Optimizing RC Tree Delay in High Speed ASICs Through Repeater Insertion

Victor Adler and Eby G. Friedman
Department of Electrical Engineering
University of Rochester
Rochester, New York 14627   USA
adler@ee.rochester.edu   friedman@ee.rochester.edu

Abstract

One method of overcoming wire delay due to long resistive interconnect is to insert repeaters in the line. Analytical expressions describing a CMOS inverter driving an RC load have been integrated into a global optimization algorithm for inserting repeaters into RC trees. The timing model predicts results generally within 10% of SPICE. The global optimization method exhibits total delay improvements of up to 86% over typical cascaded buffer insertion methods. The repeater timing model, global insertion methodology and algorithm, and software implementation are summarized in this paper.

I. INTRODUCTION

Interconnect delay has become a dominant limitation in high performance ASIC design. A common method of driving long interconnect is to insert a buffer at the beginning and the end of the interconnect line to improve the delay and slew rate of the signal. This method, however, does not necessarily minimize the delay caused by the large resistance encountered in long lines.

Bakoglu presents a methodology for inserting repeaters in a line to overcome the quadratic increase in delay due to a linear increase in interconnect length so that the RC interconnect impedance does not dominate the delay of a critical path [1]. Extensions to this repeater insertion methodology have also been reported in [2, 3]. In [4, 5], buffer placement methodologies for optimal interconnection based on minimizing the Elmore delay have been developed.

In this paper, the propagation delay and transition time characteristics of a CMOS repeater driving an RC line structure are presented. This timing model permits the development of a repeater design methodology and related algorithm for efficiently driving an RC tree structure, such as a clock distribution network, so as to reduce both the delay and the slew rate. In this methodology, the number and size of the repeaters to minimize the propagation delay and transition time are determined. The design expressions are based on an analytical expression derived from the α-power law model for short-channel CMOS devices [6]. The algorithm and software implementation of a global RC tree-based repeater insertion methodology are described in this paper. The efficacy of this repeater insertion methodology is also compared to a typical standard cascaded buffer methodology.

This paper is organized as follows: in Section II, the timing model for repeaters driving RC lines and branches is presented. This approach forms the basis for the methodology for determining an optimal repeater placement within an RC tree. The global repeater insertion algorithm is discussed in Section III. The accuracy and efficacy of the repeater insertion model is also discussed in this section. Finally, some concluding comments are offered in Section IV.

II. Analytical Delay Model for RC Trees

An analytical model for determining the delay and location of uniformly sized and spaced repeaters in RC trees is presented in this section [6–9]. This model assumes that the transistor operates in the linear region when driving an RC load since the linear region is the dominant region of operation when operating with fast input signals.

The structure of an RC tree is composed of a primary trunk with branching points. Each branch is modeled as a lumped resistance and capacitance, exemplified by Fig. 1. The total path delay is from the signal input at the root of the trunk to each end point of the tree (or leaf node).

The time required to drive a branch of an RC tree using uniform repeaters is

$$t_{branch} = t_{first\ stage} + (n - 2)t_{int.\ stage} + t_{final\ stage} . \quad (1)$$

The first component $t_{first\ stage}$ is the time required for the output of the first repeater in a branch to reach the turn-on voltage of the second repeater. The $t_{int.\ stage}$ component describes the time required for each repeater between the first and last stage to transition from $V_{DD} + V_{TP}$ to V_{TN} or vice versa. The last component, $t_{final\ stage}$, is the time required to
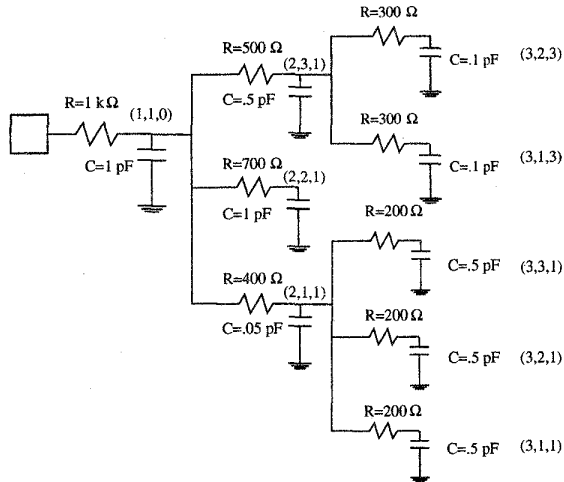
Fig. 1. An example of an RC tree. Ordered triplets (i, j, k) are used to identify specific branches (note that the downstream nodes are to the right of the upstream nodes).

Fig. 2. The total delay for a branch as a function of the number of repeaters and repeater sizes. 0.8 μm CMOS technology, $C_{rep} = 0$, $R = 1$ kΩ, $C = 1$ pF.

reach a given output voltage from either $V_{DD} + V_{TP}$ or V_{TN} [7, 8, 10]. $t_{final\ stage}$ also considers the effect of the additional capacitance C_{branch} of the downstream repeaters at a branching point.

The delay components, $t_{first\ stage}$, $t_{int.\ stage}$, and $t_{final\ stage}$, are based on an expression for the signal response of a CMOS inverter reaching an output voltage V_{out} given a step input [8],

$$t_{out} = \frac{(1 + \mho_{do}R)(C_{rep/branch} + C_{int})}{\mho_{do}} \ln\left(\frac{V_{DD}}{V_{out}}\right) . \quad (2)$$

$\mho_{do}$ is the saturation conductance, a device parameter from the α-power law model derived from $\frac{I_{do}}{V_{do}}$. I_{do} is the saturation current of the device when $V_{DS} = V_{DD}$. V_{do} is the voltage at which the device begins to operate in the saturation region [6, 8]. $C_{rep/branch}$ and C_{int} are the capacitances of the following inverting repeater and the interstage load capacitance, respectively. Additional details regarding (2) are presented in [8]. The accuracy of (2) is within 10% of SPICE simulations, with greater accuracy when the devices operate between the N and P threshold voltages.

A plot of t_{branch} derived from (1) versus the size and number of repeater stages n in a branch is shown in Fig. 2 for $C_{rep} = 0$. The optimal implementation of a repeater system for a specific RC load in terms of the number and geometric size of each repeater is represented by the minimum point on the graph. A similar graph can be drawn for any RC branch or line. Because of the difficulty in determining the derivative of (1) for each RC line or tree, the optimal number of repeaters inserted within a branch to minimize the total delay is determined, as illustrated in Fig. 2, by a numerical solution.
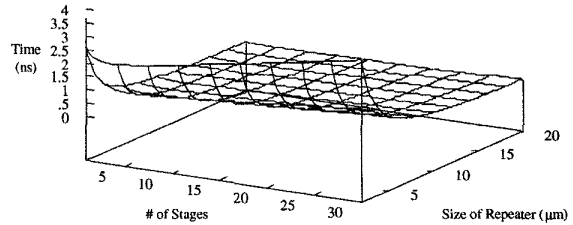
Each term in (1) is characterized by a step input to a single inverter driving an RC load, permitting a tractable solution of the delay time. This assumption permits the output waveform to be approximated by (2). The output waveform of the first stage is the input waveform of the following repeater assuming that the second repeater turns on quickly when its input threshold is reached. An example of this series of piecewise connections is shown in Fig. 3. The information describing the waveform shape permits a more accurate delay estimation as compared to estimating the path delay based on the classical Elmore delay [11]. Since the Elmore delay adds the products of a resistor (composed of the sum of the linearized model of a repeater and the interconnect resistance) and all of its downstream capacitors, the Elmore delay does not consider the shape of the output signal waveform. Thus, by integrating a more accurate timing model of the CMOS repeater into the algorithm for inserting repeaters into an RC tree, a more efficient circuit implementation can be achieved.
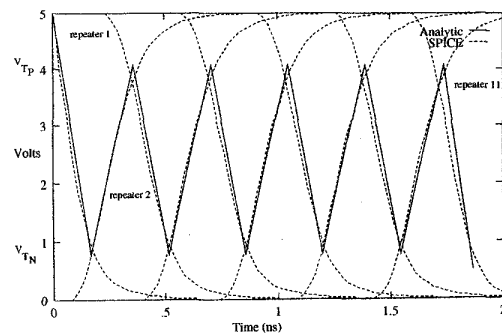

Fig. 3. The analytic and SPICE derived output waveforms of an 11-stage repeater chain driving an evenly distributed RC load of 1 KΩ and 1 pF.

III. Global Tree Repeater Insertion Algorithm

A global optimization algorithm to determine the size and number of uniform repeaters inserted within each branch of an RC tree is described in this section.

The timing model as described in Section II is used in the global optimization algorithm. The downhill simplex method of Nelder and Mead [12,13] is used to implement the multidimensional optimization algorithm.

The flow of the repeater insertion methodology for determining the optimal size and location of each repeater is shown in Fig. 4. In the downhill simplex method, each parameter variable being optimized is an element in an n-dimensional vector, $\mathbf{x}$. To insert repeaters into an RC tree, the vector $\mathbf{x}$ contains the width and number of uniformly sized and spaced repeaters in each branch. For example, in the RC tree depicted in Fig. 1, $x[1]$ is the width and $x[2]$ is the number of repeaters to be inserted into branch $(1,1,0)$. In this example, 18 elements are in $\mathbf{x}$, nine repeater widths and numbers, one pair for each of the nine branches.
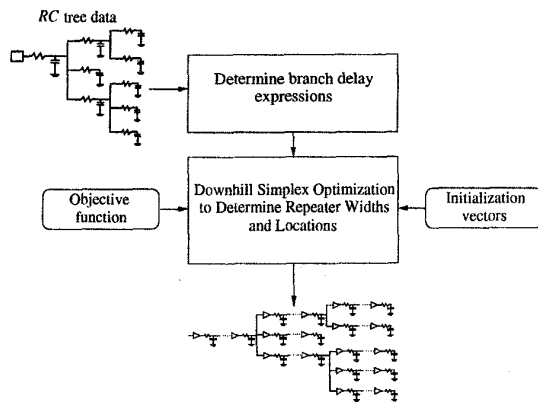

Fig. 4. A flow diagram of the methodology for globally optimal repeater insertion.

The RC tree data is converted to a set of analytical expressions describing the delays from the root node to each leaf. This set of analytical expressions, in addition to the initializing set of vectors and the objective function, are the inputs to the optimizer. In order to initialize the downhill simplex algorithm, not just one starting point, but (n+1) different arbitrary vectors are required. The n-dimensional initialization vectors are not permitted to lie along a straight line. The other input, the objective function, is the single value to be minimized. Two possible objective functions appropriate for a repeater insertion algorithm are to either minimize or specify the delay from the trunk node to the final leaf nodes as in a clock distribution network [14]. The former objective function is specified by minimizing the average delay at each leaf node while the latter objective function is achieved by minimizing the standard deviation of the predicted delay minus the target delay at each leaf node.

In the example RC tree shown in Fig. 1, the objective function is the average of the delays of each of the leaf nodes of the RC tree, which tends to minimize the delay through the trunk of the RC tree. The resulting circuit developed from the repeater insertion process is shown in Fig. 5. The repeater implementation shown in Fig. 5 does not balance the total path delay but rather minimizes the average path delay. Therefore, the minimum simplex found nearest to the starting simplex may result in different repeater sizes and spacing despite the branches being similar.

A comparison of the effectiveness of the global repeater insertion and a typical system of cascaded buffers inserted at the source of each branch is characterized in Table I for three different RC trees. The structure and impedances of the trees are described in the first three columns. In column four, the path delay for the passive RC tree with no active devices is shown. The path delay of the cascaded buffer (CB) system is shown in column five. The buffer system used for comparison is a series of optimally tapered cascaded buffers (assuming a tapering factor of three [15,16]) placed at the input of each branch so as to drive the capacitive load of each branch without considering the interconnect resistance.

The repeater insertion (RI) implementation as determined by the downhill simplex method is shown in columns six through ten. The error between SPICE simulations and the analytically derived delay is also listed and is generally within 10%. In comparing the resulting branch delays determined from the global optimization repeater insertion methodology to the cascaded buffer optimization process, the total root-to-leaf path delays are decreased between 33% to 60%. Hence the repeater insertion methodology is both effective and accurate.

IV. Conclusions

A design system for determining the optimal number and size of the uniform repeaters inserted into an RC tree is presented. An accurate timing model which considers the shape of the waveform is used within this system to achieve a more efficient repeater implementation. Analytical estimates of the total propagation delay of an example RC tree with inserted repeaters generally agree within 10% of SPICE. An algorithm is also presented for globally optimizing the size and spacing of the repeaters inserted into an RC tree in order to minimize delay. Delay improvements of 33% to 60% over a typical cascaded buffer insertion methodology are demonstrated. Thus, an integrated design system for accurately inserting repeaters into an RC tree is described in this paper.
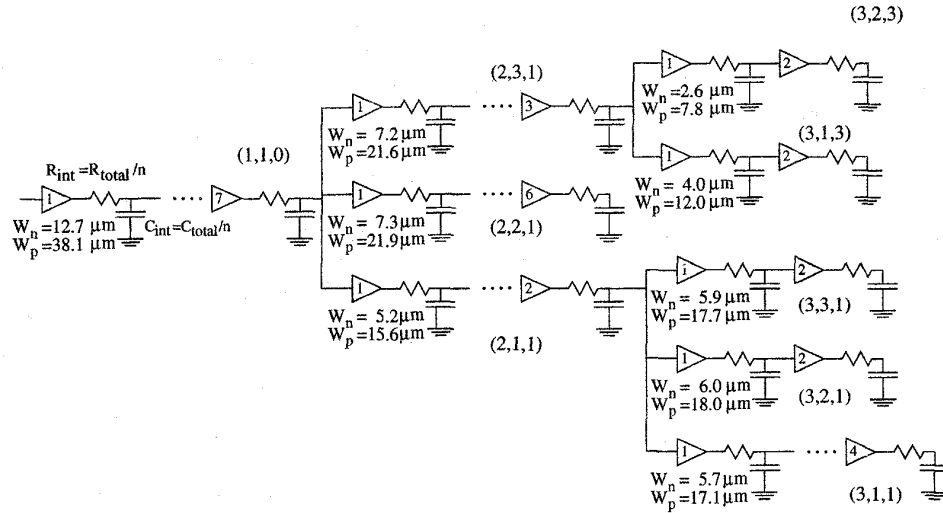
Fig. 5. The *RC* tree shown in Fig. 1 synthesized by the repeater insertion system. The transistor widths are included below the first repeater of each branch, and the number of repeaters per branch is shown inside the last repeater of each branch.

Acknowledgments

The authors would like to thank Dr. Sani Nassif of IBM, Austin Research Labs for his advice on the global optimization methodology.

REFERENCES

[1] H. B. Bakoglu, *Circuits, Interconnections, and Packaging for VLSI.* Addison-Wesley Publishing Company, 1990.

[2] C. Y. Wu and M. Shiau, "Accurate Speed Improvement Techniques for RC Line and Tree Interconnections in CMOS VLSI," *Proceedings of the IEEE International Symposium on Circuits and Systems*, pp. 2.1648–2.1651, May 1990.

[3] M. Nekili and Y. Savaria, "Optimal Methods of Driving Interconnections in VLSI Circuits," *Proceedings of the IEEE International Symposium on Circuits and Systems*, pp. 21–23, May 1992.

[4] L. P. P. P. van Ginneken, "Buffer Placement in Distributed RC-tree Networks for Minimal Elmore Delay," *Proceedings of the IEEE International Symposium on Circuits and Systems*, pp. 865–868, May 1990.

[5] J. Lillis, C.-K. Cheng, and T.-T. Y. Lin, "Optimal Wire Sizing and Buffer Insertion for Low Power and a Generalized Delay Model," *IEEE Journal of Solid-State Circuits*, Vol. SC-31, No. 3, pp. 437–446, March 1996.

[6] T. Sakurai and A. R. Newton, "Alpha-Power Law MOSFET Model and its Applications to CMOS Inverter Delay and Other Formulas," *IEEE Journal of Solid-State Circuits*, Vol. SC-25, No. 2, pp. 584–594, April 1990.

[7] V. Adler and E. G. Friedman, "Repeater Design to Reduce Delay and Power in Resistive Interconnect," *Proceedings of the IEEE International Symposium on Circuits and Systems*, pp. 2148–2151, June 1997.

[8] V. Adler and E. G. Friedman, "Delay and Power Expressions for a CMOS Inverter Driving a Resistive-Capacitive Load," *Analog Integrated Circuits for Signal Processing*, Vol. 14, No. 1/2, pp. 29–40, September 1997.

[9] V. Adler and E. G. Friedman, "Repeater Design to Reduce Delay and Power in Resistive Interconnect," *IEEE Transactions on Circuits and Systems: II – Analog and Digital Signal Processing*, Vol. CAS-45, No. 5, pp. 607–616, May 1998.

[10] V. Adler and E. G. Friedman, "Repeater Insertion to Reduce Delay and Power in RC Tree Structures," *Proceedings of the Asilomar Conference on Signals, Systems, and Computers*, November 1997.

[11] W. C. Elmore, "The Transient Response of Damped Linear Networks with Particular Regard to Wideband Amplifiers," *Journal of Applied Physics*, Vol. 19, No. 1, pp. 55–63, January 1948.

[12] J. A. Nelder and R. Mead, "A Simplex Method for Function Minimization," *Computer Journal*, Vol. 7, No. 1, pp. 308–313, 1965.

[13] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes in C, The Art of Scientific Computing.* Cambridge University Press, 1988.

[14] E. G. Friedman, *Clock Distribution Networks in VLSI Circuits and Systems.* IEEE Press, 1995.

[15] R. C. Jaeger, "Comments on 'An Optimized Output Stage for MOS Integrated Circuits'," *IEEE Journal of Solid-State Circuits*, Vol. SC-10, No. 3, pp. 185–186, June 1975.

[16] B. S. Cherkauer and E. G. Friedman, "Design of Tapered Buffers with Local Interconnect Capacitance," *IEEE Journal of Solid-State Circuits*, Vol. SC-30, No. 2, pp. 151–155, February 1995.

378

THE SIZE AND NUMBER OF REPEATERS AS DETERMINED BY THE DOWNHILL SIMPLEX ALGORITHM FOR THREE DIFFERENT RC TREE TOPOLOGIES. (t_{PD} IS THE PROPAGATION DELAY IN NANOSECONDS, # IS THE NUMBER OF REPEATERS IN A BRANCH, SIZE IS THE GEOMETRIC WIDTH OF THE N-CHANNEL DEVICE OF THE UNIFORM REPEATER FOR THAT BRANCH, AND THE P-CHANNEL TO N-CHANNEL RATIO IS 3:1.)

Branch	R	C	t_{PD} Passive	t_{PD} Buffers (CB)	Downhill Simplex (RI)					RI vs. CB Improvement
					t_{PD} Analytical	t_{PD} SPICE	Error	#	Size μm	
(1,1,0)	1 kΩ	1 pF	9.0	1.95	.88	.93	5%	7	12.7	52%
(2,1,1)	400Ω	.05 pF	9.7	1.73	1.14	1.15	1%	2	5.15	34%
(3,1,1)	200Ω	.5 pF	9.75	2.41	1.53	1.61	5%	4	5.74	33%
(3,2,1)	200Ω	.5 pF	9.75	2.41	1.51	1.57	4%	2	6.02	35%
(3,3,1)	200Ω	.5 pF	9.75	2.41	1.51	1.58	4%	2	5.89	35%
(2,2,1)	700Ω	1 pF	9.45	2.98	1.67	1.76	5%	6	7.32	41%
(2,3,1)	500Ω	.5 pF	9.28	2.46	1.35	1.45	7%	3	7.19	41%
(3,1,3)	300Ω	.1 pF	9.3	2.57	1.48	1.53	3%	2	4.05	40%
(3,2,3)	300Ω	.1 pF	9.3	2.57	1.52	1.59	4%	2	2.62	38%
(1,1,0)	700Ω	.8 pF	7.95	1.36	.72	.85	15%	7	16.39	38%
(2,1,1)	100Ω	.5 pF	7.98	1.70	.98	1.08	9%	2	8.37	36%
(2,2,1)	200Ω	.7 pF	8.18	1.86	1.05	1.14	8%	3	17.12	39%
(3,1,2)	700Ω	.6 pF	8.44	3.02	1.56	1.61	3%	5	9.60	47%
(3,2,2)	100Ω	.1 pF	8.19	2.14	1.13	1.20	6%	2	6.38	44%
(2,3,1)	1kΩ	1.6 pF	9.70	3.87	1.79	1.80	1%	11	17.22	53%
(3,1,3)	300Ω	.5 pF	9.79	3.72	2.10	2.11	0%	2	11.32	43%
(3,2,3)	600Ω	.1 pF	9.74	3.39	1.94	1.93	1%	2	5.63	43%
(1,1,0)	200Ω	5 pF	5.73	2.14	.82	.86	5%	8	75.37	60%
(2,1,1)	1 kΩ	1 pF	9.34	3.62	1.72	1.77	3%	8	15.39	51%
(3,1,1)	400Ω	.8 pF	9.54	3.95	2.19	2.22	1%	5	10.86	44%
(3,2,1)	1.5 kΩ	.1 pF	9.43	3.45	1.9	1.95	3%	3	3.06	43%
(3,3,1)	1.5 kΩ	.1 pF	9.43	3.45	1.9	1.95	3%	3	3.06	43%
(3,4,1)	400Ω	.8 pF	9.54	3.95	2.19	2.22	1%	5	10.79	44%
(2,2,1)	2 kΩ	.5 pF	7.45	3.61	1.69	1.72	2%	8	7.5	53%
(3,1,2)	800Ω	.2 pF	7.55	3.57	1.99	2.01	1%	4	4.66	44%
(3,2,2)	800Ω	.2 pF	7.55	3.57	1.99	1.96	2%	3	5.06	45%
(2,3,1)	2 kΩ	.5 pF	7.45	3.61	1.70	1.74	3%	8	7.19	53%
(3,1,3)	800Ω	.2 pF	7.55	3.57	1.99	1.97	1%	3	5.37	45%
(3,2,3)	800Ω	.2 pF	7.55	3.57	1.99	1.97	1%	3	5.28	45%
(2,4,1)	1 kΩ	1 pF	9.34	3.62	1.75	1.77	1%	8	14.83	52%
(3,1,4)	400Ω	.8 pF	9.54	3.95	2.32	2.50	7%	8	14.00	37%
(3,2,4)	1.5 kΩ	.1 pF	9.43	3.45	2.02	1.96	3%	3	3.04	43%
(3,3,4)	1.5 kΩ	.1 pF	9.43	3.45	2.03	1.96	3%	3	3.01	43%
(3,4,4)	400Ω	.8 pF	9.54	3.95	2.19	2.42	10%	4	10.36	39%