RC on-chip interconnect Performance revisited

Philippe Maurine, Nadine Azemard, Daniel Auvergne

RC on-chip interconnect Performance revisited

P. Maurine, N. Azémard, D. Auvergne

LIRMM, UMR CNRS/Université de Montpellier II, (C5506),161 rue Ada, 34392 Montpellier, France

pmaurine, azemard, auvergne@lirmm.fr

Abstract- The delay of on-chip interconnect wiring is having an important influence on the timing performance of logic path. This is particularly true where drivers are connected through non-negligible length of wire. If the Elmore resistance-capacitance delay model remains popular due to its simple formulation, limitations have been shown in sub-micrometer domain due to the inability of a so simple model in capturing input slope effects.

This paper presents an analytical expression for the transition time and the switching delay of an RC interconnect, including the line input and output drivers. Based on a previously developed model of inverter transition time and switching delay, we propose a model of the shielding capacitance effect on the input driver. We determine the transition time of the output driver and the switching delay of the complete structure for various sets of line parameters and different size of input drivers. We validate these analytical expressions with respect to electrical simulations, on a 0.13µm process, using the ELDO's transmission line model.

I. INTRODUCTION

Due to the increased integration density allowed by today ultra deep sub micrometer process the contribution of RC interconnect loading is increasing, mostly for inter block communications where drivers may be overloaded by long interconnect lines. As feature sizes are scaling down, increasing the circuit density, the gate delay is getting faster and the RC delay longer, due to higher value of the wire resistance and fringing capacitance. As a result the RC interconnect contribution delay increases with respect to the gate delay. Consequently, physical design optimization tools must be able to efficiently compute the delay across a huge number of wires and interconnect levels.

Circuit simulators such as SPICE, ELDO or model-order reduction techniques such as AWE [1,2] can be used to compute delays with a good accuracy, but are too expensive for evaluation at the optimization phase of the physical design. At this level, closed form delay equations or metrics for the delay are preferable. The most widely applied interconnect delay metric is the Elmore delay metric [3] applied to a lumped interconnect model. The accuracy of this model is not sufficient for actual processes, it does not capture the effect on delay of the transition time of the edge controlling the interconnect line and ignores the gate reduction delay induced by the resistive shielding of downstream capacitance [4].

Extending the Elmore's based work, several authors have proposed metrics based on higher circuit moments [5]. Main concern in these works is to model the delay on the interconnect line, modeling the input driver as a voltage generator supplying the line through an equivalent resistance, with no consideration, except in [6], on the interconnect slew rate.

The operating mode of an interconnect line is to propagate a signal from the output of a transmitter (line input driver) to the input of the receiver (line output driver). The important parameters, for the corresponding path, are the total delay value between the transmitter input and the receiver output, together with the transition time value of the signal at the receiver output. The accurate determination of the output transition time is of fundamental importance, it has a non-negligible contribution to the delay of the subsequent gate. These parameter values depend on the design of the drivers and the structure of the wiring.

The primary contribution of this work is to present an analytical modeling of the interconnect delay and transition time, including the contribution of the input and output drivers. We focus on RC lines and establish the limiting conditions in using purely capacitive or RC representation of interconnect wire. The remainder of the paper is organized as the following. Section II provides background material on this study. Section III summarizes the delay and transition time models previously defined. Section IV presents the equivalent line model, considered in this evaluation, with emphasis on the modeling of the resistance shielding effect. Experimental results, and validations are given in Section V, before to conclude in Section VI.

II. BACKGROUND

Addressing the delay modeling in interconnect line implies to consider the propagation of signals from the input of the driver supplying the line, to the output of the driver at the termination of the line. This is illustrated in Fig.1, where the different performance terms are defined.

The fundamental problem is then to accurately determine:

- the switching delay, $T_{OUT} - T_{IN}$, (Fig1) between the input and the termination driver,
- the output transition time value, $T_{OUTDRIVE2}$ at the termination driver.

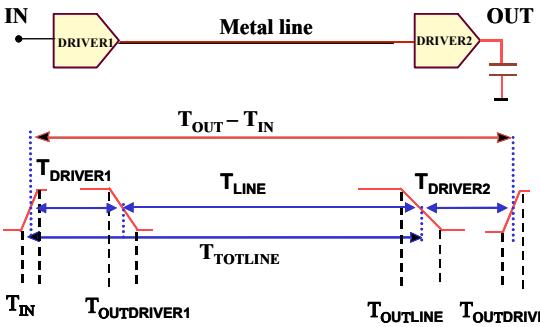

Fig.1. Identification of the different delay terms to be considered.

The first parameter ($T_{OUT} - T_{IN}$), completely characterizes the performance of the line while the second one ($T_{OUTDRIVE2}$) gives the sensitivity of the output driver to the line performance, which may produces great variations in the output driver timing. This sensitivity is illustrated in Table 1 that shows examples of the value of the different line performance parameters simulated on the general structure illustrated in Fig.1. The simulations are obtained using the ELDO's transmission line model (LDTL) [7], for a 0.13µm process. The line parameters are R_{LINE}=115 Ω/mm, C_{LINE}=472fF/mm, they correspond to a higher level of metal used for long-range interconnection. The output driver is a minimum size driver (C_{IN} = 0.82 fF), loaded by 5fF. The input driver is controlled by a short duration input ramp, three different sizes of input drivers are considered. We have limited the length line to 5mm that is longer, in high performance designs, than the maximum length allowed between repeaters.

T_{xx} specify the 50% delay and T_{OUTxx} the output transition time, determined on the output edge between 40 and 60% of its variation and extrapolated on the 100% range. The different terms are defined in Fig.1. As shown these parameters are strongly dependent on the length line and the size of the input driver. Moreover the transition time of the input driver ($T_{OUTDRIVER1}$) is far to be proportional to the line length as it could be expected. This is an evidence of the load shielding effect. The transition time value of the output driver ($T_{OUTDRIVER2}$) exhibits a great sensitivity to the transition time at the output of the line ($T_{OUTLINE}$). As shown despite of its constant loading the DRIVER2 transition time value can be multiplied by 4 in the considered range of line length. This gives evidence of the necessity to accurately determine the total switching delay (T_{OUT}-T_{IN}) and the output driver transition time that are the fundamental parameters to be determined for any timing evaluation or optimization on a path involving the RC interconnect.

For that different sub-models must be clearly established:

- the switching delay and transition time of an inverter (gate) and its sensitivity to the input controlling ramp,
- the loading and propagation effect produced by the interconnection line.

Concerning the input driver, its timing performance depends on its size and the equivalent load seen from its output. This completely determines the transition time of the signal edge supplied to the line. The problem is then to calculate the time spent by the line driver output edge to propagate across the line, and the evolution of this edge transition time. As shown in Table I, he value of the line output transition time contributes in a non-negligible part to the switching delay and transition time ($T_{OUTDRIVER2}$) of the terminal driver.

III. INVERTER MODEL

The modeling of the transition time and switching delay of inverters (gates) has been the subject of numerous works. A realistic delay model must be input slope dependent and distinguish between falling and rising signals [8]. We use a model developed for deep sub-micrometer process [9], in which the switching inverter (gate) is considered as a current generator supplying the output load. The elementary switching process of a CMOS structure can then be resumed to an exchange of charge between the structure and its output loading capacitance. The output transition time (defining the input transition time of the following cell) can then be directly obtained from the modeling of the charging (discharging) current that flows during the switching process of the structure and from the amount of charge ($C_{LTOT}.V_{DD}$) to be exchanged with the output node as

$$T_{OUTHL} = \frac{C_{LTOT} \cdot V_{DD}}{I_{NMax}} \qquad T_{OUTLH} = \frac{C_{LTOT} \cdot V_{DD}}{I_{PMax}} \tag{1}$$

where V_{DD} is the supply voltage value and C_{LTOT} the total output load, including the parasitic and input-to-output coupling capacitance [8]. In this expression the output voltage variation is supposed linear and the driving element considered as a current generator.

The key point here is to determine the realistic value of this current. Two controlling conditions are considered.

The fast input control range, in which the input signal reaches its maximum value before the output begins to vary, in this case the switching current exhibits a constant and maximum value. The transition time expression for a fast input control condition is easily obtained as:

$$T_{OUTHL}^{Fast} = \tau \cdot (1+k) \cdot \frac{C_{LTOT}}{C_{IN}} \qquad T_{OUTLH}^{Fast} = \tau \cdot \frac{(1+k)}{k} \cdot R \cdot \frac{C_{LTOT}}{C_{IN}} \tag{2}$$

where τ is a unit delay characterizing the process, $k=W_P/W_N$ the inverter configuration ratio, R is the speed ratio between electron and hole, C_{LTOT} the purely capacitive load and C_{IN} the inverter input capacitance.

In the slow input range the input and output voltages vary simultaneously. The duration of the input ramp applied to the N and P transistors decreases the current value available at the output. This is the main input slope effect on the transition time. Considering the symmetry property of the current wave shape, this new value of the current has been shown input transition time dependent resulting in [9]

$$T_{OUTHL}^{Slow} = \sqrt{\frac{(V_{DD} - V_{TN})}{V_{DD}}} \cdot \sqrt{T_{OUTHL}^{Fast} \cdot \tau_{INLH}} \quad (3)$$

with an equivalent expression for a rising edge. $V_{TN,P}$ represents the threshold voltage value of the N,P transistors.
In a realistic switching delay evaluation, consideration must be given to the finite value of the input controlling edge. As developed in [8] we introduce the input slope effect and the related input-to-output coupling in the model as:

$$T_{HL,LH}(i) = \frac{v_{TN,P}}{2} T_{INLH,LH}(i-1) + (1 + \frac{2C_M}{C_M + C_L}) \frac{T_{OUT\ HL,LH}(i)}{2} \quad (4)$$

where $v_{TN,P}$ are the reduced values ($V_{TN,P}/V_{DD}$) of the threshold voltage of the N,P transistors. $\mathbf{T}_{INHL,LH}$ is the duration time of the input signal. C_M is the coupling capacitance between the input and output nodes and C_L the output loading capacitance. Indexes (i), (i-1) specify the switching and the controlling gates, respectively.
Note that the value of each parameter strongly depends on the transition time value of the signal edge applied to the input of the device under consideration.

IV. LINE PERFORMANCE MODEL

The time and space variation of the voltage on a wire described as a RC distributed line is governed by the well-known diffusion equation that has no closed form solution [10].
An equivalent lumped RC model can be used with a satisfactory accuracy if the equivalent capacitive and resistive components are clearly determined. As observed in Table I, it has been shown in [5] that when the driver equivalent resistance value becomes comparable or smaller than the line resistance, part of the load capacitance is shielded from the line driver, resulting in large discrepancies in evaluating its output transition time and switching delay. For that let us consider a line modeled by distributed RC elements.

If the time constant of the line is smaller than some value to be determined, let us say the driver transition time, it is easy to understand that the driver equivalent output load is only constituted by the total line capacitance. In that case an accurate 50% delay prediction can be obtained as:

$$T_{OUT} - T_{IN} = T_{TOTLINE} + T_{DRIVER2} \quad (5)$$

where $T_{TOTLINE}$ is the 50% propagation delay between the input driver and the output of the line. The input driver propagation delay is given by (4) and the line propagation delay can be approximated by $R_{Line}.C_{Line}/2$, [11]. The delay across the line structure is the sum of the driver and the interconnect delay. Moreover, the transition time value at the input and output of the line are identical. This is illustrated in Fig.2 that represents the simulated wave shapes at the input (Driver1 output) and output of a 0.5mm long line in the configuration given in Fig.1.

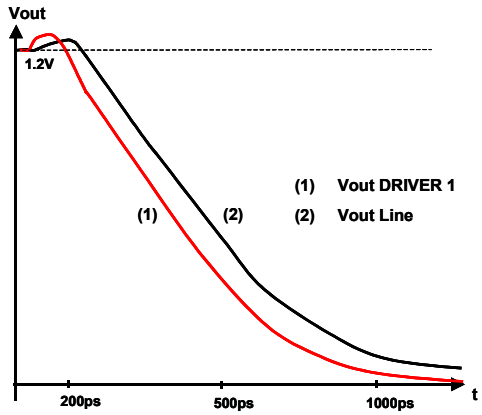

Fig.2. Simulated output wave shapes at the DRIVER1 output (Drive C) and the line termination for a 0.5mm line in the configuration given in Fig.1.

In this configuration $T_{OUTDRIVER1} = T_{OUTLINE} = 50$ ps. As shown the wave shapes are parallel, the line input driver see the total line capacitance and the driver and line output transition times can easily be deduced from the simple capacitance model (2).
If the resistance of the line has a sufficiently important value such as the line propagation delay becomes greater than the output transition time of the input driver (DRIVER1), a shielding effect can be observed. Its equivalent loading capacitance becomes smaller than the total line capacitance. This can be observed in Fig.3 that represents, for a 1.5mm long line, the evolution of the output wave shapes given in Fig.2. In this case $T_{OUTDRIVER1}$ = 185 ps, $T_{OUTLINE}$ = 1352 ps. As shown the wave shapes are no more parallel, the driver output edge is much faster than the line output one. Moreover the output transition time of the input driver is reduced from 625ps (12.5 times the value for the 0.5mm long line) to 185 ps, giving evidence of the shielding effect introduced by the line resistance.
This result can be understood, considering that the current supplied by the input driver firstly load an equivalent capacitance (part of the line capacitance)

of which loading voltage supplies the rest of the line.

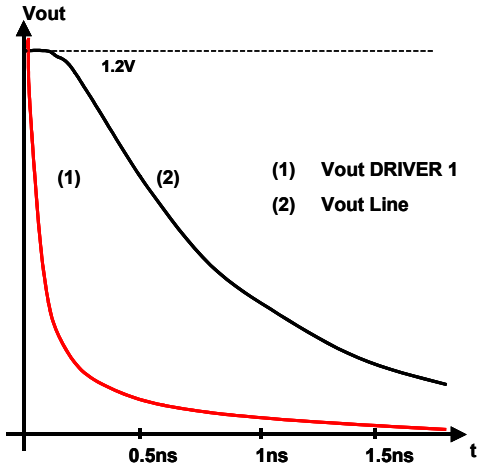

Fig.3. Simulated output wave shapes at the DRIVER1 output (Drive C) and the line termination for a 1.5mm line in the configuration given in Fig.1.

We get the superimposition of two mechanisms:

- a constant current loading of the capacitance C_{LINE} (1-x),

- diffusion of this loading on the rest of the line $R_{LINE}.C_{LINE}$ .x.

Where x appears as a shielding parameter. The determination of this parameter is of prime importance. For that let us consider the limiting conditions defined with respect to the ratio of the driver transition time to the line output transition time. When this ratio is greater than unity the input driver see the complete line capacitance. Whereas if this ratio is smaller than unity ,the input driver see part of the line capacitance. Following these considerations the output transition time of the input driver and of the line can be obtained from

$$T_{OUTDriver1}(lineload) = T_{OUTFASTDRIVER1} \cdot \left[1 - \exp\left(-\frac{T_{OUTFASTDRIVER1}}{R_{LINE} \cdot C_{LINE}}\right)\right] \quad (6)$$

and

$$T_{OUTLINE} = T_{OUTFASTDRIVER1} + (R_{LINE} \cdot C_{LINE}) \cdot \left[\exp - \frac{T_{OUTFASTDRIVER1}}{R_{LINE} \cdot C_{LINE}}\right] \quad (7)$$

Validation of this shielding model has been obtained by comparing the transition time value simulated with the ELDO's transmission line model and calculated from eq.6 and 7. As specified in Table I we use three different sizes of input driver (DRIVER1 A,B,C), a minimum size output driver and two line parameters: R_{LINE}=232 Ω/mm, C_{LINE}=352 fF/mm and R_{LINE}=115 Ω/mm, C_{LINE}=472fF/mm. The output driver is kept at minimum size with a 5fF load. Fig.4 and 5 illustrates the validation obtained using (6,7) to model the output transition time variation at the input line driver and at the output of the line, including the resistance shielding effect on the line. As shown the agreement between calculated and simulated values is very good, we observe less than 5% of discrepancy in modeling the line output transition time. The divergence obtained in modeling the output transition time of the input driver A is the consequence of its heavy output loading (C_L/C_{IN} =216) for which the driver can not be considered as a current generator on the full range of line length. However this condition is far away the normal design conditions.

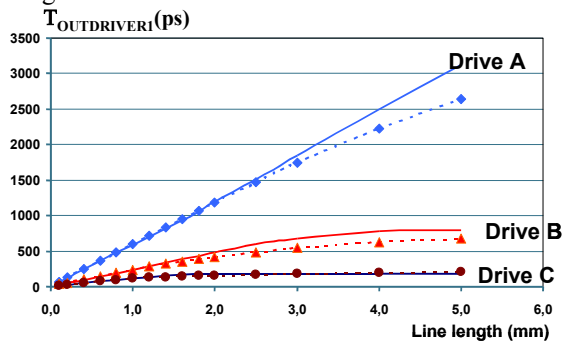

Fig.4. Simulated (full line) and calculated (dashed line) values of the input driver transition time for the configuration specified in Table I, with different size of input drivers (R_{LINE}=115 Ω/mm, C_{LINE}=472fF/mm).

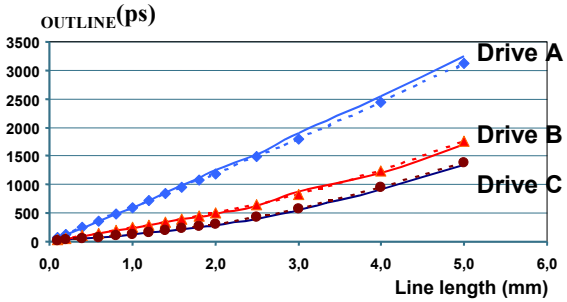

Fig.5. Simulated (full line) and calculated (dashed line) values of the line output transition time for the configuration specified in Table I, with different size of input drivers (R_{LINE}=115 Ω/mm, C_{LINE}=472fF/mm).

Note that eq.6,7, although much simpler than that proposed in [4], allows to estimate the line output transition time with a satisfactory accuracy. As shown in Table I, the influence of the shielding parameter is line length and input driver dependent. It induces a 30% load reduction when $T_{OUTFASTDRIVER1}/R_{LINE}.C_{LINE}$ = 1 and increases quickly when this ratio value decreases. As a result this ratio of transition time value can be used as a robust metric for considering shielding effects.

V. EXPERIMENTAL RESULTS

Using (6,7) the total delay on the line and the total delay between the input of DRIVER1 and the output of DRIVER2 can easily be obtained as

$$T_{TOTLINE} = T_{DRIVER1} + T_{Propag.line} \quad (8)$$

$$T_{OUT} - T_{IN} = T_{TOTLINE} + T_{DRIVER2} \quad (9)$$

where $T_{DRIVER2}$ is obtained from (4), considering the input transition time defined by the line output

transition time. The DRIVER2 transition time is calculated from (2) and (3) and includes, as well, the input transition time effect induced by the line. When not correctly captured this effect introduces the maximum error in timing evaluation.

Comparison between simulated and calculated values of the total line delay ($T_{TOTLINE}$), the DRIVER2 output transition time ($T_{OUTDRIVER2}$) and the overall delay across the complete configuration (T_{OUT}-T_{IN}) is given in Fig. 6-8. The agreement observed is very good (less than 5% discrepancy) over all the considered design range.

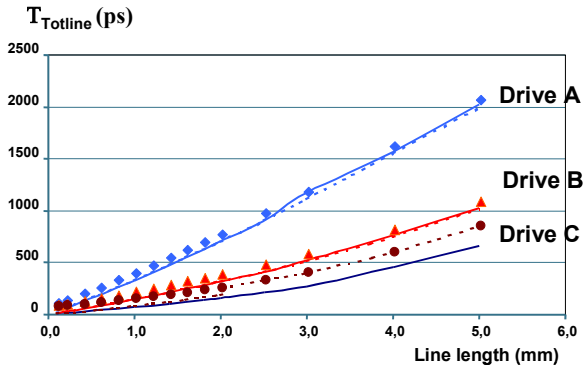

Fig.6. Simulated (full line) and calculated (dashed line) values of the total line propagation delay for the configuration specified in Table I, with different size of input drivers (R_{LINE}=115 Ω/mm, C_{LINE}=472fF/mm).

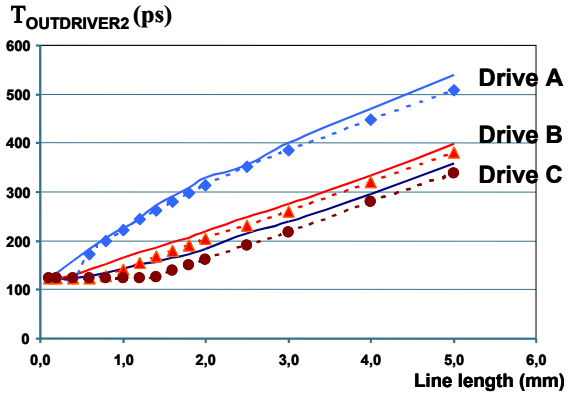

Fig.7. Simulated (full line) and calculated (dashed line) values of the DRIVER2 output transition time, for the same configuration.

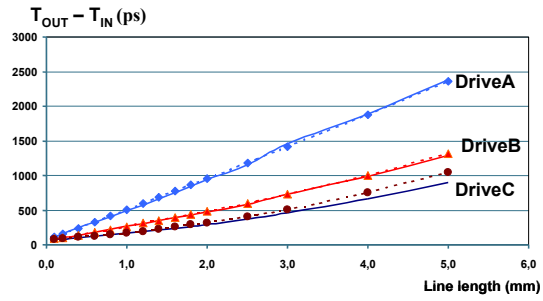

Fig.8. Simulated (full line) and calculated (dashed line) values of the total delay between the input of DRIVER1 and the output of DRIVER2, for the same configuration.

For illustration, in Table 2 we directly compare the calculated and simulated values of the total delay and the output driver transition time for the second set of line parameters we have considered: R_{LINE}=232 Ω/mm, C_{LINE}=352 fF/mm.

As shown the agreement between calculated and simulated values is excellent, the maximum discrepancy is smaller than 5% on the full-considered range. Note here the effect of the line output transition time on the output driver. Smaller is the transmitter (DRIVER1) greater is the value of the line output transition time (columns 3,4) and that of the output driver (columns 5,6). This is directly reflected in the value of the total delay (columns 7,8). This justifies the need for an accurate modeling of the line output transition time.

VI. CONCLUSION

We have proposed a complete modeling of the resistance-capacitance (RC) effect on the delay of an interconnect link between two inverters. Considering a previously developed deep sub-micrometer model, in which the inverter (gate) is considered as a current generator, we have obtained an analytical expression allowing to estimate the loading effect of the line on the transition time and the delay of the line input and output drivers. Load shielding of the input inverter has been captured with respect to the input drive-line transition time ratio, which can be used as an efficient metric for considering shielding effects.

Line output transition time effect on the transition time and delay of the output driver has been clearly defined. Validations with respect to transmission line simulations have demonstrated the potential application of this model in estimating the RC interconnect impact on circuit performance. Applications to be considered in using this analytical model are in driver selection and in line repeater insertion.

References

[1] L. T. Pillage and R. A. Rohrer, "Asymptotic Waveform Evaluation for Timing Analysis", IEEE Trans. Computer-Aided Design, vol. 9, no.4, 1990.

[2] F. Dartu, N. Menezes and L. T. Pileggi, "Performance Computation for Precharacterized CMOS Gates with RC loads", IEEE Trans. Computer-Aided Design, vol. 15, no. 5, may 1996.

[3] W. C. Elmore, "The transient response of damped linear network with particular regard to wide band amplifier", J. Appl. Physics,vol.19, pp.55-63,1948.

[4] J. Quian, S. Pullela, L. Pillage, "Modeling the effective capacitance for the RC interconnect of CLOS gates", IEEE trans. on Computer Aided Design, Vol. 13, n°12, pp.1526-1535, 1994

[5] M.Celik, L. Pileggi, A. Odabasioglu, "IC Interconnect analysis" Kluwer Acad. Press, 2002.

[6] K. Agarwal, D. Sylvester, D. Blauw, "Simple metrics for slew rate of RC circuits based on two circuit moments", Proc. Design Automation Conference, pp.950-953, 2003.

[7] ANACAD EES: Eldo User's Manual, Document number 8301, issue 4.2 june 93.

[8] K. O. Jeppson, "Modeling the influence of the transistor gain ratio and the input-to-output coupling capacitance on the CMOS inverter delay", IEEE J. Solid State Circuits, vol.29, pp.646-654, 1994.
[9] xxxx for anonymous review IEEE trans. on Computer Aided Design, 2002.
[10] Y. I. Ismail, E. G. Friedman, "On-chip inductance in high speed integrated circuits" Kluwer Acad. Publishers, 2001.
[11] J. M. Rabaey, "Digital Integrated Circuits, A Design Perspective" Prentice Hall, Inc., New Jersey, 1999.

Table I

	Line lgth mm	$T_{OUTDRIVER1}$ ps	$T_{OUTLINE}$ ps	$T_{TOTLINE}$ ps	$T_{OUTDRIVER2}$ ps	T_{OUT}-T_{IN} ps
DRIVER1 A	0.1	63	63	33	125	102
C_{IN} = 11fF	1	591	592	327	226	505
	3	1850	1900	1175	400	1462
	5	3126	3243	2024	508	2383
DRIVER1 B	0.1	33	33	16	123	82
C_{IN} = 28fF	1	239	241	148	141	265
	3	670	879	518	261	727
	5	795	1700	1025	398	1295
DRIVER1 C	0.1	16	16	12	126	75
C_{IN} = 56fF	1	115	125	71	143	167
	3	185	562	273	240	455
	5	185	1352	660	358	904

Table II

	Line lgth mm	$T_{OUTLINE}$ Sim.(ps)	$T_{OUTLINE}$ Calc. (ps)	$T_{OUTDRIVER2}$ Sim.(ps)	$T_{OUTDRIVER2}$ Calc. (ps)	T_{OUT}-T_{IN} Sim.(ps)	T_{OUT}-T_{IN} Calc.(ps)
DRIVER1 A	0.1	49	52	124	124	99	102
C_{IN} = 11fF	1	452	450	204	193	375	377
	3	1624	1466	380	348	1057	1008
	5	3015	2984	525	496	1911	1846
DRIVER1 B	0.1	23	27	123	123	79	86
C_{IN} = 28fF	1	189	192	158	126	214	189
	3	870	867	300	267	622	650
	5	2197	2140	449	420	1343	1388
DRIVER1 C	0.1	13	16	123	123	72	79
C_{IN} = 56fF	1	113	120	144	123	161	147
	3	697	690	276	239	517	537
	5	2015	1882	429	394	1209	1245