

# Partition Length between Board-Level Electrical and Optical Interconnects

Azad Naemi, Anthony V. Mule, and James D. Meindl

Microelectronic Research Center, Georgia Institute of Technology, 791 Atlantic Dr. NW, Atlanta, GA, 30332, USA

Phone: (404) 894-9457 Fax: (404) 894-0462, Email: azad@ece.gatech.edu

## Abstract

The lengths beyond which board-level optical interconnects are capable of transferring a larger number of bits per second in comparison with electrical interconnects are found for different technology generations. At the 130 nm node, the partition length is 29 cm which reduces to 8.3 cm at the 45 nm because of 7 times faster drivers and 40% finer waveguide pitch.

## I. Introduction

While optical interconnection has found mainstream application in telecommunications, it is still unclear at which generation optics will be used for chip-to-chip or intra-chip communication. It is projected that as technology advances and demand for bandwidth grows, optical interconnects will be used for shorter and shorter distances. Previous quantitative analyses have identified lengths at which optical interconnects outperform electrical interconnects in terms of power dissipation and latency [1-3]. In this paper, a new partition length is identified that illustrates the length at which optical interconnects can offer a larger aggregate inter-chip bandwidth in comparison to electrical interconnects for a fixed routing area.

In this analysis, it is rashly assumed that optical drivers and receivers eventually will mature and become comparable with their electrical counterparts in terms of power, size, and cost. The comparison, therefore, is between the interconnect media or “wires versus waveguides” rather than the interface circuits. Initial opportunities to use optical interconnects appear to be more promising for chip-to-chip interconnection rather than on-chip interconnection; therefore, in this paper, board-level interconnects are considered. The partition lengths that are finally found confirm this assumption.

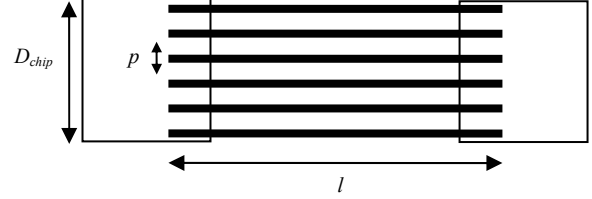
To have a fair comparison between optical and electrical interconnects, bandwidth and interconnect density should be considered simultaneously. To do so, data flux densities that electrical and optical interconnects can offer should be compared [4]. Data flux density is interconnect bandwidth per unit width of interconnect (Fig. 1). It is desired to have a large data flux density to transfer as many bits per second as possible using a constant wiring area. In this paper, data flux density is found for electrical interconnects, and it is shown which wire width maximizes data flux density. After finding maximum data flux density for optical interconnects, the lengths beyond which optical interconnects offer larger data flux density are identified.

## II. Electrical Interconnects

Off-chip or on-board wires are usually fat and work in the RLC regime, where the frequency dependent attenuation which is caused by skin effect limits the bandwidth. The bit-rate limit of a distributed RLC interconnect is rigorously found [5] as

$$B_{elect} = \frac{1}{k} B_0 \frac{P^2}{l^2}, \quad (1)$$

where  $B_0$  is a constant determined by the conductor material,  $P$  is the cross-sectional perimeter,  $l$  is the wire length, and  $k$  is a factor between 60 to 120 depending on the “eye opening” that is



$$\Phi_D = B/p \quad B_{tot} = \Phi_D \cdot D_{chip}$$

Figure 1: A schematic view of the interface between two chips which shows how data flux density,  $\Phi_D$  is defined. Data flux density is bandwidth per unit width of interconnect and for a constant wiring area it determines the aggregate bandwidth [4].

desired.  $B_0$  is  $1.846 \times 10^{18}$  (bit/s) for copper, and  $k$  is chosen equal to 120, a conservative value.

At the board level, wire thickness,  $T$ , is smaller than wire width  $W$  ( $T/W \sim 0.2-0.6$ ), and because of *proximity effect*, current flows mostly through the lower and upper regions of wires that are close to ground planes [6]. Hence, the effective perimeter of a wire can be taken as  $2W$ , and (1) can be written as

$$B_{elect} = K_0 \frac{W^2}{l^2}, \quad (2)$$

where  $K_0$  is  $6.152 \times 10^{16}$ . It should be noted that (2) is valid if all cross-sectional dimensions scale according to wire width and the line’s characteristic impedance remains constant. Data flux density, or bandwidth per unit width of interconnect, is therefore

$$\phi_{D-elect} = \frac{1}{2} K_0 \frac{W}{l^2}. \quad (3)$$

Equation (3) shows that data flux density increases linearly with increasing width. For instance, by doubling the width of a wire, the bandwidth of that wire becomes 4 times larger, but wiring density decreases by a factor of 2. In this manner, data flux density doubles. This analysis suggests that on-board electrical interconnects should be made as wide as possible to maximize the aggregate bandwidth. The maximum frequency that a driver can switch, however, introduces a limit on maximum wire width. The optimal wire width, therefore, is the width at which interconnect bandwidth becomes equal to the maximum frequency that the driver can switch,  $f_{elect-max}$ :

$$W_{opt} = \sqrt{\frac{f_{elect-max}}{K_0}} l. \quad (4)$$

By substituting (4) into (3), the maximum data flux density that electrical interconnects can offer is

$$\phi_{D-elect-max} = \frac{\sqrt{K_0 f_{elect-max}}}{2l}. \quad (5)$$

Equation (5) shows that the maximum data flux density is inversely proportional to  $l$ , not  $l^2$ . The reason is that for longer interconnects, wider wires can be used to enable a bandwidth

equal to the maximum driver frequency. It also shows that as technology advances, transistors become faster and  $f_{elect-max}$  increases, increasing in turn the maximum data flux density that can be achieved by electrical interconnects.

To estimate how the maximum data flux density changes with technology, we assume that  $f_{elect-max}$  will be equal to the International Technology Roadmap for Semiconductors (ITRS) [7] projections for the chip-to-board clock frequency. Figure 2 plots the maximum data flux density for different technology nodes for 20, 30 and 40 cm long interconnects implemented at different technology nodes. Figure 3 shows the maximum data flux density and the optimal wire width versus interconnect length at the 45 nm technology node. Figure 3 shows that as interconnect length decreases, the optimal wire width decreases, and in this manner, data flux density increases. The wire width, however, is assumed to be limited by the minimum line width resolvable on-board,  $W_{min}$ , implying that a minimum length exists below which data flux density remains constant. The value of minimum wire width is taken as 36  $\mu m$  as projected by the ITRS for the 45 nm technology node [7].

Power consumption is a big concern for GSI chips and it is important to send bits of information in a manner that minimizes energy dissipation. The energy that should be dissipated to send one bit of data can be written approximately as

$$E_b = \frac{C_p V_{dd}^2}{2} + \frac{V_{dd}^2}{R_r + Z_0} T_b, \quad (6)$$

where  $C_p$  is the parasitic capacitance associated with an I/O pad and its via,  $R_r$  is the driver resistance,  $Z_0$  is the line's characteristic impedance, and  $T_b$  is the bit duration time. The first term in (6) is the energy required to charge the parasitic pad capacitance, and the second term is the energy dissipated to transfer a pulse through the transmission line. Typically, the first term is smaller than the second term, especially for the new packaging technologies with small I/O parasitics. Equation (6) shows that to have a small energy per bit,  $T_b$  should be small, and to have valid data at the end of the line,  $T_b$  cannot be smaller than  $1/B_{elect}$ . Earlier, it was shown that increasing wire width increases the interconnect bandwidth until wire width becomes equal to  $W_{opt}$ , and interconnect bandwidth becomes equal to

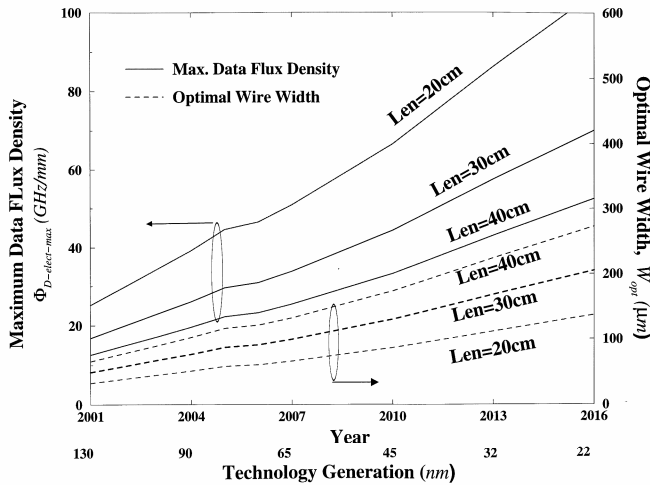


Figure 2: Maximum data flux density and optimal wire width for 20, 30 and 40 cm long interconnects implemented in different technology generations. Both maximum data flux density and optimal wire width increase in future generations because of faster transistors.

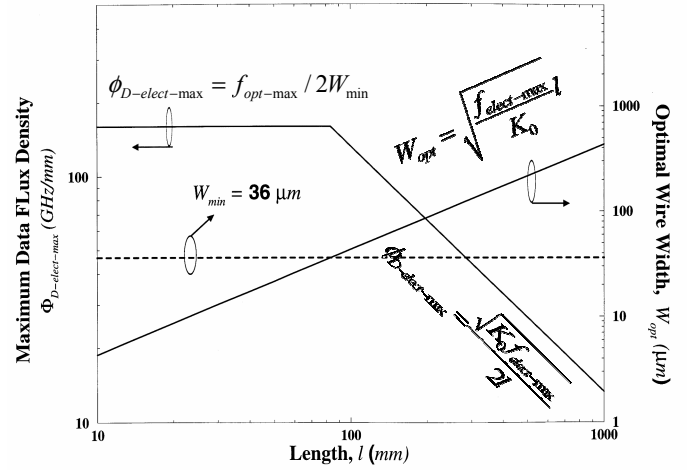


Figure 3: Maximum data flux density and optimal wire width versus interconnect length for the 45nm technology node. The dashed line shows the minimum line width resolvable on-board [7].

$f_{elect-max}$ . Hence, optimal wire width not only maximizes data flux density, but also minimizes energy-per-bit. Assuming that driver resistance is equal to the line's characteristic impedance, the minimum energy per bit is

$$E_{b-min} = \frac{C_p V_{dd}^2}{2} + \frac{V_{dd}^2}{2Z_0} \frac{1}{f_{elect-max}}. \quad (7)$$

### III. Optical Interconnects

While bandwidth of an electrical interconnect is determined by its length and cross-sectional dimensions, in practice physical dimensions of an optical waveguide have no impact on data bandwidth. Indeed, the maximum number of bits per second that an optical link can transfer is determined by the driver and receiver [2]. In this way, the maximum data flux density is

$$\phi_{D-optic} = \frac{f_{opt-max}}{p_{opt}}, \quad (8)$$

where  $f_{opt-max}$  is the maximum frequency at which optical drivers and receivers can switch, and  $p_{opt}$  is the pitch of optical waveguides. Although optical wavelength and waveguide technology influence optical crosstalk and hence the minimum value of  $p_{opt}$  [8],  $p_{opt}$  is set to  $2W_{min}$ . The maximum data flux density for the optical interconnects, therefore, is

$$\phi_{D-optic-max} = \frac{f_{opt-max}}{2W_{min}}, \quad (9)$$

where  $W_{min}$  is the minimum line width that the PWB technology permits. Figure 4 illustrates data flux density for both electrical and optical interconnects at 130 nm and 45 nm technology nodes assuming that maximum switching frequency is the same for optical and electrical transceivers.

Although arbitrary routing is relatively easy for electrical interconnects because of orthogonal routing, optical waveguides require less-complicated routing schemes to avoid lossy sharp bends and the need for inter-level communication. Although mirrors and grating couplers have been realized within a PWB [9], it is beneficial to minimize their use due to power budget constraints and cost. Hence, as Fig. 5 shows, optical I/Os should be carefully placed such that waveguides do not block one another.

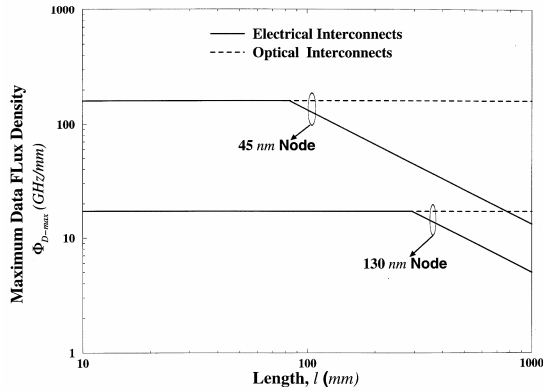


Figure 4: Maximum data flux density for electrical and optical interconnects implemented in 130 and 45 nm technology nodes.

#### IV. The Partition Length between Electrical and Optical Interconnects

Although  $f_{opt-max}$  and  $f_{elec-max}$  can be different, they cannot be completely unrelated, since an optical driver is driven by an electrical driver and an optical receiver feeds an electrical driver. In this analysis, the maximum switching frequencies for optical and electrical transceivers are assumed to be equal. A small modification would be necessary if these two frequencies are different.

By comparing (9) and (5), a partition length for electrical and optical interconnects can be found as

$$l_{part} = \sqrt{\frac{K_0}{f_{max}}} W_{min}, \quad (10)$$

which shows the interconnect length beyond which optical waveguides offer a larger data flux density compared to electrical interconnects. Figure 6 plots the partition length for different technology nodes. From Fig. 4, it is evident that the data flux density of optical waveguides is constant and larger than that of electrical interconnects for interconnects longer than  $l_{part}$ . As technology advances, the partition length decreases due to decreases in line width and increases in clock frequency. For instance, the partition length decreases from 29 cm at the 130 nm to 8.3 cm at the 45 nm because of 7 times faster transistors and 40% smaller line width.

By pushing the PWB technology and achieving minimum line widths smaller than those the ITRS has projected, smaller partition lengths can be attained and optical waveguides outperform electrical interconnects to an even greater extent. The minimum waveguide pitch, however, is limited by the crosstalk between adjacent lines. The minimum waveguide pitch constrained by crosstalk is different for various waveguide technologies. For instance, the minimum waveguide pitch for optical waveguides composed of  $\text{SiO}_2/\text{Air}/\text{MSQ}$  and operating at 632 nm wavelength is 5  $\mu\text{m}$  [8].

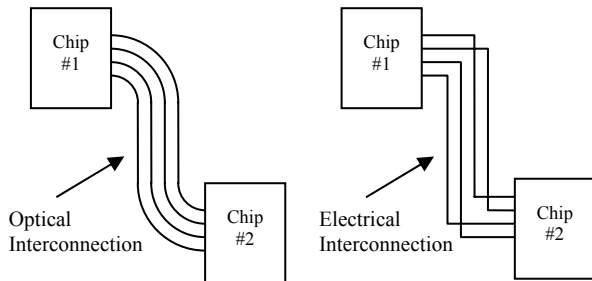


Figure 5: Unlike electrical interconnects, optical waveguides can not be routed arbitrarily because sharp bends and inter-level communication should be minimized due to power budget constraints.

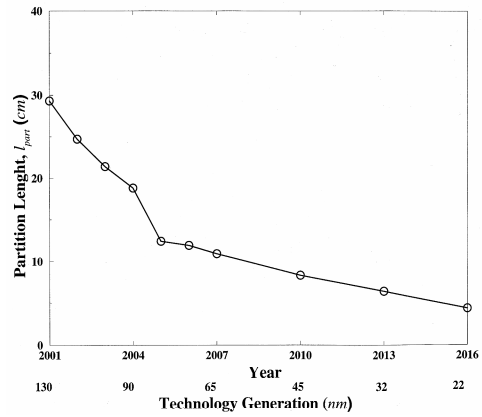


Figure 6: Partition length between optical and electrical interconnects for different generations of technology. Partition length decreases in future generations because of faster transistors and finer PWB line width.

Energy dissipation in optical interconnects is length independent and is determined by drivers and receivers. Efficiency of optical transceivers has been improved significantly and innovative devices will improve their efficiency even more in future. For each specific technology, energy-per-bit for an optical link can be compared with what (7) predicts to determine whether optical or electrical interconnects dissipate less power.

#### V. Conclusions

For different generations of technology, a partition length between electrical and optical interconnects is identified based on aggregate bandwidth constrained by a fixed routing area. For electrical interconnects, wire width is optimized such that data flux density, which is the bandwidth per unit width, is maximized. The maximum data flux density of electrical interconnects is inversely proportional to length while data flux density of optical interconnects is length independent. At the partition length, electrical and optical interconnects have equal data flux density, while for lengths greater than this length, data flux density decreases for electrical interconnects and remains constant for optical interconnects. This partition length is 29 cm for the 130 nm technology node which decreases to 8.3 cm at the 45 nm technology node. Although power dissipation in optical and electrical interconnects is not fully investigated in this paper, it is shown that the optimal wire width found for electrical interconnects not only maximizes data flux density but also minimizes energy-per-bit.

#### References:

- [1] D. A. B. Miller, "Optics for low-energy communication inside digital processors," *Optics Letters*, Jan. 1989, pp. 146-149.
- [2] A. Z. Shang and F. A. P. Tooley, "Digital optical interconnects for networks and computing systems," *Journal of Lightwave Technology*, Dec. 2000, pp. 2086-2094.
- [3] E. D. Kyriakis-Bitxaros et al., "Realistic end-to-end simulation of the optoelectronic links and comparison with the electrical interconnects for system-on-a-chip applications," *Journal of Lightwave Technology*, Oct. 2001, pp. 1532-1542.
- [4] A. Naeemi, J. A. Davis, and J. D. Meindl, "Optimal global interconnecting devices for GSI," *IEDM Tech. Digest.*, Dec. 2002, pp. 319-322.
- [5] J. Xu and R. Mooney, "Bit-rate limit of distributed RLC interconnects based on skin-effect," submitted to *IEEE Intl. Interconnect Technology Conf.*, June 2003.
- [6] M. Mizuno et al., "Clock distribution networks with on-chip Transmission lines," *IEEE Intl. Interconnect Technology Conf.* June 2000, pp. 3-5.
- [7] "International Technology Roadmap for Semiconductors," Edition 2001.
- [8] A. V. Mule et al., "Towards a comparison between chip-level optical interconnection and board-level interconnection," *IEEE Intl. Interconnect Technology Conf.*, June 2002, pp. 92-94.
- [9] R. T. Chen et al., "Fully embedded board-level guided-wave optoelectronic interconnects," *Proc. of IEEE*, Vol. 88, June 2000.