Optimal Interconnection Circuits for VLSI

H. B. BAKOGLU, STUDENT MEMBER, IEEE, AND JAMES D. MEINDL, FELLOW, IEEE

Abstract—**The propagation delay of interconnection lines is a major factor in determining the performance of VLSI circuits because the *RC* time delay of these lines increases rapidly as chip size is increased and cross-sectional interconnection dimensions are reduced. In this paper, a model for interconnection time delay is developed that includes the effects of scaling transistor, interconnection, and chip dimensions. The delays of aluminum, WSi_2, and polysilicon lines are compared, and propagation delays in future VLSI circuits are projected. Properly scaled multilevel conductors, repeaters, cascaded drivers, and cascaded driver/repeater combinations are investigated as potential methods for reducing propagation delay. The model yields optimal cross-sectional interconnection dimensions and driver/repeater configurations that can lower propagation delays by more than an order of magnitude in MOSFET circuits.**

I. Introduction

EFFICIENT packaging and interconnection of silicon chips are significant challenges in high-speed digital-system design. With the evolution of computers, packaging has become the most critical aspect of the central processor in a fast mainframe computer and, as a result, sophisticated packaging technologies are required to achieve high performance [1]. Once considered to be a routine task, interconnecting the transistors of a VLSI circuit is also becoming a major concern because of the long design time required to lay out the interconnections and the interconnection propagation delay that limits overall chip performance. These problems become more severe as the minimum feature size is scaled down to the submicrometer level and chip size is increased potentially to wafer dimensions. In this paper, an *RC* model is introduced that incorporates the resistance and capacitance of the interconnection, load capacitance, and resistance of the transistor that drives the line. This model is then used to calculate propagation delay and develop strategies to minimize it and also to predict future trends.

II. Scaling of Interconnection Dimensions

In ideal scaling of MOS transistors, all linear dimensions and voltages are scaled by $1/S$ and all doping levels are scaled by S, where S is the scaling factor ($S > 1$). As a result, gate delay decreases by $1/S$, power density remains the same, and the power-delay product per device is reduced by $1/S^3$ [2], [3].

Manuscript received July 2, 1984; revised September 12, 1984. H. B. Bakoglu is supported by an IBM Fellowship. This work was supported by DARPA under Contract MDA903-80-C-0432. Some of the earlier results of this work were presented at the IEEE International Solid-State Circuits Conference, San Francisco, CA, February 1984.

The authors are with the Integrated Circuits Laboratory, Stanford University, Stanford, CA 94305.

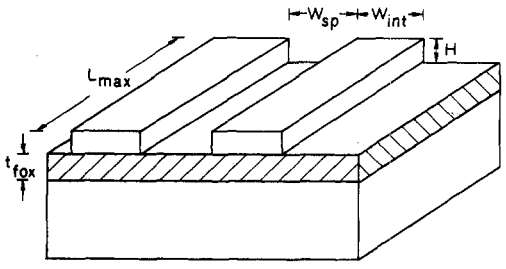

Fig. 1. Interconnection dimensions.

As device dimensions are miniaturized, the interconnection dimensions (Fig. 1) must also be reduced to take full advantage of the scaling process.

A straightforward approach to satisfying the processing and layout constraints is to scale all cross-sectional interconnection dimensions (W_{int}, W_{sp}, H, t_{fox}) by the same factor as used for transistors (ideal scaling). The devices and interconnections would then require the same relative accuracy from lithography, pattern-etching, and material-deposition technologies, and the aspect ratios of interconnections and steps will not change as sizes are reduced. The effects of ideal scaling on local and long-distance interconnections are listed in Tables I and II [4]. Although gate delay decreases by $1/S$ in ideal scaling, it can be seen that the response time of local interconnections remains the same. Even more troublesome, the delay time of long-distance interconnections (such as those extending from corner to corner on a die) increases by $S^2 S_c^2$, where S_c is the scaling factor for the chip size. The chip scaling factor is included to account for the increase in die size from one generation of integrated circuits to the next.

Deviation from ideal scaling can be advantageous. Scaling interconnections and field-oxide thickness by factors smaller than S will lower R_{int} and C_{int}. Two alternative approaches to minimizing propagation delay are also presented in Tables I and II. In "quasi-ideal" scaling of local interconnections, the horizontal dimensions are scaled by $1/S$ (like transistors) to improve overall packing density by a factor S. On the other hand, the vertical dimensions are reduced only by $1/\sqrt{S}$ to maintain a small *RC* time constant and, as a result, delay decreases by $1/\sqrt{S}$. In "constant-R" scaling, all cross-sectional dimensions are reduced only by $1/\sqrt{S}$ and, consequently, propagation delay is lowered by $1/S$. Long-distance interconnections are more difficult because of the additional burden introduced by the increasing chip size. In "constant-dimension" scaling of long-distance interconnections, all cross-sectional dimensions are held constant, and propagation delay rises by S_c^2

0018-9383/85/0500-0903$01.00

TABLE I
SCALING OF LOCAL INTERCONNECTIONS

Parameter	Ideal Scaling	Quasi-Ideal Scaling	Constant-R Scaling	Generalized Scaling
Thickness (H)	$1/S$	$1/\sqrt{S}$	$1/\sqrt{S}$	$1/S_H$
Width (W_{int})	$1/S$	$1/S$	$1/\sqrt{S}$	$1/S_{int}$
Separation (W_{sp})	$1/S$	$1/S$	$1/\sqrt{S}$	$1/S_{sp}$
Field-oxide thickness(t_{fox})	$1/S$	$1/\sqrt{S}$	$1/\sqrt{S}$	$1/S_{fox}$
Length (L_{max})	$1/S$	$1/S$	$\approx 1/S$	$1/S$
Resistance (R_{int})	S	$\sqrt{S}$	1	$S_{int}\,S_H/S$
Capacitance between line and substrate($C_{int-sub}$)	$1/S$	$1/S^{3/2}$	$\approx 1/S$	S_{fox}/SS_{int}
Capacitance between neighboring lines($C_{int-int}$)	$1/S$	$1/\sqrt{S}$	$\approx 1/S$	S_{sp}/SS_H
RC Delay (T)	1	$\approx 1/\sqrt{S}$	$\approx 1/S$	$\approx S_{int}\,S_H/S^2$
Voltage drop (IR)	1	$1/\sqrt{S}$	$1/S$	$S_{int}\,S_H/S^2$
Current density (J)	S	$\sqrt{S}$	1	$S_{int}\,S_H/S$

S : Scaling factor for device dimensions

TABLE II
SCALING OF LONG-DISTANCE INTERCONNECTIONS

Parameter	Ideal Scaling	Constant Dimension	Constant Delay	Generalized Scaling
Thickness (H)	$1/S$	1	S_c	$1/S_H$
Width (W_{int})	$1/S$	1	S_c	$1/S_{int}$
Separation (W_{sp})	$1/S$	1	S_c	$1/S_{sp}$
Field-oxide thickness(t_{fox})	$1/S$	1	S_c	$1/S_{fox}$
Length (L_{max})	S_c	S_c	S_c	S_c
Resistance (R_{int})	S^2S_c	S_c	$1/S_c$	$S_{int}\,S_H\,S_c$
Capacitance (C_{int})	S_c	S_c	S_c	$\approx S_c$
RC Delay (T)	$S^2S_c^2$	S_c^2	1	$S_{int}\,S_H\,S_c^2$

S : Scaling factor for device dimensions
S_c : Scaling factor for chip size

because of the growth in chip size. In "constant-delay" scaling, the cross-sectional interconnection dimensions are increased such that the improvement in the RC delay per unit length cancels the effect of increasing chip size and total delay remains constant.

As the ratios H/W_{int} and t_{fox}/W_{int} get larger, the two-dimensional fringing fields and capacitances between neighboring lines become important and, after a certain point, larger aspect ratios yield no additional advantage. As a result, the capacitance per unit length of the wires in a multilevel interconnectional scheme approaches a lower limit of 2 pF/cm with SiO_2 as the dielectric material; an ultimate limit of 1 pF/cm is projected with improved dielectric materials [5]. To obtain accurate results, two-dimensional fringing fields and the contribution of the neighboring lines must be included in determining the capacitance of the interconnections; neglecting these effects can introduce substantial error. Expressions that are used to calculate the capacitances in the following derivations have less than 5-percent error over a wide range of interconnection parameters; a set of design rules and interconnection dimensions is assumed, and a capacitance of 3 pF/cm is achieved independent of minimum feature size with ideal scaling. This result is in agreement with detailed analytical and computer-aided models developed to determine the capacitance of VLSI interconnections [5]-[14]. In the following calculations, conductivity is assumed to remain constant, and this is a valid assumption until the mean-free path of the carriers becomes comparable to interconnection thickness and width; however, this does not become a problem until the interconnection width is scaled down to approximately 0.1 μm [15].

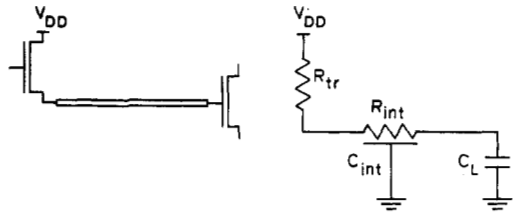

Fig. 2. Interconnection-delay model.

III. MODEL FOR INTERCONNECTION PROPAGATION DELAY

Interconnection delay is modeled as illustrated in Fig. 2. Under step-voltage excitation, the times (T) required for the output voltage of distributed and lumped RC networks to rise from 0 to 90 percent of their final values are $1.0RC$ and $2.3RC$, respectively [16]. Accordingly, a very good approximation for delay is obtained by combining the resistive and capacitive terms and weighting them by the appropriate factors as described earlier

$$T = 1.0R_{\text{int}}C_{\text{int}} + 2.3(R_{tr}C_{\text{int}} + R_{tr}C_L + R_{\text{int}}C_L) \tag{1}$$

$$T \approx (2.3R_{tr} + R_{\text{int}})C_{\text{int}}. \tag{2}$$

This approximation is in agreement with Sakurai's expression [17] reported to have less than 4-percent error over the entire range of parameters. The on-resistance of a MOS transistor is

$$R_{tr} \approx \frac{L/W}{\mu C_{gox} V_{DD}} \tag{3}$$

which remains constant with ideal scaling. Interconnection resistance is

$$R_{\text{int}} = \rho \frac{L_{\max}}{W_{\text{int}}H} \tag{4}$$

and, as shown in Table II, it increases as S^2S_c when ideal scaling is applied. The capacitance of the center one of three adjacent lines above a ground plane is expressed [7] as

$$\frac{C_{\text{int}}}{\epsilon_{\text{ox}}L_{\max}} = 1.15\left(\frac{W_{\text{int}}}{t_{fox}}\right) + 2.80\left(\frac{H}{t_{fox}}\right)^{0.222} + \left(0.06\left(\frac{W_{\text{int}}}{t_{fox}}\right) + 1.66\left(\frac{H}{t_{fox}}\right) - 0.14\left(\frac{H}{t_{fox}}\right)^{0.222}\right)\left(\frac{t_{fox}}{W_{sp}}\right)^{1.34} \tag{5}$$

with an error of less than 10 percent over a wide range of H/t_{fox}, W_{int}/t_{fox} and W_{sp}/t_{fox} ratios.

Fig. 3 plots the propagation delay for various chip dimensions as a function of minimum feature size and for two driver resistances (1 and 10 kΩ). It can be seen that the delay is dominated by $R_{\text{int}}C_{\text{int}}$ as dimensions are scaled aggressively. With the existing aspect ratios and design rules, the capacitive term is determined by the capacitance between the interconnection and substrate ($\approx \epsilon_{\text{ox}}L_{\max}\ W_{\text{int}}/t_{fox}$). In modest VLSI circuits and for aluminum lines, the resistive term is dominated by R_{tr} and, in larger VLSI circuits and for polysilicon lines, it is con-

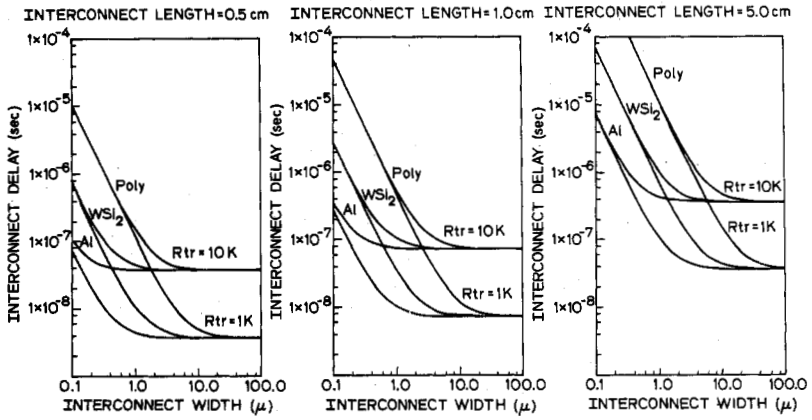

Fig. 3. Interconnection delay as a function of width for three materials (Al, WSi_2, polysilicon), two driver resistances $1k\Omega(W/L = 10)$ and $10k\Omega(W/L = 1)$, and three lengths (0.5, 1.0, 5.0 cm). It is assumed that $W_{sp} = W_{int}, H = W_{int}/3, t_{fox} = W_{int}/5$, and ideal scaling is applied. Also, $\rho_{Al} = 3\ \mu\Omega \cdot$ cm, $\rho_{WSi_2} = 30\ \mu\Omega \cdot$ cm, and $\rho_{Poly} = 500\ \mu\Omega \cdot$ cm.

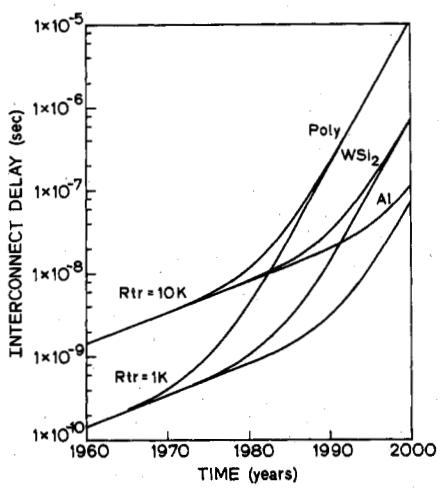

Fig. 4. Projected long-distance propagation delays of future integrated circuits with single-layer metal interconnections. Equations (7) and (8) are used to determine λ and L_{max}. It is assumed that $W_{int} = \frac{3}{2}\lambda, W_{sp} = \frac{3}{2}\lambda, H = \frac{1}{4}\lambda, t_{fox} = \frac{1}{6}\lambda$, and ideal scaling is applied.

trolled by R_{int}. The optimal interconnection dimensions for a given length can be obtained from the corner point of the curves where $R_{int} \approx 2.3R_{tr}$. The optimal cross-sectional dimensions, therefore, are

$$W_{int}H = \rho \frac{L_{max}}{2.3R_{tr}}. \tag{6}$$

The projected long-distance interconnection propagation delays of future integrated circuits with single-layer metal are plotted in Fig. 4. Here, the minimum feature size is assumed to drop by a factor of 2 every six years and chip dimension is assumed to double every eight years as

$$\lambda = 2.0 \times 2^{-(t-1983)/6}\ \mu\text{m} \tag{7}$$

$$L_{max} = 0.35 \times 2^{(t-1983)/8}\ \text{cm}. \tag{8}$$

It can be seen that, in the early 1970s, all interconnections were pure capacitive loads, and time delay was determined by the size of the driver and the length of the interconnection. As the chip dimension increases and the minimum feature size decreases, R_{int} becomes important and propagation delay rises rapidly. Today, aluminum lines still can be approximated as capacitive loads, whereas polysilicon lines have exhibited an RC behavior for some time. If the conventional methods for interconnecting the transistors are not improved, longer interconnection lengths and smaller widths and thicknesses will result in unacceptably long propagation delays. Since in the past the channel of the transistor dominated the resistive component of delay and the interconnections were virtually perfect conductors with respect to the channel and the parasitic capacitances of all components (gate, diffusion, polysilicon, aluminum) have been significant, the major challenges were to minimize these capacitances and to drive large capacitive loads. With smaller feature sizes and greater chip dimensions, the parasitic resistances of interconnections are becoming comparable to channel resistance and, as a result, the design challenges are to reduce parasitic resistances and capacitances and to drive large RC rather than large capacitive loads.

IV. Methods for Improving Propagation Delay

Two approaches to shortening interconnection delay are investigated. The first is to reduce interconnection resistance by using only aluminum lines for long-distance communication and by forming multilayers of interconnections with thicker and wider lines in the upper levels. The second is to improve the driver circuit through cascaded drivers that increase in size until the last device is large enough to drive the line and/or by using repeaters that divide the interconnection into smaller subsections [18].

A. Multilayer Interconnections

It is not possible to interconnect the transistors with only one level of aluminum without polysilicon or diffusion crossovers; moreover, even aluminum lines will introduce excessive delay when chip dimension and minimum feature size are scaled. Multilayers of interconnections are partial solutions to both

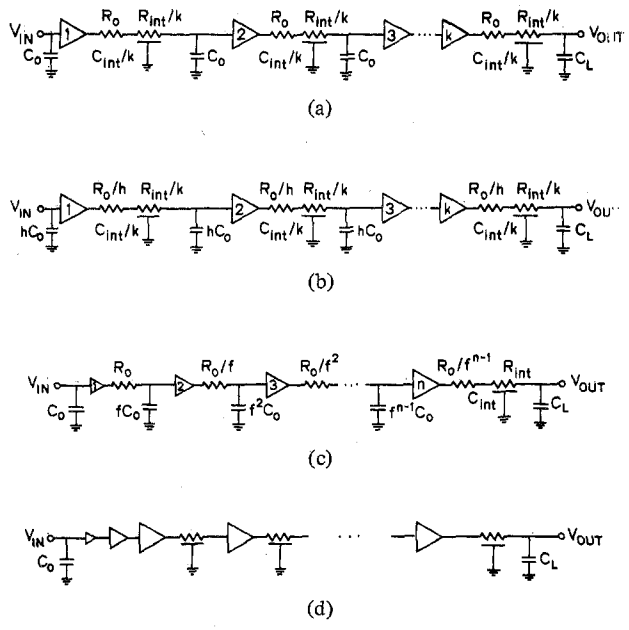

Fig. 5. Methods for driving interconnections. (a) Minimum size repeaters. (b) Optimal repeaters. (c) Cascaded drivers. (d) Optimal repeaters with a cascaded first stage.

problems. Layers of aluminum in the x- and y-directions interconnected through vias between the levels enable long-distance communication without the need for polysilicon or diffusion crossovers. Based on (6), the cross-sectional dimensions of the upper layers can also be optimized to reduce propagation delays; the local interconnections can use the scaled down first-level metal, and the long-distance interconnections can use the upper levels with wider and thicker lines that yield shorter propagation delays. Because much of the chip area is occupied by interconnections, multilayers of metal can also reduce chip size and further improve the time delay because the average interconnection length is inversely proportional to the number of levels [5], [19].

B. Repeaters

When the resistance of the interconnection is comparable to or larger than the on-resistance of the driver, propagation delay increases as the square of the interconnection length because both capacitance and resistance increase linearly with length. The use of repeaters makes time delay linear with length by dividing the interconnection into smaller subsections. Based on the model developed in the preceeding section, the propagation delay of an interconnection with k minimum-size inverters as repreaters (Fig. 5(a)) can be expressed as

$$T = k\left(2.3R_0 + \frac{R_{\text{int}}}{k}\right)\left(\frac{C_{\text{int}}}{k} + C_0\right) \tag{9}$$

where C_0 and R_0 are the input capacitance and output resistance of the minimum size inverter. By setting $dT/dk = 0$

$$\frac{R_{\text{int}}C_{\text{int}}}{k^2} = 2.3R_0C_0. \tag{10a}$$

To achieve the shortest total delay, therefore, the delay of the segments connected by the repeaters should be equal to that of a repeater. The number of repeaters is

$$k = \sqrt{\frac{R_{\text{int}}C_{\text{int}}}{2.3R_0C_0}}. \tag{10b}$$

Total delay is then expressed as

$$T = (\sqrt{2.3R_0C_{\text{int}}} + \sqrt{R_{\text{int}}C_0})^2. \tag{11}$$

For long-distance interconnections, C_{int} is on the order of picofarads and C_0 is on the order of femptofarads, and R_{int} and R_{tr} have values around kilohms; therefore, $R_0C_{\text{int}} >> R_{\text{int}}C_0$, and the delay expression can be further simplified to

$$T \approx 2.3R_0C_{\text{int}}. \tag{12}$$

As a result, repeaters can effectively "transform" the interconnection into a capacitive load.

Propagation time can be further improved by increasing the size of the repeaters because the current-drive capability of the repeaters is directly proportional to the W/L ratios of its devices. When the W/L ratios of the transistors are increased by a factor h, the output resistance and input capacitance become R_0/h and hC_0, respectively (Fig. 5(b)), and the delay expression then takes the form of

$$T = k\left(2.3\frac{R_0}{h}\left(\frac{C_{\text{int}}}{k} + hC_0\right) + \frac{R_{\text{int}}}{k}\left(\frac{C_{\text{int}}}{k} + 2.3hC_0\right)\right). \tag{13}$$

By setting dT/dk and dT/dh to zero, optimal values for k and h are obtained as

$$k = \sqrt{\frac{R_{\text{int}}C_{\text{int}}}{2.3R_0C_0}} \tag{14}$$

$$h = \sqrt{\frac{R_0C_{\text{int}}}{R_{\text{int}}C_0}} \tag{15}$$

and resulting delay expression becomes

$$T = 7.6\sqrt{R_0R_{\text{int}}C_0C_{\text{int}}} \tag{16}$$

which is shorter than the delay achieved by minimum-size repeaters because $R_0C_{\text{int}} >> R_{\text{int}}C_0$.

C. Cascaded Drivers

Just as repeaters are good in driving large RC loads, cascaded drivers are good in driving large capacitive loads. Instead of a single minimum-size driver, a chain of drivers can be used that increase in size until the last device is large enough to drive the load [20]. This is necessary because, if the load is driven by a large transistor which, in turn, is driven by a small device, the turn-on time of the large transistor dominates the delay term. The optimal delay is obtained with a sequence of drivers that increase gradually in size. When applied to interconnections (Fig. 5(c)), this method optimizes the sum of the delay caused by charging the input capacitances of the cascaded drivers and the interconnection propagation delay. Total delay is then expressed as

$$T = 2.3(n-1)fR_0C_0 + \left(\frac{2.3R_0}{f^{n-1}} + R_{\text{int}}\right)C_{\text{int}} \tag{17}$$

and, by setting dT/dn and dT/df to zero

$$f = e \tag{18}$$

$$n = \ln\left(\frac{C_{\text{int}}}{C_0}\right) \tag{19}$$

$$T = 2.3eR_0C_0 \ln\left(\frac{C_{\text{int}}}{C_0}\right) + R_{\text{int}}C_{\text{int}} \tag{20}$$

where e is the base of the nautral logarithm. Note that cascaded drivers minimize the $R_{tr}C_{\text{int}}$ term in (2), but the $R_{\text{int}}C_{\text{int}}$ term remains unchanged. As a result, this method is very useful when R_{int} is small and C_{int} and R_0 are dominant; however, it is not adequate when R_{int} is comparable to or larger than R_0. In addition, because the total time delay increases slowly with increased f, in practice a size ratio f larger than e is used to save chip area with little increase in propagation time.

As illustrated in Fig. 5(d), the first stage of the optimal-size repeaters must be a cascaded driver to lower the input capacitance of the structure and to optimize the total propagation delay when it is driven by a minimum-size transistor.

Figs. 6 and 7 compare the different driving schemes. It can be seen that repeaters eliminate the R_{int} term and that delay increases linearly with $L_{\max}$ because C_{int} is linearly proportional to length. Cascaded drivers are preferred when $L_{\max}$ is short because interconnection resistance is low for small $L_{\max}$; however, for longer $L_{\max}$, delay increases rapidly because cascaded drivers do not improve the $R_{\text{int}}C_{int}$ term. Optimal size repeaters with a cascaded first stage combine the advantages of the two and obtain the shortest delay under all conditions. The ultimate lower limit is the propagation delay of a lossless transmission line; here the transmission speed is

$$v = \frac{c}{\sqrt{\epsilon_r}} \tag{21}$$

where c is the speed of light in vacuum and ϵ_r is the dielectric constant of the medium in which the line is buried.

D. Additional Considerations

All of the preceding methods introduce area penalties to improve performance, and the models reported here do not include the possible increase in the average interconnection length and resulting complications. In addition to interconnection delay, the designer must consider the total area required to implement the circuit and the yield variation caused by process complexity introduced by some of the described methods (multilevel interconnections). After these approaches have been modified to meet the constraints of a specific IC fabrication line, they can then be included in the computer-aided design tools to optimize propagation delays automatically.

Source and drain resistances in addition to the contact resistances of MOS transistors become major concerns as minimum feature size and junction depths are scaled down [13], [21], [22]. These problems can be resolved by silicidation of the source and drain areas and by new contact materials and methods [23], [24]. These parasitic resistances can also be included in the delay model by combining them with the on-resistance of the transistor.

The interconnection resistance is larger than the value obtained in (4) because of the reduced cross-sectional area at the steps. This is dependent to the specific metal deposition system, and the model can be adjusted to include this effect by experimentally determining a factor by which $L_{\max}$ can be multiplied to obtain an effective $L_{\max}$.

As W_{sp} is reduced and H is increased the capacitive crosstalk between interconnections becomes significant. In addition, the inductive coupling increases because inductance rises logarithmically as the linewidth of the interconnection is scaled down, and magnetic couplings have a longer range than do capacitive couplings [6]. Because this may become very important in future VLSI circuits, the capacitive and inductive couplings between interconnections must be analyzed carefully as the linewidth is reduced and chip size increases.

V. Effects of Interconnection Time Delay

In the model developed in this investigation, a single transistor drives one interconnection line; however, a single transistor must often drive many signal lines. A dramatic example is the clock driver and clock lines. As chip size increases, the network of clock lines constitute a large RC load that produces excessive clock skew and sets an upper limit on clock frequency. Clock frequency determines the overall performance of the chip (cycle time, access time, instructions per second) because, despite its fast switching speed, a transistor must normally wait until the next clock cycle before it can change its state. In MSI or LSI circuits, maximum clock frequency is mainly determined by the switching speed of the gates; however, in advanced VLSI circuits, clock skew caused by interconnection delay becomes the dominant delay term and limits overall chip performance. Radical changes in design methods (such as self-timed systems) have been proposed as possible solutions to this problem [25]; however, it would be difficult to address the design complexity of VLSI circuits without a clock signal that serves as a convenient sequence and time reference [26]. To enhance performance, an intermediate approach may be a highly partitioned integrated circuit with a collection of clock signals (fastest clock for the smallest partition, slowest clock for overall chip communication). These limitations and methods will become even more significant with the advent of wafer-scale integration.

In addition, by using the model for interconnection propagation delay, the optimal chip size for a state-of-the-art multichip packaging technology can be determined. Given the characteristic parameters of chip- and package-level interconnections, the best performance will be achieved when the interconnection delay on the chip is roughly equal to the delay of the chip-to-chip interconnections. If the chips are made larger, the system will become slower; if they are made smaller, the complexity of the package will increase with no speed improvement. Chip- and package-level delays can be set equal, assuming optimal driving schemes, to obtain

$$7.6\sqrt{R_0R_{\text{int}}C_0C_{\text{int}}} = 2.3eR_0C_0 \ln\left(\frac{C_{\text{INT}}}{C_0}\right) + R_{\text{INT}}C_{\text{INT}}. \tag{22}$$

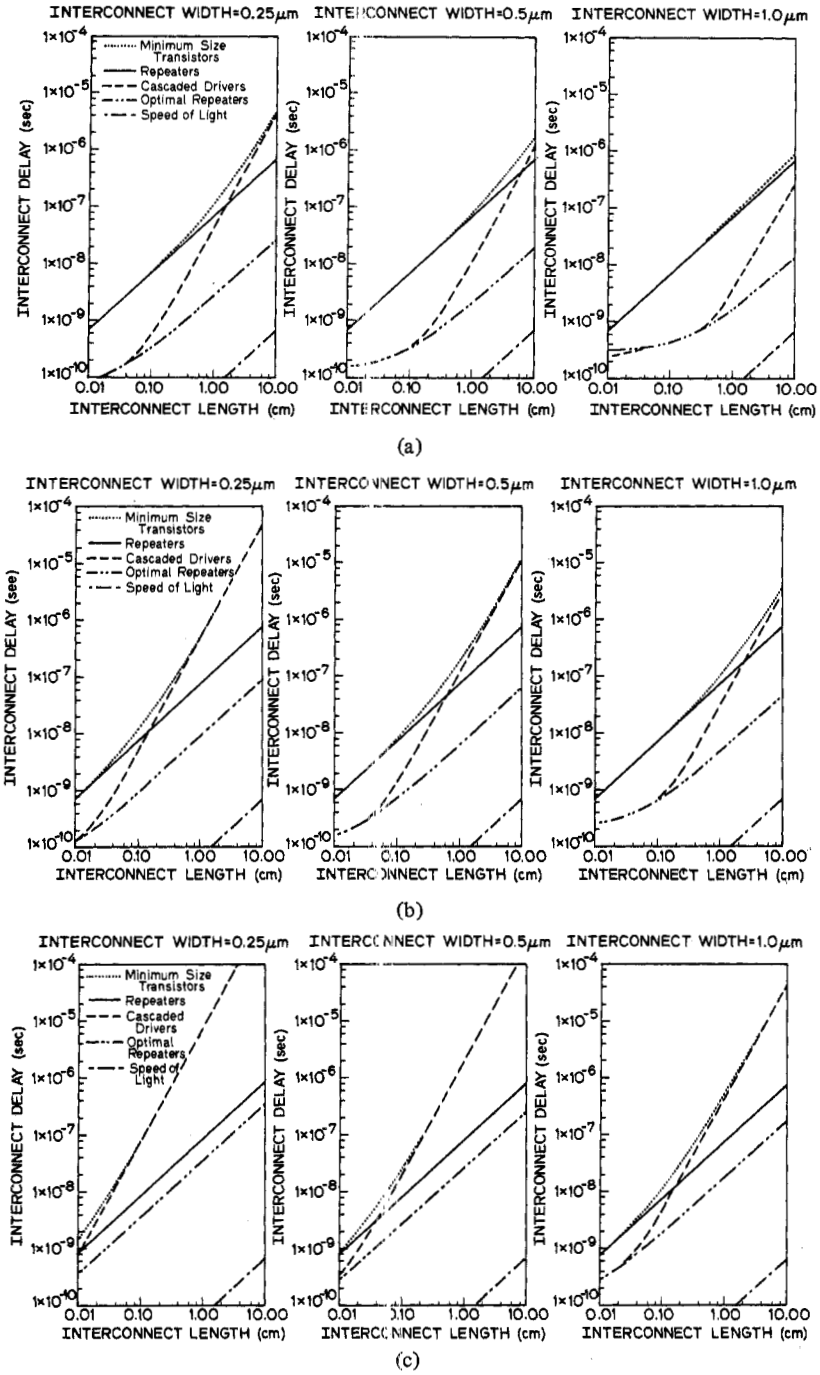

Fig. 6. Comparison of interconnection driving methods. Propagation delay is plotted for four driving mechanisms as a function of interconnection length and for three widths (0.25, 0.5, 1.0 μm). It is assumed that $W_{sp} = W_{int}$, $H = W_{int}/3$, $t_{fox} = W_{int}/5$, $t_{gox} = W_{int}/75$, and the minimum gate dimension is $2W_{int}/3$. As a result, $R_0 = 10\ k\Omega$, $C_0 = 1.17 \times 10^{-15}$ F/μm × W_{int}, $C_{int} = 3.0 \times 10^{-12}$ F/cm × L_{max}, and $R_{int} = 3.0 \times \rho L_{max}/W_{int}^2$. (a) Aluminum. (b) WSi_2. (c) Polysilicon.

Because the package level R_{INT} is usually negligible

$$\sqrt{R_{int}C_{int}} = 0.82\sqrt{R_0C_0}\ \ln\left(\frac{C_{INT}}{C_0}\right) \tag{23}$$

$$L_{max} = 0.82\sqrt{\frac{R_0C_0}{r_{int}c_{int}}}\ \ln\left(\frac{c_{INT}}{C_0}\ L_{MAX}\right) \tag{24}$$

where r_{int}, c_{int}, and c_{INT} are the resistance and capacitance per unit length, and the lower- and upper-case subscripts refer to the chip and package levels, respectively. Using $L_{max} = \sqrt{A}/2$, where A is the area of the die or package [27], the optimal chip size is obtained as a function of the package size and characteristic interconnection parameters at the chip and package levels

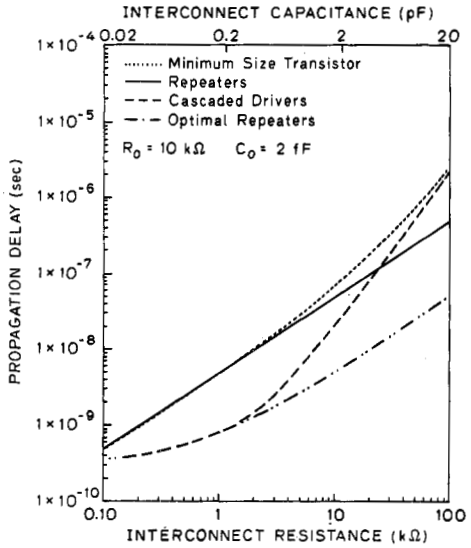

Fig. 7. Comparison of interconnection driving methods as a function of R_{int} and C_{int}.

$$\sqrt{A_{chip}} = 1.6\sqrt{\frac{R_0C_0}{r_{int}c_{int}}}\ln\left(\frac{c_{INT}}{C_0}\frac{\sqrt{A_{package}}}{2}\right). \quad (25)$$

VI. Conclusion

The propagation delay of interconnections is a major factor in determining the performance of VLSI circuits because the RC time constant of these lines increases rapidly as chip and device dimensions are scaled. This behavior affects bipolar circuits and GaAs-based IC's even more than it does MOS circuits. The performance advantage of bipolar transistors and GaAs devices over MOS transitors is their low output impedance which is significant if the load is capacitive; however, it is less important if the transistor drives an RC load with a resistance larger than the output resistance of the transistor.

Multilayers of interconnections improve propagation delay significantly and are essential for high-performance VLSI. With multilayers, all the long-distance interconnections can be formed from low-resistivity aluminum lines, and the cross-sectional dimensions of the upper layers can be adjusted to yield shorter time delays. Multilayers also occupy less chip area, which reduces interconnection length and again improves propagation delay.

Repeaters eliminate the $R_{int}C_{int}$ term in the delay expression and effectively "transform" the interconnection into a capacitive load. Cascaded drivers optimize the driving of a capacitive load, but the $R_{int}C_{int}$ term remains unchanged. Optimal size repeaters with a cascaded first stage combine the advantages of the two methods and obtain the shortest delay under all conditions. The ultimate lower limit for the time delay is set by the propagation speed of a signal in a *lossless* transmission line, and this limit is approached as parasitic resistances are eliminated.

Acknowledgment

The encouragement of P. Losleben and R. Reynolds are gratefully acknowledged.

References

[1] A. J. Blodgett, "Microelectronic packaging," *Sci. Amer.*, pp. 86-96, July 1983.
[2] R. H. Dennard, F. H. Gaensslen, H. N. Yu, V. L. Rideout, E. Bassous, and A. R. LeBlanc, "Design of ion implanted MOSFET's with very small physcial dimensions," *IEEE J. Solid-State Circuits*, vol. SC-9, pp. 256-268, Oct. 1974.
[3] VLSI Laboratory, Texas Instruments Inc., "Technology and design challenges of MOS VLSI," *IEEE J. Solid-State Circuits*, vol. SC-17, pp. 442-448, June 1982.
[4] K. C. Saraswat and F. Mohammadi, "Effect of scaling of interconnections on the time delay of VLSI circuits," *IEEE J. Solid-State Circuits*, vol. SC-17, pp. 275-280, Apr. 1982.
[5] P. M. Solomon, "A comparison of semiconductor devices for high speed logic," *Proc. IEEE*, vol. 70, pp. 489-509, May 1982.
[6] H. Yuan, Y. Lin, and S. Chiang, "Properties of interconnection on silicon, sapphire, and semi-insulating gallium arsenide substrates," *IEEE Trans. Electron Devices*, vol. ED-29, pp. 639-644, Apr. 1982.
[7] T. Sakurai and T. Tamaru, "Simple formulas for two- and three-dimensional capacitances," *IEEE Trans. Electron Devices*, vol. ED-30, pp. 183-185, Feb. 1983.
[8] A. E. Ruehli, "Survey of computer-aided electrical analysis of intergrated circuit interconnections," *IBM J. Res. Develop.*, vol. 23, pp. 626-639, Nov. 1979.
[9] A. E. Ruehli and P. A. Brennan, "Accurate metallization capacitances for integrated circuits and packages," *IEEE J. Solid-State Circuits*, vol. SC-8, pp. 289-290, Aug. 1973.
[10] A. E. Ruehli andx P. A. Brennan, "Capacitance models for integrated circuit metallization wires," *IEEE J. Solid-State Circuits*, vol. SC-10, Dec. 1975.
[11] R. L. M. Dang and N. Shigyo, "Coupling capacitances for two-dimensional wires," *IEEE Electron Device Lett.*, vol. EDL-2, pp. 196-197, Aug. 1981.
[12] W. H. Chang, "Analytical IC metal-line capacitance formulas," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-24, pp. 608-611, Sept. 1976.
[13] D. B. Scott, W. R. Hunter, and H. Shichijo, "A transmission line model for silicided diffusions: Impact on the performance of VLSI circuits," *IEEE Trans. Electron Devices*, vol. ED-29, pp. 651-661, Apr. 1982.
[14] A. K. Sinha, J. A. Cooper, and H. J. Levinstein, "Speed limitations due to interconnect time constants in VLSI integrated circuits," *IEEE Electron Device Lett.*, vol. EDL-3, pp. 90-92, Apr. 1982.
[15] A. Mayadas and M. Sahtzkes, "Electrical resistivity model for polycrystalline films: The case of arbitrary reflection at external surface," *Phys. Rev.*, B-1, pp. 1382-1389, 1970.
[16] A. Wilnai, "Open-ended RC line model predicts MOSFET IC response," *EDN*, pp. 53-54, Dec. 1971.
[17] T. Sakurai, "Approximation of wiring delay in MOSFET LSI," *IEEE J. Solid-State Circuits*, vol. SC-18, pp. 418-426, Aug. 1983.
[18] H. B. Bakoglu and J. D. Meindl, "Optimal interconnect circuits for VLSI," in *ISSCC Dig. Tech. Papers*, pp. 164-165, Feb. 1984.
[19] R. W. Keyes, "The wire-limited logic chip," *IEEE J. Solid-State Circuits*, vol. SC-17, pp. 1232-1233, Dec. 1982.
[20] C. Mead and M. Rem, "Minimum propagation delays in VLSI," *IEEE J. Solid-State Circuits*, vol. SC-17, pp. 773-775, Aug. 1982.
[21] P. I. Suciu and R. L. Johnson, "Experimental derivation of the source and drain resistance of MOS transistors," *IEEE Trans. Electron Devices*, vol. ED-27, pp. 1846-1848, Sept. 1980.
[22] C. Y. Chang, Y. K. Fang, and S. M. Sze, "Specific contact resistance of metal-semiconductor barriers," *Solid-State Electron.*, vol. 14, p. 541, 1971.
[23] D. Pramanik and A. N. Saxena, "VLSI metallization using aluminum and its alloys: Part I," *Solid State Technol.*, pp. 127-133, Jan. 1983.
[24] —, "VLSI metallization using aluminum and its alloys: Part II," *Solid State Technol.*, pp. 131-138, Mar. 1983.
[25] C. L. Seitz, "Self-timed VLSI systems," in *Proc. Caltech Conf. VLSI*, C. L. Seitz, Ed., pp. 345-355, Jan. 1980.
[26] C. Mead and L. Conway, *Introduction to VLSI Systems.* Reading, MA: Addison-Wesley, 1980, ch. 7.
[27] R. W. Keyes, "The evolution of digital electronics towards VLSI," *IEEE Trans. Electron Devices*, vol. ED-26, pp. 271-278, 1979.