Optimization of Digital BiCMOS Circuits, An Overview

M. S. Elrabaa Student Member, IEEE, and M. I. Elmasry, Fellow, IEEE.

VLSI Research Group, Dept. of Electrical and Computer Engineering, University of Waterloo, Waterloo, Ontario, Canada, N2L 3G1.

ABSTRACT

An overview of the optimization of buffer chains and multilevel logic in a BiCMOS environment, including scaling effects, is presented. The BiCMOS speed-up contours are reviewed. The use of these contours and analytical delay expressions in the design and optimization of BiCMOS buffer chains is also reviewed. The performance differences between different types of multi-stage mixed CMOS/BiCMOS buffers are summarized. Different BiCMOS CML circuits, such as the multi-emitter BiCMOS CML circuits, are considered. The performance advantages of using such circuits in implementing multilevel logic are summarized.

I. Introduction

BiCMOS circuit designers are usually faced with the task of selecting the best combination of CMOS /BiCMOS/bipolar circuit structures for the design of critical paths that would render the optimum system performance in terms of speed, power, and area. This problem has no easy answer since the BiCMOS technology offers circuit designers an environment that is very rich with different circuit structures to implement buffer chains and logic circuits. This, in turn, complicate both the selection and design processes. In this paper, some of the recent design techniques for BiCMOS circuit design that are comprehensive and yet, easy to implement are reviewed. The BiCMOS speed-up contours are first reviewed and their use in the quick design and optimization of buffer chains is presented in section II. The effects of scaling on BiCMOS performance are shown in section III. Different implementations of multi-level logic and their relative performance are summarized in section IV.

This work is supported in part by NSERC, ITRC and MICRONET grants.

0-7803-0510-8/92$03.00 ©1992IEEE

II. BiCMOS Speed-up Contours

The contours of BiCMOS speed-up factors (*s.f.*) are defined as [1]:

$$S.f. = \frac{\textit{Optimized CMOS Buffer Chain Delay}}{\textit{BiCMOS Buffer Delay}}$$

similarly the area contours are defined as the ratio of CMOS area to BiCMOS area. The speed-up contours of the BiCMOS buffers of figure 1 are shown in figures 2 and 3. Also, on the same figures some area contours are shown. These figures show that BiCMOS speed-up over CMOS at a certain C_{in} increases with increasing C_L but it begins to decrease as C_L increases further. This is because as the number of stages in the CMOS buffer increases, its delay sensitivity to the load capacitance decreases, while the BiCMOS delay sensitivity remain constant, figure 4. However, at large C_L, the CMOS area is much greater than that of BiCMOS, figures 2 and 3. The speed-up contours could be generated using analytical expressions of CMOS and BiCMOS delays. The delay of an N stage CMOS buffer chain with a tappering factor f and input capacitance C_{in} can be expressed as :

$$T_D(CMOS) = NC_{in}(g+f)\alpha \qquad (1)$$

where g is the ratio of input capacitance to output capacitance of an unloaded CMOS inverter. α is a constant that depends on the technology and the first stage input capacitance, C_{in},

$$\alpha = \left[\frac{G_N + B_P}{K_N} + \frac{G_P + B_N}{K_P}\right]_{1st\ stage} \qquad (2)$$

where $G_{N,P}$, $B_{N,P}$ and $K_{N,P}$ are constants and their values depend on whether the devices have long channels [2] or short channels [3]. The BiCMOS delay could be fitted accurately by the following equation [1] :

$$T_D(BiCMOS) = A + BC_L \ , \quad (3)$$

$$A = \left[A_1(C_{in})^{-1} + A_2(C_{in})^{-1/2}\right] \ , \quad (4)$$

$$B = \left[\frac{A_3 + A_4(C_{in})^{-1/2}}{C_{in} + A_5(C_{in})^{1/2} + A_6}\right] \quad (5)$$

where A_{1-6} are fitting parameters. Once, the speed-up contours are generated, the designer will have an efficient way of determining which circuit will be faster for a certain C_{in}, and C_L combination. Specially if the area contours are laid over the speed-up contours, then the designer would be able to meet different speed/area design requirements. A gate C_{in} is usually determined by the driving gate speed requirements. A good example where the speed-up contours may come handy, is the design of a critical logic path. The mixing of CMOS and BiCMOS gates to meet certain design specifications becomes easier. Also the analytical delay expressions can be used to select and design optimum circuits under any design situation [1].

Multi-Stage Buffer Chains

The performance of several combinations of CMOS/BiCMOS buffer chains is summarized in table 1. The BiCMOS-BiCMOS chain has the minimum delay, however, its area is very large. Also, the BiCMOS-CMOS option is obsolete, while the CMOS-BiCMOS chain achieved a speed close to that of the BiCMOS-BiCMOS at a much smaller area, rendering this option most attractive. These results could be predicted from the speed-up contours, since for each buffer stage, the speed-up could be determined from the C_{in} and C_L (next stage C_{in}) values.

III. Scaling Effects

As the technology scales down, BiCMOS speed-up over CMOS diminishes. Figure 5 shows the cross-over capacitance (unity speed-up contours) of both types BiCMOS of buffers for three BiCMOS technologies; a $1\mu m$,$5V$ technology, a $0.6\mu m$,$3V$ technology, and a $0.2\mu m$,$2V$ technology. The cross-over capacitance C_{xov} increases rapidly with scaling, and at 2 V, the FSBiCMOS is faster than CMOS for a small range of C_{in} and C_L.

IV. Multi-Level Logic Implementations

The implementations considered are; CMOS, CMOS with a FSBICMOS as a last stage, and CML (current mode logic). As an example, a multi-level logic (MLL) consisting of cascaded 4-input NAND gates is considered. The fan_{out} for each level of logic is one except for the last level where the fan_{out} and the wire capacitance are represented by a load capacitance C_L. The CML implementation using gates similar to that (MCSL) in [4] is shown in figure 6. The speed-up of the CMOS+BiCMOS and CML implementations over CMOS for different C_L's and BiCMOS technologies are shown in figures 7 and 8, respectively. For high C_L and low number of logic levels, the CMOS+BiCMOS implementation has a good speed-up factor. However, this factor diminishes with scaling. The CML implementation has a good speed-up factor even at low C_L. This factor increases with the number of logic levels and does not suffer significantly with scaling. Another CML option is the Multi-emitter BiCMOS CML circuits [5]. This type has two structures; the Multiemitter Merged MOS/Bipolar CML (M^3BiCML) and the Low Power Multiemitter Merged MOS/Bipolar CML (LPM^3BiCML). An XOR implementation using both structures is shown in figure 9. These circuits are suitable for low voltages and can implement complex gates at smaller area.

Conclusions

The BiCMOS speed-up contours could be a very helpful tool for circuit designers a certain design situation. They could also be used in the design of buffer chains and multi-level logic. Using BiCMOS gates at high capacitance nodes in a CMOS logic can enhance the performance greatly with a small increase in area, however, at low loads and/or low voltages, CML has to be used to gain any speed-up over CMOS.

Refferences

[1] M. S. Elrabaa and M. I. Elmasry, "Design and Optimization of Buffer Chains and Logic Circuits in a BiCMOS Environment," IEEE JSSC, vol. 27, pp. 792-801, May 1992.

[2] N. Hedenstierna and K. Jeppson, "CMOS Circuit Speed and Buffer Optimization," IEEE Trans. Computer-Aided Design, Vol. CAD-6, pp. 270-281, 1987.

[3] T. Sakurai and A. Newton, "Delay Analysis of Series-Connected MOSFET Circuits," IEEE JSSC, vol. 26, pp. 122-131, 1991.

[4] W. Heimsch *et al.*, "Merged CMOS/Bipolar Current Switch Logic (MCSL)," IEEE JSSC, vol. SC-24, pp. 1307-1311, 1989.

[5] M. S. Elrabaa and M. I. Elmasry, "Multi-Emitter BiCMOS CML Circuits," IEEE JSSC, vol. 27, PP. 454-458, Mar. 1992.

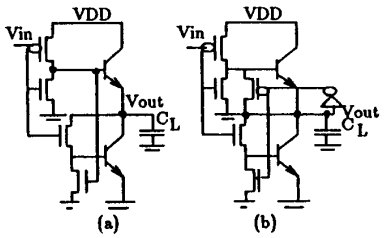

Fig.1. (a) A partial swing BiCMOS (PSBiCMOS) Buffer, (b) A full swing BiCMOS (FSBICMOS) buffer.

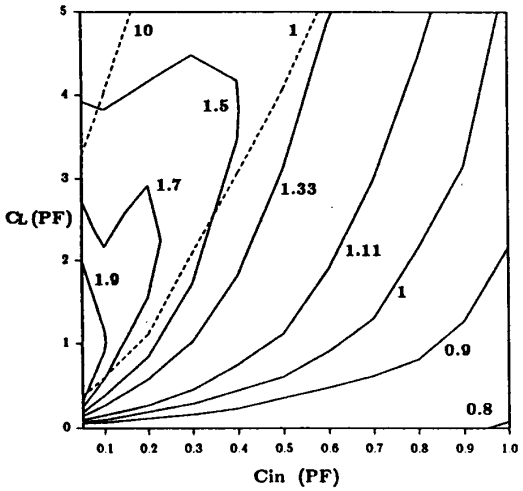

Fig.2. Speed-Up contours of the PSBiCMOS (solid lines). Also shown are some area contours (dashed lines).

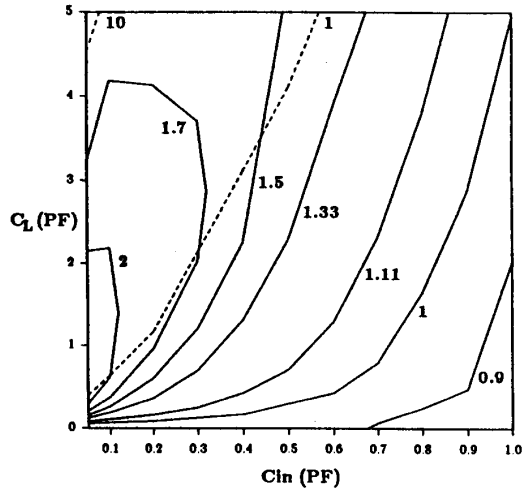

Fig.3. Speed-Up contours of the FSBiCMOS (solid lines). Also some area contours are shown too (dashed lines).

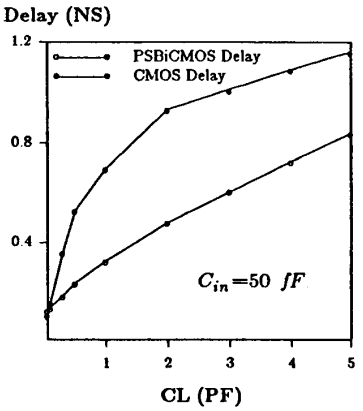

Fig.4. Delay vs load capacitance of the CMOS buffer chain and the PSBiCMOS buffer.

Table 1. The performance of the different CMOS/BiCMOS buffer chains and the single stage BiCMOS buffer.

	CMOS Chain	Single BICMOS	BICMOS-BICMOS	CMOS-BICMOS	BICMOS-CMOS
Delay (NS)	1.465	1.270	0.877	1.176	1.050
Normalized Area *	58.1	3.6	29.3	9.2	33.6

* Normalized to the area of the 1st stage of the CMOS buffer chain.

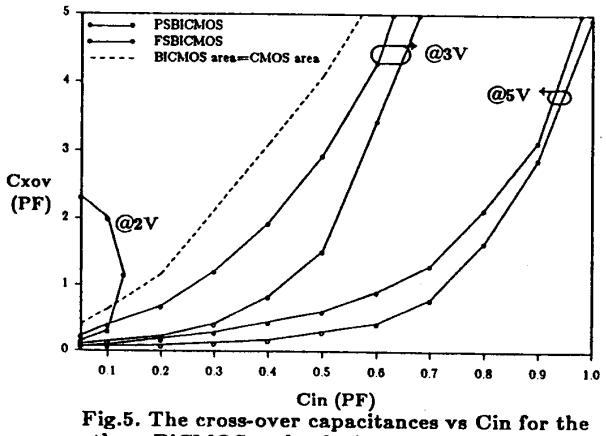

Fig.5. The cross-over capacitances vs Cin for the three BiCMOS technologies.

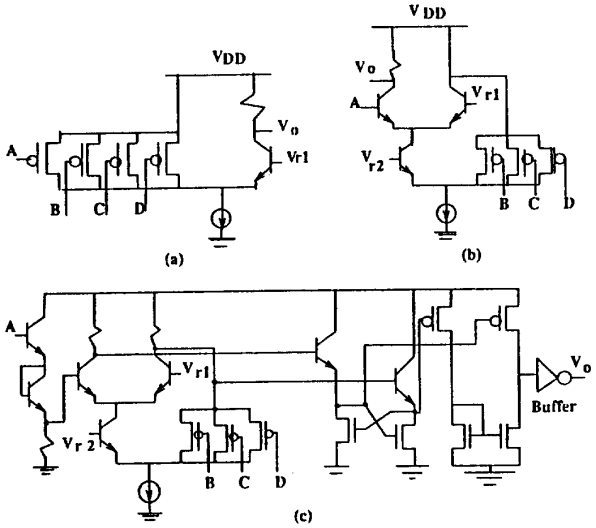

Fig.6. The CML implementation of the MLL: (a) the first gate, (b) an internal gate, and (c) the last gate with the CML/CMOS level conversion.

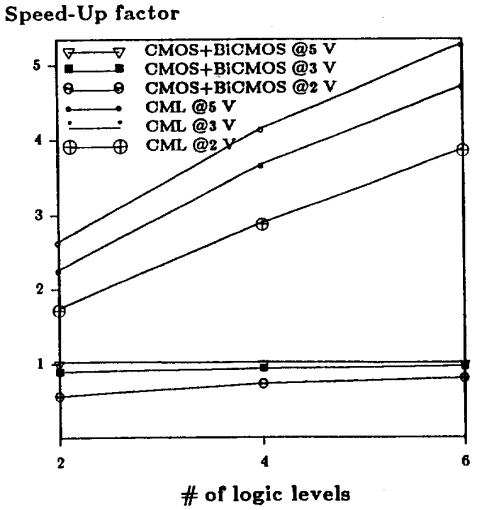

Fig.8. Speed-up factors of the MLL implementations for the three technologies and at low load capacitance.

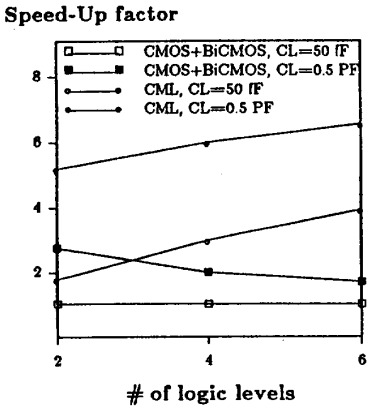

Fig.7. Speed-up factors of the MLL implementations for different load capacitances.

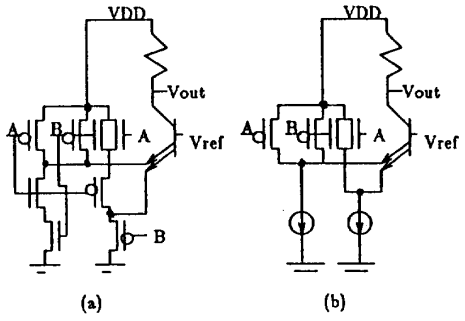

Fig.9. XOR implementation using; (a) a *LPM³BiCML* , and (b) a *M³BiCML* structures.