POWER CHARACTERISTICS OF INDUCTIVE INTERCONNECT

Magdy A. El-Moursy and Eby G. Friedman

Department of Electrical and Computer Engineering
University of Rochester
Rochester, New York 14627-0231

ABSTRACT

The power characteristics of an inductive interconnect lines is presented. The matching condition between the driver and the load has an effect on the power consumption since the short-circuit power dissipation may decrease with increasing line width and the dynamic power increases. A tradeoff exists between the short-circuit and dynamic power in inductive interconnects. The short-circuit power increases with wider line widths if the line is underdriven. As the short-circuit power is directly proportional to the signal transition time, an analytic solution for the transition time with an error of less than 15% is presented. The solution can be used in wire sizing synthesis techniques to decrease the overall power dissipation.

1. INTRODUCTION

With the decrease in the feature size of CMOS circuits, on-chip interconnect now dominates circuit delay and power dissipation. Furthermore, the inductive behavior of the interconnect can not be neglected, particularly in long interconnect lines. The interconnect inductance increases on-chip noise as well as the computational complexity of the design process. Furthermore, on-chip inductance may affect certain design techniques such as repeater insertion [1] and wiresizing [2].

Interconnect design has, therefore, become an important issue in high complexity, high speed integrated circuits (IC). Many algorithms have been proposed to determine the optimum wire size that minimizes a cost function such as delay [3] or power. As the inductance becomes important, specific algorithms have been enhanced that consider RLC impedances [4, 5].

Previous research in wiresizing does not consider the inductive impedance characteristics, which may lead to non-optimal circuits. For example, the research described in [2] neglects the impedance matching characteristics between the line driver and the load. In [2] wider wires are used to reduce power dissipation. It is shown in this paper that this technique may actually increase the total power dissipation.

This research was supported in part by the Semiconductor Research Corporation under Contract No. 2003-TJ-1068, the DARPA/ITO under AFRL Contract F29601-00-K-0182, the National Science Foundation under Contact No. CCR-0304574, the Fulbright Program under Grant No. 87481764, grants from the New York State Office of Science, Technology & Academic Research to the Center for Advanced Technology - Electronic Imaging Systems and to the Microelectronics Design Center, and by grants from Xerox Corporation, IBM Corporation, Intel Corporation, Lucent Technologies Corporation, and Eastman Kodak Company.

In this paper, the effect of sizing an inductive interconnect line on the short-circuit power dissipation of a driven CMOS gate is presented. The change in the power dissipation with the interconnect width is characterized in Fig. 1. It is shown that the impedance characteristics of an inductive interconnect line are sensitive to changes in the line width. As the short-circuit power depends directly on the signal transition time, an analytic solution for the signal transition time at the far end of an inductive interconnect line is presented. This solution can be used in design techniques such as wiresizing to optimize the overall transient power dissipation.

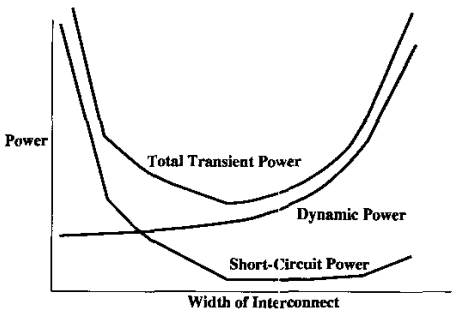

Figure 1. Short-circuit, dynamic, and total power dissipation as a function of the interconnect line width

The paper is organized as follows. In section 2, the power characteristics of an inductive interconnect line are discussed. Furthermore, the effects of changing the line width on the line impedance characteristics are presented. In section 3, an analytic solution for the transition time at the far end of an inductive interconnect is provided. A comparison between the analytical expression and simulation is presented in section 4. Some conclusions are discussed in last section.

2. POWER CHARACTERISTICS OF INDUCTIVE INTERCONNECT

The power characteristics of inductive interconnect line are discussed in this section. As shown in Fig. 1, the short-circuit power decreases as the interconnect width is increased until a minimum power level is reached. If the

interconnect width exceeds a certain size, the short-circuit power increases. The dynamic power increases with line width as the line capacitance increases. As shown in Fig. 1, a tradeoff exists between the dynamic and short-circuit power in sizing inductive interconnect.

For the circuit shown in Fig. 2, a long interconnect line between two CMOS gates is modeled as a lossy transmission line. The short-circuit power dissipation within the load gate P_{sc} is directly proportional to the input signal transition time, which is the transition time at the far end of an interconnect line. Regardless of the load characteristics, the short-circuit power P_{sc} can be represented as

$$P_{sc} = G(V_{dd}, V_t, K, C_L)\ \tau_0\ f_{clk}, \qquad (1)$$

where τ_0 is the transition time of the input signal at the load gate, f_{clk} is the transition frequency, and G is a function of the supply voltage V_{dd}, threshold voltage V_t and transconductance K of the load gate, and the capacitive load C_L. Different techniques have been developed to characterize G under different load models. The general form of (1) is valid whether the load is modeled as a capacitive load [6], a lossless transmission line [7], or a lossy transmission line [8]. G may also be a function of τ_0; however, the dependence of G on τ_0 is small.

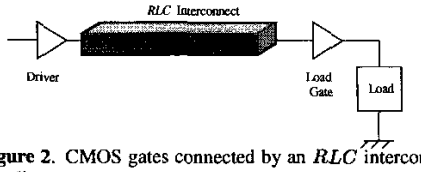

Figure 2. CMOS gates connected by an RLC interconnect line

At small interconnect widths, the characteristic line impedance Z_{lossy} is large as compared to the equivalent output resistance of the transistor R_{tr}. Thus, the line is overdriven (the underdamped condition). Z_{lossy} decreases with increasing line width. The line remains underdamped until Z_{lossy} equals R_{tr}. A further increase in the line width underdrives the line as Z_{lossy} becomes less than R_{tr}. As the line width is increased, the line driving condition changes from overdriven to matched to underdriven [5].

To characterize this behavior, an analytic solution for the signal transition time is presented in section 3. For an overdriven line, the short-circuit power dissipation changes with line width as shown in section 4. For an underdriven line, however, an increase in the line width increases the short-circuit power.

3. ANALYTIC SOLUTION FOR THE TRANSITION TIME

As described in section 2, short-circuit power is directly proportional to the signal transition time. Hence, an analytic expression for the signal transition time characterizes the short-circuit power of the load gate. To determine an analytic solution for the transition time of the signal at the far end of an inductive interconnect, a lumped RLC model of the interconnect impedance is assumed. The total line resistance, capacitance, and inductance are $R_t = Rl$, $C_t = Cl$, and $L_t = Ll$, respectively, where l is the line length. Adding the gate capacitance C_g to the line capacitance, the total load capacitance is $C_p = C_t + C_g$. The input ramp signal is

$$V_{in}(t) = \begin{cases} \frac{t}{\tau_r} V_{dd} & \text{for } 0 \le t \le \tau_r, \\ V_{dd} & \text{for } t > \tau_r, \end{cases} \qquad (2)$$

where τ_r is the transition time of the input signal. The line is assumed to be driven by a CMOS inverter.

When V_{in} changes from low-to-high, the PMOS transistor initially operates in the triode region, then enters the saturation region. When the input reaches $V_{dd} - |V_{tp}|$, the PMOS transistor turns off, and the charge on the capacitive load discharges through the NMOS transistor. The NMOS transistor initially operates in the saturation region, then moves into the triode region. For a comparable transition speed of the input and output signal of a CMOS inverter, the operation can be divided into four regions as listed in Table 1.

The PMOS and NMOS transistors are modeled by equivalent resistances $R_p = 1/\gamma_p$ and $R_n = 1/\gamma_n$, respectively. According to the n_{th}-power law MOSFET model, $\gamma_{p,n}$ is

$$\gamma_{p,n} = \alpha_{p,n} K_{p,n} (|V_{gsp,n} - V_{tp,n}|)^{m_{p,n}}, \qquad (3)$$

where $K_{p,n}$ and $m_{p,n}$ control the triode region characteristics of the transistor, $\alpha_{p,n}$ is a constant between one and two, representing the dependency of the MOSFET equivalent resistance on the drain-to-source voltage drop V_{ds}, and V_{gs} is the gate-to-source voltage of the transistor.

In region I, after V_{in} exceeds the NMOS transistor threshold voltage V_{tn}, the saturation current of the NMOS transistor is

$$I_{nI}(t) = B_n (V_{in}(t) - V_{tn})^{n_n} \text{ for } t \ge \tau_{nON}, \qquad (4)$$

where B_n and n_n determine the saturation region characteristics of the NMOS transistor and τ_{nON} is given by $V_{tn} \frac{\tau_r}{V_{dd}}$. At the output node of the driver, KCL and KVL equations are

$$I_p + I_l = I_n, \qquad (5)$$
$$V_o = V_c - V_r - V_l, \qquad (6)$$

respectively, where V_o is the voltage at the output node and V_r, V_l, and V_c are the voltage across the resistance, inductance, and capacitance, given by (7), (8), and (9), respectively. I_p, I_n, and I_l are the current through the PMOS transistor, the NMOS transistor, and the load capacitor, respectively,

$$V_r(t) = I_l(t)\, R_t, \qquad (7)$$

$$V_l(t) = L_t \frac{dI_l(t)}{dt}, \qquad (8)$$

$$V_c(t) = \frac{-1}{C_p} \int I_l(t) \qquad (9)$$

In region I, I_p, I_l, and V_c are given by (10), (11), and (12), respectively,

$$I_{pI}(t) = \gamma_p (V_{dd} - V_o), \qquad (10)$$

$$I_{lI}(t) = A + B\,t + D_1\, e^{-\alpha_{p1} t} + D_2\, e^{-\alpha_{p2} t}, \qquad (11)$$

Table 1. Different regions of operation for a CMOS inverter with a ramp input signal

Region	Condition		NMOS	PMOS
I	$V_{dd} - \|V_{tp}\| > V_{in} \geq V_{tn}$	$V_o > V_{in} + \|V_{tp}\|$	Saturation	Triode
II		$V_o > V_{in}$ - V_{tn}	Saturation	Saturation
III		$V_o \leq V_{in}$ - V_{tn}	Triode	Saturation
IV	$V_{in} \geq V_{dd} - \|V_{tp}\|$		Triode	Cut-off

$$V_{cI}(t) = V_{dd} - \frac{1}{C_p}[At + B\frac{t^2}{2} + D_1(1 - e^{-\alpha_{p1}t}) + D_2(1 - e^{-\alpha_{p2}t})], \quad (12)$$

where $A, B, D_1, D_2, \alpha_{p1}$, and α_{p2} are constants given by

$$A = -(D_1 + D_2),$$
$$B = q C_p L_t,$$
$$D_{1,2} = q\frac{1}{\alpha_{p1,2}^2(\alpha_{p2,1} - \alpha_{p1,2})},$$
$$\alpha_{p1,2} = \frac{\frac{1+R_t\gamma_p}{L_t\gamma_p} \pm \sqrt{(\frac{1+R_t\gamma_p}{L_t\gamma_p})^2 - \frac{4}{L_t C_p}}}{2},$$

where $\beta = \frac{\tau_r V_{tn}}{V_{dd}}$, $q = (n_n)!(\frac{V_{dd}}{\tau_r})^{n_n}$.

Region II starts when $V_o(t)$ reaches $V_{in} + |V_{tp}|$. In this region, the PMOS transistor is saturated. The output voltage in this region is determined using a Newton-Raphson iteration. $V_c(\tau_{psat})$ is determined by (12), where τ_{psat} is the initial time of this region.

Since both transistors have the same drain voltage, the second region of operation in which both transistors are saturated is quite short, allowing the change in V_c during that region to be neglected.

During region III, the NMOS transistor operates in the triode region, and the PMOS transistor is saturated. Expressions for $I_l(t)$ and $V_c(t)$ are obtained in the same way as discussed in region I, and are given by (13) and (14), respectively.

$$I_{lIII}(t) = A_1 + B_1 t + E_1 e^{-\alpha_{n1}t} + E_2 e^{-\alpha_{n2}t}, \quad (13)$$

$$V_{cIII}(t) = V_c(\tau_{psat}) - \frac{1}{C_p}\int_{\tau_{psat}}^{t} I_{lIII}(t)\,dt, \quad (14)$$

where A_1, B_1, E_1, and E_2 are constants given by

$$A_1 = I_{lIII}(\tau_{psat}) - (B_1\tau_{psat} + E_1 e^{-\alpha_{n1}\tau_{psat}} + E_2 e^{-\alpha_{n2}\tau_{psat}}),$$
$$B_1 = -2C_p\frac{b}{\gamma_n},$$
$$E_{1,2} = \frac{\alpha_{n1,2}^2(\gamma_n V_{oIII}(\tau_{psat}) - a^2 b) - 2b - 2ab\alpha_{n1,2}}{\alpha_{n1,2}^2(\alpha_{n2,1} - \alpha_{n1,2})\gamma_n L_t},$$

where $b = B_p(\frac{V_{dd}}{\tau_r})^{n_p}$, $a = \frac{(V_{dd} - V_{tp})\tau_r}{V_{dd}}$, B_p and n_p are parameters that determine the characteristics of the saturation region of a PMOS transistor, and α_n is similar to α_p for an NMOS transistor.

As V_{in} reaches $V_{dd} - |V_{tp}|$, the PMOS transistor turns off, initiating region IV. The time at which this region begins is $\tau_{pOFF} = \frac{(V_{dd} - |V_{tp}|)\tau_r}{V_{dd}}$, where $V_o(\tau_{pOFF})$ is obtained from (6). After the PMOS transistor turns off, the NMOS transistor continues to operate in the triode region. An expression for $V_c(t)$ in this region is

$$V_{cIV}(t) = V_c(\tau_{pOFF})e^{-\alpha_{n2}(t - \tau_{pOFF})}. \quad (15)$$

The transition time τ_0 is $\tau_0 = \frac{(t_{10\%} - t_{90\%})}{0.8}$ where $t_{10\%}$ and $t_{90\%}$ are the times at which the far end voltage reaches 10% and 90% of the output voltage, respectively. These expressions can be obtained from (15) by substituting $V_c(t_{10\%})$ with $0.1V_{dd}$ and $V_c(t_{90\%})$ with $0.9V_{dd}$.

4. SIMULATION RESULTS

In order to verify the analytical expressions derived in section 3, the line resistance is expressed by $R = \frac{\rho}{W_{INT}T}$, where ρ, T, and W_{INT} are the line resistivity, thickness, and width, respectively.

The values of R, C, and L are determined based on the following physical parameters, $\rho = 2.5\,\mu\Omega cm$ and $l = 5\,mm$. A 0.24 μm CMOS inverter with $W_n = 15\,\mu m$ and $W_p = 30\,\mu m$ is assumed. As listed in Table 2, the transition time is determined from the analytical equation described by (15) and compared with SPICE. The line width is varied from 0.1 μm to 20 μm. The maximum error in the analytical expression as compared to SPICE is less than 15% and is typically around 8%. The transition time based on this analytical expression is compared to SPICE in Fig. 3. It is shown in the figure that the transition time decreases with increasing interconnect width if the line is overdriven. After exceeding the line width at which the line is matched with the driver, the transition time increases as the line is underdriven. The short-circuit power has the same behavior as measured by SPICE and shown in Fig. 4.

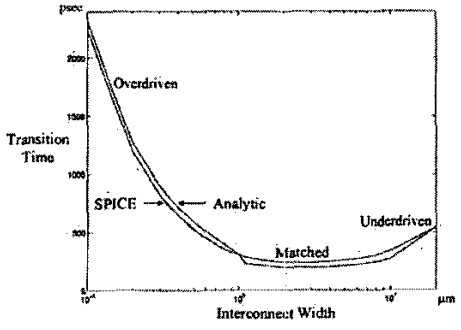

Figure 3. Analytical solution of the transition time as compared with SPICE for different line widths.

Table 2. Simulation and analytical transition times of the signal at the end of a long interconnect line

$W_{INT}(\mu m)$	R_t (Ω)	L_t (nH)	C_p (fF)	τ_0		Error (%)
				Spice	Analytic	
0.1	1250	9.62	628.1	2386.25	2510.71	5.22
0.2	625.0	9.53	652.02	1349.13	1431.8	6.13
0.3	416.67	9.45	670.56	999.88	1050.92	5.11
0.4	312.5	9.37	686.80	826.25	870.73	5.38
0.5	250.0	9.30	701.72	725.0	761.57	5.04
0.6	208.33	9.24	715.80	657.50	688.53	4.72
0.7	178.57	9.18	729.27	614.0	636.50	3.66
0.8	156.25	9.12	742.30	582.44	597.86	2.65
0.9	138.89	9.07	754.98	559.66	568.38	1.56
1	125.0	9.02	767.38	543.03	545.54	0.46
2	62.5	8.61	883.13	507.73	460.45	-9.31
3	41.67	8.32	991.97	537.98	465.83	-13.41
4	31.25	8.10	1097.83	581.59	497.43	-14.47
5	25.00	7.92	1202.00	629.75	538.98	-14.41
10	12.50	7.31	1711.30	886.04	781.18	-11.83
20	6.25	6.67	2709.85	1411.56	1273.64	-9.77

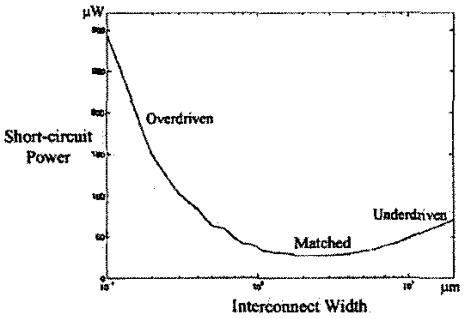

Figure 4. Short-circuit power of the driven gate.

5. CONCLUSIONS

It is shown in this paper that the short-circuit power within a load gate driven by an inductive interconnect line is dependent upon the impedance characteristics of the line. The matching condition between the driver and the load is shown to have an important effect on the line impedance characteristics. If the line is overdriven, the short-circuit power decreases with increasing line width. When the line exceeds the matched condition, the short-circuit power increases with increasing signal transition time.

An analytic solution of the signal transition time at the far end of an inductive interconnect line is presented. The solution can be used to estimate the transition time at the far end of an inductive interconnect and exhibits an error of less than 15%. The analytic solution can be used to both estimate and optimize the power dissipation within high speed CMOS circuits.

6. REFERENCES

[1] Y. I. Ismail, E. G. Friedman, and J. L. Neves, "Effects of Inductance on the Propagation Delay and Repeater Insertion in VLSI Circuits," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Vol. 8, No. 2, pp. 195-206, April 2000.

[2] Y. I. Ismail, E. G. Friedman, and J. L. Neves, "Exploiting On-Chip Inductance in High Speed Clock Distribution Networks," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Vol. 9, No. 6, pp. 963 - 973, December 2001.

[3] J. J. Cong, K. Leung, and D. Zhou, "Performance-Driven Interconnect Design Based on Distributed *RC* Delay Model," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 14, No. 3, pp. 321-336, March 1995.

[4] Q. Zhu and W. M. Dai, "High-Speed Clock Network Sizing Optimization Based on Distributed *RC* and Lossy *RLC* Interconnect Models," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 15, No. 9, pp. 1106-1118, September 1996.

[5] M. A. El-Moursy and E. G. Friedman, "Optimizing Inductive Interconnect for Low Power," *Canadian Journal of Electrical and Computer Engineering*, Vol. 27, No. 4, pp. 183-187, October 2002.

[6] L. Bisdounis, S. Nikolaidis, and O. Koufopavlou, "Propagation Delay and Short-Circuit Power Dissipation Modeling of the CMOS Inverter," *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, Vol. 45, No. 3, pp. 259-270, March 1998.

[7] Y. I. Ismail, E. G. Friedman, and J. L. Neves, "Dynamic and Short-Circuit Power of CMOS Gates Driving Lossless Transmission Lines," *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, Vol. 46, No. 8, pp. 950-961, August 1999.

[8] K. T. Tang and E. G. Friedman, "Delay and Power Expressions Characterizing a CMOS Inverter Driving an *RLC* Load," *Proceedings of the IEEE International Symposium on Circuits and Systems*, Vol. III, pp. 283-286, May 2000.