

[11] Nieweglowski, K., Henker, R., Ellinger, F., and Wolter, K.-J., "Performance of step index multimode waveguides with tuned numerical aperture for on-board optical links," in [*Proceedings of SPIE Vol. 8991, Optical Interconnects XIV*], **8991**, 899105– 1–10 (February 2014).

[12] Henker, R., Pliva, J., Khafaji, M., Ellinger, F., Toifl, T., Offrein, B., Cevrero, A., Oezkaya, I., Seifried, M., Ledentsov, N., Kropp, J.-R., Shchukin, V., Zoldak, M., Halmo, L., Turkiewicz, J., Meredith, W., Eddie, I., Georgiades, M., Charalambides, S., Duis, J., and van Leeuwen, P., "Adaptive optical interconnects: the ADDAPT project," in [*Proc. SPIE*], **9662**, 966207–966207–11 (2015).

[13] Stepniak, G., Lewandowski, A., Kropp, J. R., Ledentsov, N. N., Shchukin, V. A., Ledentsov, N., Schaefer, G., Agustin, M., and Turkiewicz, J. P., "54 Gbit/s OOK transmission using single-mode VCSEL up to 2.2 km MMF," *Electronics Letters* **52**(8), 633–635 (2016).

[14] Ramirez, R. P., Agustin, M., Chorchos, L., Tonski, J., Kropp, J. R., Lendentov, N., Shchukin, V., Ledentsov, N., Henker, R., Monroy, I. T., Olmos, J. J. V., and Turkiewicz, J. P., "Effective 100 gb/s IM/DD 850 nm multi- and single-mode VCSEL transmission through OM4 MMF," *Journal of Lightwave Technology* (2016).

[15] https://ams-ix.net/technical/statistics (Jan 2017).

[16] Ellinger, F., Henker, R., Eickhoff, R., and Rieske, R., "Verfahren zur Steuerung oder Regelung eines Energiebedarfs und/oder einer Leistungsfhigkeit einer elektronischen Schaltung." European Patent EP 2762989 B1 (April 2016).

[17] Schoeniger, D., Henker, R., and Ellinger, F., "A high-speed energy-efficient inductor-less transimpedance amplifier with adjustable gain for optical chip-to-chip communication," in [*SBMO/IEEE MTT-S International Microwave and Optoelectronics Conference (IMOC)*], (Nov 2015).

[18] Schoeniger, D., Henker, R., and Ellinger, F., "High-speed transimpedance amplifier with runtime adaptive bandwidth and power consumption in 0.13 μm SiGe BiCMOS," *Electronics Letters* **52**(2), 154–156 (2016).

[19] Szilagyi, L., Schoeniger, D., Henker, R., and Ellinger, F., "Optical receiver amplifier with adaptive power and bandwidth for up to 30 Gbit/s in 28 nm CMOS," in [*11th European Microwave Integrated Circuits Conference (EuMIC)*], 105–108 (Oct 2016).