A Scaling Scheme for Interconnect in Deep-Submicron Processes

K. Rahmat O. S. Nakagawa S-Y. Oh J. Moll
ULSI Laboratory, Hewlett-Packard Co., Palo Alto, CA 94304

W. T. Lynch
SRC, Research Triangle Park, North Carolina, NC. 27709

Abstract— In this paper, we study the requirements for interconnect in deep-submicron technologies and identify critical factors that will require innovations in process technology, process integration and circuit-and-system design techniques. We also propose a scaling scheme for global lines to optimize the interconnect for a given application domain such as microprocessors, ASIC's or memory. For local interconnect we demonstrate that cross-talk is the major challenge which can be addressed by selectively using larger drivers to reduce cross-talk noise when necessary.

I. Introduction

It is well known that interconnect will play a larger role in integrated circuit design and manufacture as technology is scaled [1]. This paper analyzes the requirements for interconnect in deep-submicron technologies and identifies critical factors that will require innovations in process technology, process integration and circuit-and-system design techniques.

Two regimes of interconnect are considered here: *local interconnect* which is used within a cell or a block using the finest pitch metal allowed by the design rules; *global interconnect* which is used for inter-block communication using the upper metal layers. To identify trends in interconnect and transistor technology we use the SIA (Semiconductor Industry Association) Roadmap as a guideline [2]. The key transistor and interconnect parameters that were used in this study are shown in Tables 1 and 2 and Fig. 1. The aggressive SIA roadmap targets for clock frequency, die size, metal line widths and aspect ratios raised concerns about the ability of the interconnect system to meet these targets due to cross-talk and global line delay. The purpose of this study is to understand if indeed these targets are realizable under realistic assumptions about future technologies.

All simulations were done using a SPICE circuit consisting of a driver and a wire with two neighbors, with either one (global case) or two (local case) ground planes. Coupling capacitances were obtained from a two-dimensional field solver.

II. Global Interconnect

The major criteria for interconnect design that we considered for global interconnect are cross-talk noise, electromigration and delay.

Fig. 2 shows that for future technologies the maximum interconnect length, with a fixed geometry ($2\,\mu$m wide and $1\,\mu$m thick over $1\,\mu$m oxide dielectric), that can be switched in a clock period is a decreasing fraction of the chip-side length. The maximum interconnect length is calculated based on the required clock frequency for each technology generation, as shown in Fig. 3. Of course the actual time available for simply switching a long interconnect line is much less than the clock period as the signal must go through several stages of logic. Therefore, the results shown in Fig. 2 can be considered a best case. Note that the maximum switching length cannot be increased significantly by increasing the driver size as the limitation is in the interconnect. New interconnect materials that reduce the RC time constant will reduce the delay problem, as shown in Fig. 4, but may not be sufficient nor cost-effective to fully solve it by themselves.

In the above analysis we studied the performance of interconnect assuming a fixed geometry, another approach would be to turn the question around and determine the interconnect geometry that meets all the design criteria. This approach is described below.

A. Optimization of Global Interconnect

For each layer of interconnect, we use the space of horizontal pitch (line width and spacing) and vertical pitch (line thickness plus dielectric thickness) to determine the best interconnect parameters [3]. The interconnect design space is used to plot all the constraints imposed by material, circuit performance and reliability requirements. For example, a given delay specification is represented by a contour of constant delay (see Fig. 5). All design points on or above it meet or exceed the circuit delay requirement, but process and reliability constraints would restrict the choice of the best interconnect design point. In Fig. 5, two types of reliability constraints: (1) the maximum current density to prevent electromigration and (2) the maximum cross-talk noise are shown schematically. Other processing limits such as the maximum aspect ratio can also be included on this graph. The optimal design point is defined as the point that meets all the constraints while minimizing the metal pitch. The above approach applied to $0.18\,\mu$m technology, using realistic data, is shown Fig. 6. From this figure it can be seen that the optimal interconnect design point for a 20% Vdd noise margin would be a horizontal pitch of $5.5\,\mu$m and the sum of metal and dielectric thickness of $3.2\,\mu$m.

B. Reverse Scaling for Global Interconnect

In Fig. 7, the loci of the optimal interconnect design for each generation of technology, using the SIA clock cycle, are plotted on the same axes, thus providing a guideline for reverse scaling of global interconnect. Variations such as new materials can be easily incorporated in this scheme as shown in Fig. 8, where the use of copper lines and low permittivity dielectric allows a reduction in both the vertical and the horizontal pitches by about 40%. Changes in circuit design styles that impact the lengths of critical path can also be reflected in the choice of the optimal design point by using this method.

III. Local Interconnect

For local lines (especially in array structures such as cache RAMs) a key limitation is cross-talk. Cross-talk noise will progressively become worse in the future because of the increasing coupling capacitance among neighboring lines as shown in Fig. 9. In Fig. 10, the maximum allowed wire length for a signal wire is shown for different technology generations, when the noise on a neighboring wire is limited to 20% Vdd. Also shown is the maximum wire length when the average capacitive current is limited by electromigration. From the figure it is clear that electromigration, assuming a maximum average current density limit of $2 \times 10^5\mathrm{Acm}^{-2}$, is not a critical limit for signal wires for the $0.18\,\mu$m or earlier generations. For technology generations later than the $0.18\,\mu$m, electromigration would be a limiting factor for wire lengths. In Fig. 10, we also show the impact of using two driver size scaling scenarios: in one case the ratio $W_{\mathrm{eff}}/L_{\mathrm{eff}}$ is kept constant, while in the other W_{eff} is kept constant. A larger driver has a lower resistance and thus produces less cross-talk for the same coupling capacitance. If the driver width is scaled along with the technology then the allowed interconnect length can become unacceptably short. This can only be counteracted (for minimum spacing layout) by using unscaled drivers, in which case the maximum interconnect length is longer then the scaled driver case, although there is greater power dissipation, as shown in Fig. 11.

The impact of new materials on cross-talk is shown in Fig. 12. As expected the use of copper, which reduces the wire resistance by about 40% has almost no impact on cross-talk noise because the wire resistance for lines that are shorter than about 1mm is much less than the effective transistor resistance. A low permittivity dielectric on the other hand reduces cross-talk noise almost proportionally to the reduction in the permittivity and thus can be used to drive longer lines with the same cross-talk noise tolerance.

As shown in Fig. 10, the size of the driving transistor has a dramatic impact on the cross-talk noise in a wire, therefore we can systematically trade-off cross-talk noise with driver size as shown in Fig. 13, for the $0.35\,\mu$m technology. For a given interconnect length, if less noise must be tolerated than a larger driver must be used. Hence, if a 10% Vdd noise contour is chosen then all points in the design space below and to the right are acceptable from a noise standpoint, the delay constraint in a logic design must of course also be met.

The same rule holds for all the technology generations, though the length of interconnect that can be driven for a fixed noise margin (as a fraction of the power supply voltage) is reduced for future technology generations due to the increased coupling capacitance, as shown in Fig. 14. Note that this also has implications for algorithmic scaling. For example, if a design in $0.25\,\mu$m technology is shrunk, then both the driver and wire lengths would be reduced proportionally and it would still lie on the $0.25\,\mu$m contour for 20% Vdd noise. But if the shrunk design is now processed in the $0.18\,\mu$m technology, then the design point would lie above the 20% Vdd noise contour for the $0.18\,\mu$m process and thus would fail to meet the design goal.

IV. Conclusion

In this study, we have used realistic parameters for transistors, interconnect and system performance for future technology generations to show that delay and cross talk will be severe constraints for global lines. To meet these limits we propose a scheme, which includes the impact of new materials, for optimally designing global interconnect. For local interconnect, cross-talk is the major challenge which can be addressed by selectively using larger drivers to reduce cross-talk noise when necessary.

Acknowledgement

The authors would like to thank the members of the Systems Technology Department of the Computer Systems Lab for many valuable discussions. In particular, discussions with P. Raje concerning interconnect design methodology are greatfully acknowledged. The support and encouragement of H. Stork is also much appreciated.

References

[1] S-Y Oh, K-J Chang, "2001 needs for multi-level interconnect technology", *Circuits and Devices Magazine*, January 1995.
[2] *SIA semiconductor technology – Workshop Working Group Reports*, Semiconductor Industry Association, 1994.
[3] P. Raje, "A framework for insight into the impact of interconnnect on $0.35\,\mu$m VLSI performance", *Hewlett-Packard Journal*, p. 97, February 1995.

TABLE 1 Interconnect Technology Parameters

Parameter	Technology					
	0.35um	0.25um	0.18um	0.13um	0.10um	0.07um
Local Line Width (um)	0.40	0.30	0.22	0.15	0.11	0.08
Local Line Spacing (um)	0.60	0.45	0.33	0.25	0.16	0.12
Local Line Thickness (um)	0.60	0.60	0.55	0.45	0.39	0.32
Dielectric Thickness (um)	1.00	0.84	0.70	0.59	0.57	0.50
Chip-side Length (mm)	15.8	17.3	19.0	20.7	22.8	24.9
Power Supply (V)	3.3	2.5	1.8	1.5	1.2	1.0
Frequency (MHz)	300	450	600	800	1000	1100

TABLE 2 Device Technology Parameters

Parameter	Technology					
	0.35um	0.25um	0.18um	0.13um	0.10um	0.07um
L_{eff} (um) †	0.28	0.18	0.10	0.08	0.07	0.06
Gate Oxide Thickness (A)	80	60	45	40	35	30
Threshold Voltage (V)	0.65	0.60	0.50	0.45	0.40	0.35

† Note that L_{eff} is not given in the SIA roadmap but is based on extrapolation from technology trends.

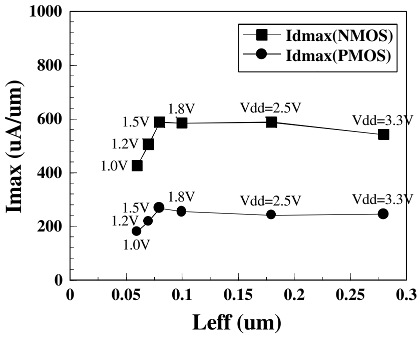

Fig. 1: Maximum drain current for NMOS and PMOS transistors of different technology generations used in this study.

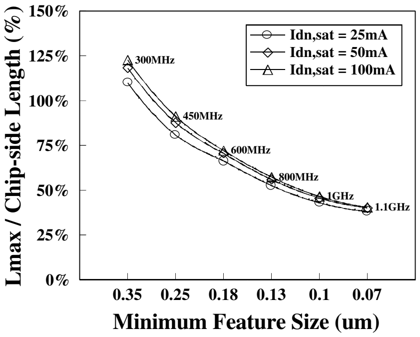

Fig.2: Maximum interconnect length / chip-side length for different driver sizes and technology generations with line width= line spacing=2um and line height=interlevel dielectric thickness=1um

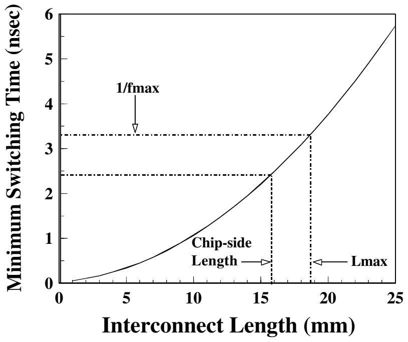

Fig. 3: Minimum switching time vs. interconnect length for 0.35um technology generation. Minimum switching time is the sum of rise and fall time. The minimum switching time = 1/f (300MHz) at Lmax.

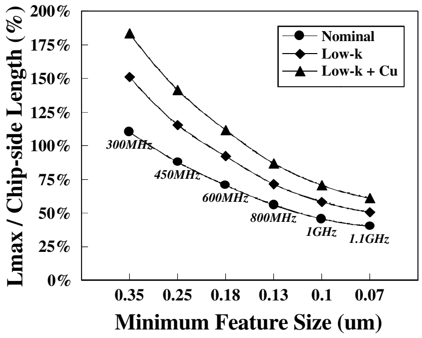

Fig. 4: Maximum interconnect length / chip-side length for new interconnect materials. Significant improvement is seen with new interconnect materials. (k=2.5)

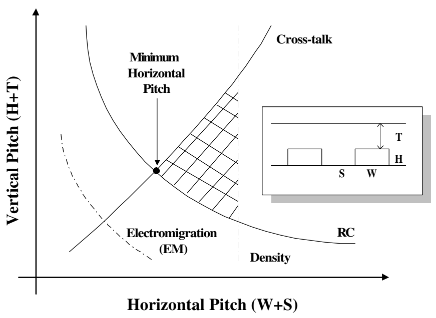

Fig. 5: Constant RC and Cross-talk noise contours in a horizontal and vertical pitch plane. H=T and W=S are assumed in the following plots for simplicity.

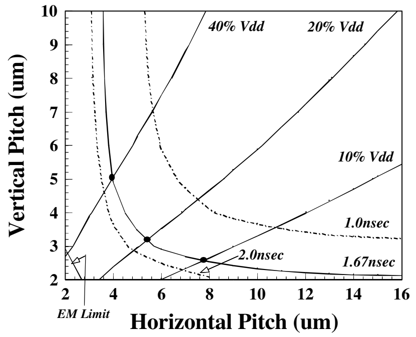

Fig. 6: Constant minimum switching time and cross-talk contours (40%, 20%, and 10% Vdd) for 0.18um technology gene-. ration. The SIA roadmap targets 1/f = 1.67nsec.

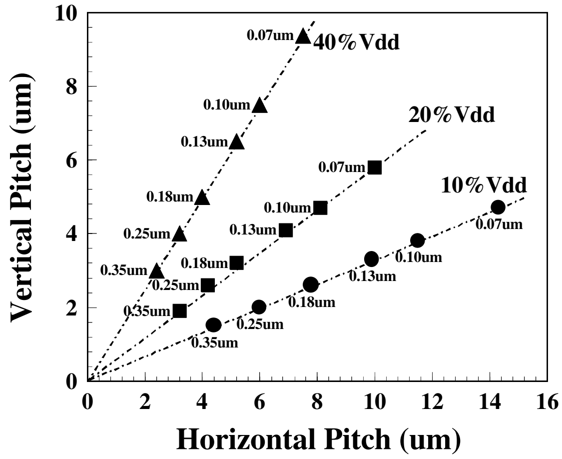

Fig. 7: Minimum-pitch design loci for chip-side length interconnect. Note that global interconnect must be scaled up to meet the SIA road map as the technology generation advances.

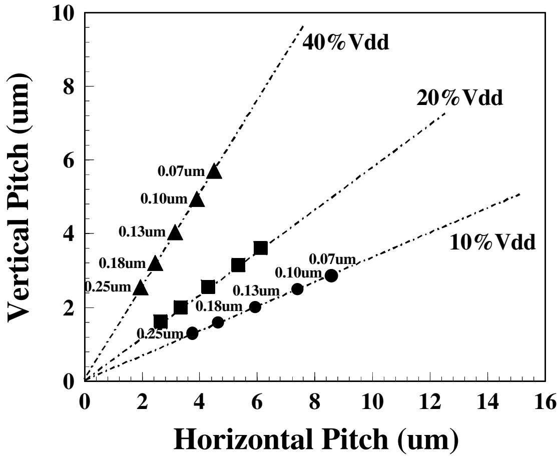

Fig. 8: Minimum-pitch design loci for chip-side length interconnect with low-k interlevel dielectric (k=2.5) and copper. Both the vertical and the horizontal pitches are approximately 40% smaller with new materials.

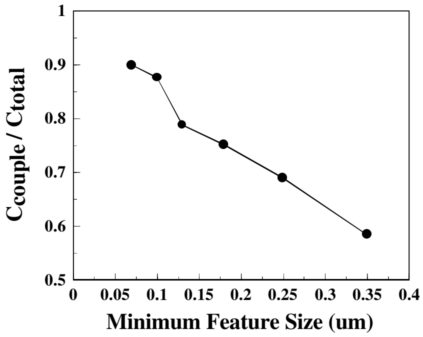

Fig. 9: Ratio of interlayer capacitance to total capacitance assuming long, uniform minimum spacing lines with ground planes above and below.

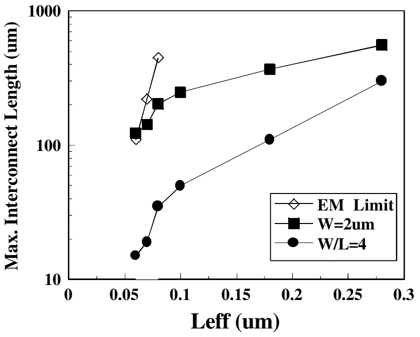

Fig.10: Maximum interconnect length for 20% Vdd cross-talk noise and the electromigration limit using a maximum average current of $2x10^5 Acm^{-2}$

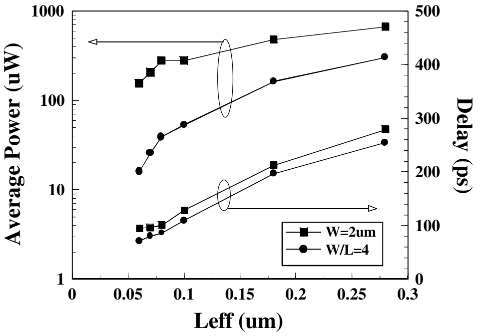

Fig. 11: Average power dissipation and delay/stage driving the max. interconnect length shown in Fig. 10. The difference in power is due to the difference in interconnect and load capacitance.

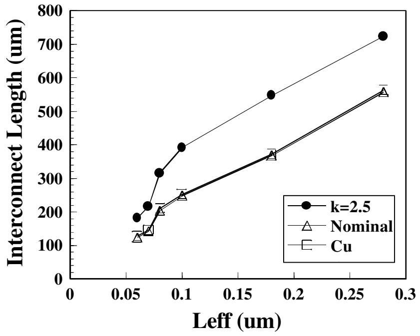

Fig12: The effect of Copper and low-k interlevel dielectric on the maximum interconnect length for cross-talk noise of 20% Vdd for a fixed driver size of 2um.

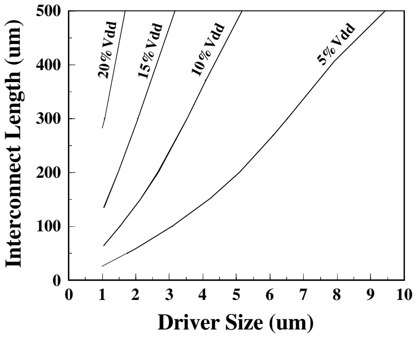

Fig. 13: Contours of constant cross-talk noise for 0.35um technology using minimum spacing lines.

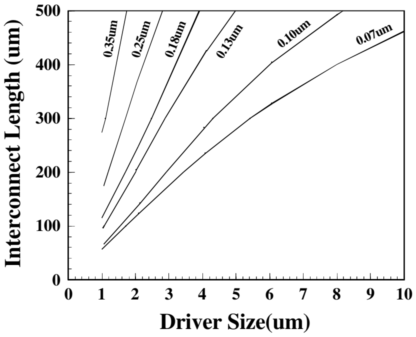

Fig.14: Contours of 20% Vdd cross-talk noise for different technology generations using minimum spacing lines.