SIMULATION OF SEMICONDUCTOR DEVICES AND PROCESSES Vol. 5
Edited by S. Selberherr, H. Stippel, E. Strasser – September 1993

A Newly Proposed Delay Improvement on CMOS/SOI Future Technology

M. Lee‡† and K. Asada‡

‡Department of Electronic Engineering, The University of Tokyo
7-3-1 Hongo, Bunkyo-ku, Tokyo 113, JAPAN
†Nippon Motorola Ltd.
3-20-1 Minami-azabu, Minato-ku, Tokyo 106, JAPAN

Abstract

Contribution of gate fringing capacitance to CMOS/SIMOX inverter time delay in deep sub-micrometer gate is propounded. Measurements of the fifty-one stage ring oscillator's time delay are completed for comparison with analytical model. Propagation Delay Times(TPD) by reducing Poly-Si gate thickness were improved up to two times in deep-submicron CMOS/SIMOX inverters. It is concluded that SOI technology is promising for a *high speed* by reducing gate fringing capacitance which is correlated to the poly-Si gate thickness.

1. Introduction

Deep sub-micrometer gate SOI devices built in ultra-thin Si film have attracted much attention for their high performance. It is generally acknowledged that the gate oxide capacitance strongly affects on the TPD [1][2][3][7], however parasitic capacitances other than the gate oxide capacitance become a major factor limiting device speed with device miniaturization [4] [5]. In this paper, we studied on the contribution of gate fringing capacitance to the time delay. Reduction of poly-Si gate thickness proportionally decreases the gate fringing capacitance as shown in Fig. 1 and Fig. 2, resulting in high speed of fabricated SOI CMOS inverters. It is indicated that a method for reducing the poly-Si gate thickness is proposed to improve the TPD as in future SOI technology.

2. Formulation for Model

It has been reported[5] that gate capacitances of CMOS/SOI inverter are composed of gate oxide capacitance, C_{ox}, gate fringing capacitance, C_f and wiring capacitance, C_w, in nMOS and pMOS devices. A suggestion for high speed has also been raised by reducing parasitic capacitances other than intrinsic gate oxide capacitance using effective time-dependent capacitance model [6]. In this study we focused on the contribution of the C_f to the delay time, reducing poly-Si gate thicknes (t_m). It is found that gate fringing capacitance (C_f) is significantly decreased by reducing

the t_m with assumption of invariable resistance for poly-Si gate as $t_m = 350$ (nm) whose assumption can be done by salicidation technique. The formulation of C_f is separated into three components at various distances from the edge which involves transforming a rectangular electrode and ground plane into the complex potential plane [6]. Introducing the parallel plate model for C_{ox} and summarizing the gate fringing capacitance, C_f, in the CMOS/SOI inverter as follows:

$$C_f = \epsilon_{\mathrm{ox}} \times \gamma \times (W_g + L_g), \tag{1}$$

where L_g is gate length, W_g is gate width, ϵ_{ox} is permittivity of oxide, ϵ_{ni} is permittivity of nitride, and

$$\gamma = \frac{\left(0.613 + \ln(\frac{u}{a}) + (\frac{\epsilon_{\mathrm{ni}}}{\epsilon_{\mathrm{ox}}}) \ln(a)\right)}{\pi}, \tag{2}$$

while constants for u and a are determined by the following equations: $a = -1 + 2K^2 + 2K(K^2-1)^{1/2}$ and $u = (R^2 a - 1)/(R^2 - 1)$, where K and R are calculated as follows:

$$K = 1 + t_m/t_{ox}, \tag{3}$$

and nonlinear expression on R as

$$\frac{\pi}{2} \cdot \frac{L_G}{t_{ox}} = \frac{a-1}{a^{1/2}} \cdot \frac{R}{(R^2-1)} + \ln\left(\frac{a^{1/2}R+1}{a^{1/2}R-1}\right) - \frac{a+1}{2a^{1/2}} \cdot \ln\left(\frac{R+1}{R-1}\right). \tag{4}$$

We analyzed the influence of the C_f via varying the poly-Si gate thickness, t_m, appeared in eq. (3). The constant, R, should first be solved numerically to evaluate the fringing parameter, γ, in Fig. 1(a) and corresponding C_f in Fig. 2(a) varying the poly-Si gate thickness, t_m. Fig. 2(b) shows the calculated results for the conventional gate oxide capacitance, C_{ox}, using the parallel-plate model, $C_g(t)$ by new model, gate fringing capacitance, C_f, in eq. (1) by the above sequence and drain-substrate capacitance, C_d, using the parallel-plate model for the buried oxide. It is noted that the gate capacitance was formulated using a Time-Dependent Gate Capacitance(TDGC) model, $C_{out}(t)$ [6] as follows:

$$C_{out}(t) = C_g(t) + C_f^{(n)} + C_f^{(p)} + C_w, \tag{5}$$

where $C_f^{(n)}$ and $C_f^{(p)}$ are gate fringing capacitances of nMOS and pMOS, respectively, C_w comprises both of metal wiring stray capacitance C_{metal} and drain-substrate capacitance C_d. However, the C_{metal} is negligible for SOI circuits. $C_g(t)$ at average of nMOS pull-down and pMOS pull-up circuit is shown in Fig. 2(b). Details for the $C_g(t)$, TDGC model, are expressed [6]. The output level $V_{out}(t)$ of an inverter during low-to-high transient process of the inverter gate potential can be represented as

$$\frac{d}{dt}\{C_{out}(t) \times V_{out}(t)\} = -I_{nMOS}(V_{out}(t)), \tag{6}$$

where I_{nMOS} is the nMOS drain current and $C_{out}(t)$ is output capacitance in eq. (5). The eq.(6) was solved numerically with dynamic change of the I_{nMOS}. The program sequence for the Propagation delay time(TPD) is depicted in Fig. 1(b). The TPD is defined as an average of a time for 50 % V_{dd} of nMOS pull-down circuit and pMOS pull-up circuit [5][6]. Details for numerical technique by multistep methods using Milne's *Predictor-Corrector*, MPC, algorithm are explained in [5]. It was found that the MPC appears to be somewhat more favorable than other techniques such as Runge-Kutta 5th order method [8] and Adams-Moulton method [8] such as better accuracy by the iterations. economing the number of corrections, automatic error estimate at each step which will minimize the computing time [9][10][11].

3. Results and Discussion

To provide experimental support for the analysis in section 2, we measured DC characteristics and TPD at various operating frequencies and power supply voltages at three-terminal SOI MOSFET's and CMOS inverters, respectively. The devices were fabricated at Nippon Telegraph and Telephone(NTT) with poly-Si gate thickness(t_m), 350(nm), same as [6]. Typical case TPDs are clearly shown as solid circles in Fig. 3 at $t_m = 350$(nm) for comparison with experimental results. Fig. 3 also shows other simulated results at different gate thicknesses and supply voltages of 1.5V and 2.0V. It is found that approximately two-time decreased speeds to the typical case TPDs were observed as the gate thickness was approached to zero. It is implied the reduction of C_f was obtained by decreasing the t_m and was significant to improve the propagation delay times for next generation in integrated circuit, *i. e., high speed* as in SOI future technology.

4. Conclusion

Contributions of gate fringing capacitance on the propagation delay times are studied. It is concluded that reduction of the poly-Si gate thickness(t_m) decreases the gate fringing capacitance(C_f), resulting in improving the propagation delays up to about two times than typically simulated TPDs. This is promising for high speed CMOS/SOI devices by reducing the t_m with current SOI technology. Thinning the t_m for high performance is proposed in CMOS/SIMOX technology. It is also predicted that this proposed reduction of the gate height leads to low power as next generation in integrated circuit.

References

[1] T. Douseki K. Aoyama and Y. Omura, "Dependence of CMOS/SIMOX inverter delay time on gate overlap capacitance," *IEICE Trans.*, ED92-62(176):39–44, Aug. 1992

[2] T. Sakurai and A. R. Newton, "Alpha-Power Law MOSFET Model and its Applications to CMOS Inverter Delay and Other Formulas," *IEEE J. Solid-State Circuits*, SC-25(2):584–594, Apr. 1990

[3] T. Sakurai and A. R. Newton, "Delay Analysis of Series-Connected MOSFET Circuits," *IEEE J. Solid-State Circuits*, SC-26(2):122–131, Feb. 1991

[4] M. Fujishima, M. Ikeda, K. Asada, Y. Omura and K. Izumi, "Analytical Modeling of Dynamic Performance of Deep Sub-micron SOI/SIMOX based on Current-Delay Product," *IEICE Trans. Electron.*, E75-C(12):1506–1514, Dec. 1992

[5] M. Lee and K. Asada, "Deep Sub-micron CMOS/SIMOX delay modeling by Time-Dependent Capacitance Model," *IEEE Trans. on Electron Devices,* To be published, July, 1993

[6] M. Lee and K. Asada, "Sub-100nm CMOS/SIMOX Delay Modeling by Time-Dependent Gate Capacitance Model," *1993 International Symposium on VLSI TSA*, Presented, Taipei, Taiwan, May 1993

[7] S. R. Vemuru and A. R. Thorbjornsen, "A MODEL FOR DELAY EVALUATION OF A CMOS INVERTER," *1990 IEEE Int. Sym. on CAS*, ISCAS-90(1):89–92, May. 1990

[8] S. D. Conte and C. Boor, *Elementary Numerical Analysis: An Algorithmic Approach*, 3rd ed., London; McGRAW-Hill:120–127, 1980

[9] R. E. Bank and D. J. Rose, "Global approximate Newton Methods," *Numer. Math.*, (37):279–295, 1981

[10] L. A. Hageman and D. M. Young, "Applied Iterative Methods," New York; Academic Press, 1981

[11] D. Luenberger, *Linear and Nonlinear Programming*, 2nd ed., Philippines; Addison-Wesley, 1984.

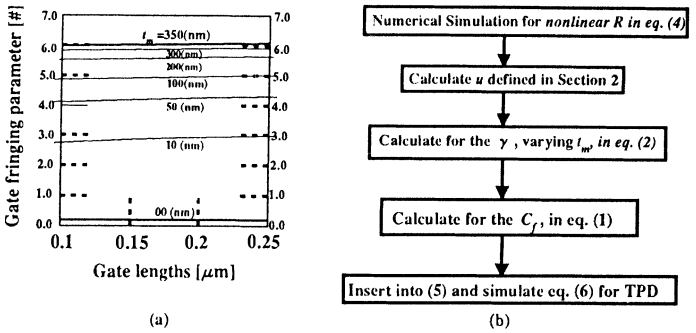

Fig. 1 (a) Simulation results for gate fringing parameter, γ, reducing poly-Si gate thickness (t_m). (b) Program sequence to obtain the gate fringing capacitance (C_f) for the delay time.

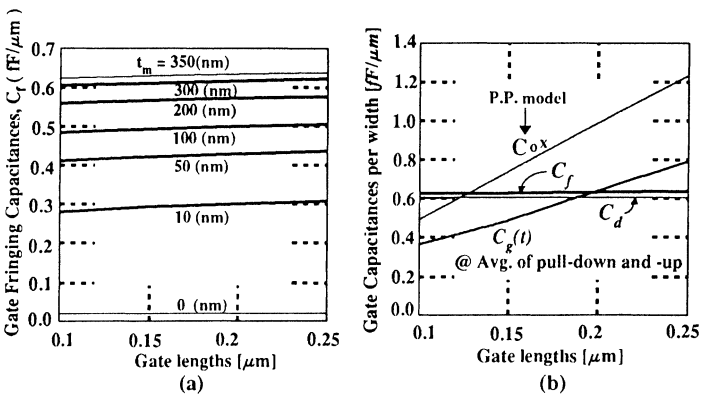

Fig. 2 (a) Normalized C_f by gate width, depending upon poly-Si gate thickness (t_m). (b) Normalized gate capacitances by the gate width based on measured parameters.

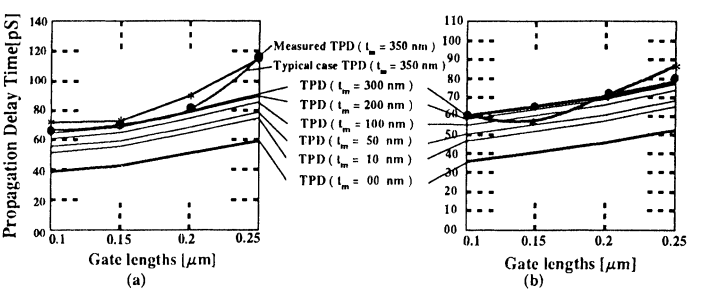

Fig. 3. The Simulated TPDs with the Measured data at (a) V_{dd} = 1.5v; (b) V_{dd} = 2.0v.

* : Measured TPDs
Solid line with filled circles : Typical case TPDs by TDGC with C_f @t_m=350nm.
Other Solid lines : TPDs with C_f @ different t_m = 300nm to 00 nm.

CMOS inverters: W_g of pMOSFET is twice larger than W_g of nMOSFET. Back-gate bias of 0.0V and