



On-chip optical interconnects versus electrical interconnects for high-performance applications[☆]



Michele Stucchi^{*}, Stefan Cosemans, Joris Van Campenhout, Zsolt Tókei, Gerald Beyer

IMEC, Kapeldreef 75, B-3001 Leuven, Belgium

ARTICLE INFO

Article history:

Available online 25 March 2013

Keywords:

On-chip optical interconnect
On-chip electrical interconnect
RC interconnect
Transmission line
Bandwidth density
Energy per bit
High performance application

ABSTRACT

In this paper, four types of on-chip point-to-point communication links are optimized for low energy per bit at sufficient bandwidth density for use in high performance applications: on-chip optical interconnect links with off-chip and with on-chip lasers, electrical full-swing RC links and electrical transmission line (TL) links. System requirements and interconnect link performance are evaluated based on the ITRS scaling roadmap. A benchmark of the link performance against system requirements indicates that full-swing RC interconnect suffices down to 12 nm. Both TL and optical interconnect can provide low-latency communication with sufficient bandwidth density and sufficiently low energy per bit until the end of the roadmap. For the system scenario studied in this work, there is no evident need to switch to the optical option.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

Potential advantages of optical interconnects over electrical interconnects have been extensively discussed in literature [1–14]. Most of the comparisons have been based on system-level analysis; the optical option is clearly beneficial for signal propagation over off-chip distances that are significantly larger than a typical die size of 1 cm and seems to become more promising for shorter distances as new low-power on-chip components become available [15].

This paper aims to provide a more up-to date, detailed and accurate evaluation of the on-chip optical interconnect option replacing global electrical interconnect. The considered application is not a full system, but a basic point-to-point, 1 cm-long link of a high performance system. This link is representative for typical data communication within ICs and allows for a direct comparison of different interconnect options without using complex system level abstractions. Two electrical options are considered for the link: full-swing RC and transmission line interconnects. Optical and electrical links are individually designed and modeled by using fine-grain circuit-level design optimization techniques. This fully optimized electrical interconnect link allows a fair comparison with the optical link in terms of propagation speed and energy con-

sumption. The metrics considered for benchmarking are the bandwidth (BW) density, expressed as Gbit/s/μm and the average energy/bit. BW density is related to the link speed per unit pitch of the link. The average energy/bit can be estimated from the energy/cycle of a random bit stream. An up-down cycle consists of a 0-to-1 and a 1-to-0 transition of the data signal. On average, a random signal makes a transition in only 50% of the clock pulses. For each stage of the link, half of these transitions take charge from the power supply, the other half dumps charge in the ground. Therefore, the average energy/bit is roughly 1/4th of the energy/cycle. The reference for the benchmarking with the proposed metrics is represented by the requirements for high performance applications expressed in BW density and average energy/bit. The benchmarking is projected in a scaling scenario derived from the ITRS Roadmap [16]. The paper is organized as follows: Section 2 describes the design, the optimization and the modeling of the electrical interconnect options; Section 3 is devoted to the on-chip optical interconnect option with both off-chip and on-chip laser source; in Section 4 the high performance requirements are derived; Section 5 presents and discusses the benchmarking results; conclusions are finally drawn in Section 6.

2. Electrical on-chip interconnects

2.1. Full-swing RC interconnect link

The full-swing RC interconnect link consists of a driver receiving an input signal and sending it into a metal interconnect with repeaters; the signal propagates in the wire modeled as distributed RC network and is finally detected by a receiver circuit. Analysis

[☆] Contains Material Reprinted with permission of the IEEE International Interconnect Technology Conference 2011.

^{*} Corresponding author. Tel.: +32 16 281681; fax: +32 16 281576.

E-mail addresses: Michele.Stucchi@imec.be (M. Stucchi), Stefan.Cosemans@imec.be (S. Cosemans), Joris.VanCampenhout@imec.be (J. Van Campenhout), Zsolt.Tokei@imec.be (Z. Tókei), Gerald.Beyer@imec.be (G. Beyer).

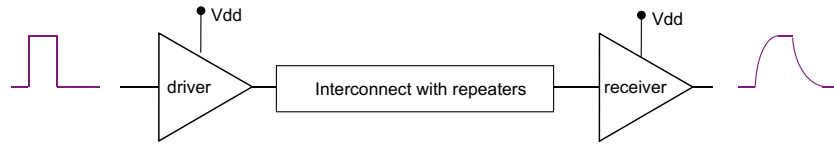


Fig. 1. Point-to-point full swing RC interconnect link with repeaters.

and optimization of RC and RLC interconnect have been extensively reported in literature; some examples can be found in [17–19]. In this paper, an RC approximation rather than an RLC model is adopted, since the fastest signal transition times are more than five times slower than the time of flight [17]. The driver is an inverter chain and the receiver is an inverter of size 4. Repeater are inserted at equal distance from each other along the interconnect length. Both the number of repeaters and the size of the repeaters are optimized to obtain the best possible combinations of bandwidth density and energy consumption per bit. The schematic of this link is presented in Fig. 1.

No pipelining is assumed for this link; the next bit is sent when the first one has been received. Hence there is only one bit of data on the interconnect link at any point in time. The use of wave pipelining could approximately double the BW without increasing the energy per bit, but this is not considered here. The RC interconnect architecture is based on Cu/low-k technology and consists of a shielded signal wire surrounded by two grounded wires in the same metal level and by two metal planes representing the metal levels above and below, Fig. 2(a). This configuration can be considered as worst-case for the parasitic capacitance. The RC link design is Pareto-optimized to provide the best combination of energy/cycle and BW density. The optimization is based on brute force spice simulations over a large parameter exploration space. The analysis is performed for a 65 nm, Low Standby power (LP) reference node.

A large parameter space is explored: wire dimensions W, H, S as indicated in Fig. 2(a), number and size of repeaters, design of the driver and supply voltage. The vertical spacing D_{top} and D_{bottom} has been fixed at 500 nm. An inverter with $W_{NMOS} = 150$ nm and $W_{PMOS} = 300$ nm is referred to as an inverter with size 1. The receiver is an inverter with size 4, thus featuring $W_{NMOS} = 600$ nm and $W_{PMOS} = 1200$ nm.

A cloud of scattered datapoints has been generated by random sampling in the exploration space. Fig. 2(b). The Pareto-optimal points for the BW density versus energy/cycle trade-off are indicated. Table 1 shows a representative set of optimal points, including the corresponding values of the variables in the exploration space.

Similar curves can be obtained for each wire length, and are plotted in Fig. 2(c). The longer the link, the higher the energy/cycle required to maintain a certain BW density. The link performance at each speed of each optimal trade-off curve is obtained with a specific combination of parameter values. For $L = 1$ cm, a BW density of 0.5 Gbit/s/um requires 0.6 pJ/cycle, hence 0.15 pJ/bit, while 1.5 Gbit/s/um requires 2 pJ/cycle, hence 0.5 pJ/bit.

High-Performance (HP) technology is more relevant for the interconnect benchmarking in the domain of high-performance applications. For this purpose, the results obtained for 65 nm have been extrapolated towards the 45 nm HP node and beyond by estimating the improvement of the transistors between the nodes based on the ITRS roadmap. The high performance system, considered in Section 5 to derive the corresponding speed and energy requirements, requires a very high bandwidth density in the order of some Gbit/s/um when scaled to very small technology nodes; therefore, the BW density is fixed at 2 Gbit/s/um in the following analysis. If the BW density would be optimized for each node separately, the resulting power consumption would be lower. The corresponding average energy/bit trend in the scaling scenario is represented in Fig. 3, starting from the 45 nm node.

2.2. Transmission line interconnect link

Transmission lines (TL) are used in high-speed off-chip interconnect, such as the memory-processor link on PCB or in USB links. TLs are based on the wave propagation of the electromagnetic field. In this paper, a coplanar waveguide is used, Fig. 4(a). The schematic of the link is shown in Fig. 5. The input signal waveform is driven onto the TL by a driver circuit; to avoid reflections, the TL is terminated after the receiver with a resistance matched to Z_0 .

The waveform propagates along the central conductor with speed $v = c/\sqrt{\epsilon_r}$, where c is the speed of light and ϵ_r is the relative dielectric constant of the material where the electromagnetic fields are confined.

The receiver is a sense amplifier (SA), which senses the voltage difference between the signal wire and the ground conductor. Signal voltage and current in TL are related by the following

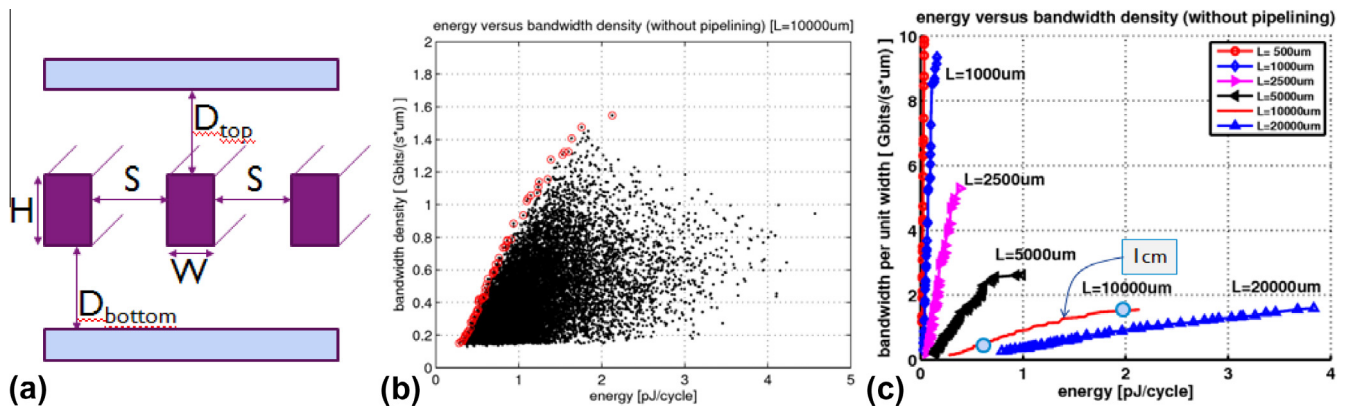


Fig. 2. (a) RC wire architecture assumed for the benchmarking; (b) Pareto-optimal speed-energy trade-off curve for $L = 1$ cm, extracted from simulations in the exploration space and (c) trade-off between BW density and energy/cycle for different interconnect lengths. Longer wires require more energy to provide the same BW; furthermore, the maximal achievable BW density reduces.

Download English Version:

<https://daneshyari.com/en/article/6943814>

Download Persian Version:

<https://daneshyari.com/article/6943814>

[Daneshyari.com](https://daneshyari.com)