ESTIMATION OF FRINGING CAPACITANCE USING RC – DELAY SIMULATIONS

Asha G H[1], Dr. Deepali Koppad[2]
[1]Associate Professor, Dept. of Electronics & Communication Engineering, Malnad College of Engineering, Hassan India
[2]Professor, Dept. of Electronics and Communication, PES University, Bangalore, India

Abstract
As transistors become smaller, they switch faster, dissipate less power, and are cheaper to manufacture! Since 1995, as the technical challenges have become greater, the pace of innovation has actually accelerated because of ferocious competition across the industry. Designers need to be able to predict the effect of this feature size scaling on chip performance to plan future products, ensure existing products will scale gracefully to future processes for cost reduction, and anticipate looming design challenges. Scaled transistors are steadily improving in delay, but scaled global wires are getting worse. Increased gate leakage is one of the main limiting factors for aggressive scaling of $Si0_2$ for deep sub-micron CMOS technology. Extracted gate to source/drain fringing capacitances using a highly accurate 3D capacitance extractor [1] has shown that the inner fringing capacitances between the gate and the source/drain play a key role in degrading the short channel performance in the case of MOSFETs with high-K gate dielectrics. Numerical simulations [2] have shown that for sub-micron gate lengths normalized total fringing capacitance associated with the transistor is greater than the intrinsic counterpart. An estimation of fringing capacitance can be carried out by taking a design and finding the power dissipation for two different technologies. If the resistance component is kept constant the difference in the power is proportional to fringing capacitance.

Keywords:Fringing capacitance, FO4-Inverter, RC-delay, technology-node, scaling

I. INTRODUCTION

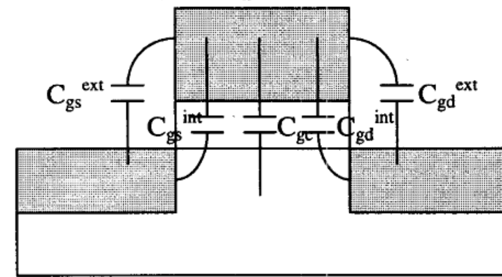

Fig. 1. Capacitances associated with a MOS - Device

For a fully scaled 70 nm gate length **MOSFETs** following the technology Roadmap having a gate length down to 70 nm, the source/drain junction depths fixed at 35 nm, peak channel doping concentration of 1.5 X 10^{18} cm^{-3} and a threshold voltage of 0.2V, all the capacitances associated with the device, which are extracted from direct Monte- Carlo simulations are shown in Fig. 1. The fringing capacitance C_f is the parallel combination of fringing capacitance on the outer side C_{of} and the fringing capacitance on the inner side C_{if} between the gate and source/drain junctions, defined as in equation (1).

$$C_f = C_{of} + C_{if}$$
$$C_{of} = C_{gs}^{ext} + C_{gd}^{ext} \quad (1)$$
$$C_{if} = C_{gs}^{int} + C_{gd}^{int}$$

The variation of C_{of} and C_{if} with K value is such that, C_{if} decrease and C_{of} increases, with an overall decrease in total capacitance with increasing K. The external fringing capacitance decreases due to the increasing distance (due to the higher physical oxide thickness) travelled by field lines from gate to source/drain for increasing K. This decrease in C_{of} results in reduced delay for circuit operation. The higher

C_{if} in high-K dielectrics is due to a larger number of field lines terminating on the edges of the source/drain depletion region. This induces an electric field from source/drain to channel reducing the barrier height. Since the threshold voltage V_t of the device is controlled by the injection of electrons over this potential barrier, V_t decreases with increasing K. The gate to channel capacitance C_{gc} reduces with higher K decreasing the control of gate over channel which increases the *2D* effects. This introduces Drain Induced Barrier Lowering (DIBL), increasing the sub-threshold slope due to lower C_{gs} and thus increasing OFF state leakage. In case of higher K more of the field lines originating from the drain terminate on the source rather than on the gate reducing the peak barrier height and increasing 2-D effects. From the circuit point of view, the lower capacitances would reduce the delay, a desirable effect, while the increased inner fringing capacitances would degrade the short-channel performance thus leading to higher leakage currents and power dissipation.

A. ON-Resistance of a Unit-nMOSFET[3]

Nano-meter processes have gate capacitance of roughly 1fF/µm. If the Fan-Out-of-4 (FO4) inverter delay [3] of a process with a feature size – f in nm is 0.5 pico seconds, then the ON – resistance of a unit (4λ wide) nMOS transistor can be obtained from FO4 inverter delay as = 5τ = 15RC

$$RC = \frac{0.5f}{15} = \frac{f\ psec}{30nm}$$

A unit transistor has a width W = 2f and thus capacitance of

$$C = 2f\ femtoF/\mu m$$

Solving for resistance R

$$R = \frac{f\ psec}{30nm}\frac{1\mu m}{2f\ femtoF} = 16.6K\Omega$$

B. Extracting effective capacitance for delay estimation[3]

Fig. 2 shows a circuit for determining the effective gate capacitance of inverter X4. The approach is to adjust the capacitance C_{delay} until the average delay from c to g equals the delay from c to d. Because X6 and X3 have the same input slope and are the same size, when they have the same delay, C_{delay} must equal the effective gate capacitance of X4. X1 and X2 are used to produce a reasonable input slope on node c. A single inverter could suffice, but the inverter pair is even better because it provides a slope on node-c that is essentially independent of the rise time at node-a. X5 is the load on X4 to prevent node-e from switching excessively fast, which would over predict the significance of the gate-to-drain capacitance in X4.

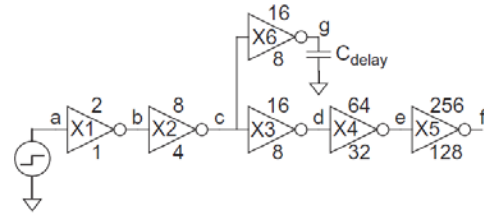

Fig. 2. Circuit for extracting effective gate Capacitance for delay estimation

C. FO4-Inverter Delay

One of the simplest measures of a process's inherent speed is the fanout-of-4 inverter delay. A circuit to measure this delay with nMOS and pMOS transistor sizes (in multiples of unit 4/2λ transistor) are listed below and above each gate, respectively. X3 is the inverter under test and X4 is its load, which is four times larger than X3. To first order, these two inverters would be sufficient. To obtain a realistic input slope it would be necessary to drive node c with a pair of FO4 inverters X1 and X2. The input capacitance of X4 depends not just on its Cgs but also on Cgd. Cgd is Miller-multiplied as node e switches and would be effectively doubled if e switched instantaneously. When e is loaded with X5, it switches at a slower, more realistic rate, slightly reducing the effective capacitance presented at node d by X4.

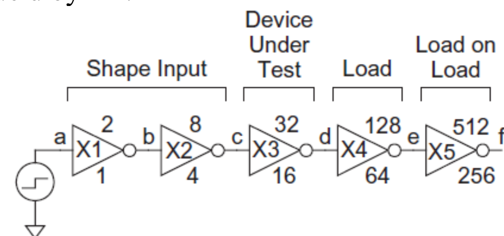

Fig. 3. Circuit for finding FO4-Inverter Delay

If a unit transistor has R = 10 kΩ and C = 0.1 fF in a 65 nm process, the RC product in the 65 nm process is (10 kΩ)(0.1 fF) = 1 ps. For h = 4, the delay is (3 + 3h)(1 ps) = 15 ps. This is called the fanout-of-4 (FO4) inverter delay and is representative of gate delays in a typical circuit.

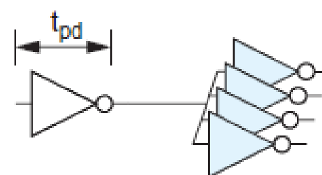

Fig. 4. Fanout-of-4 (FO4) inverter[3]

It is often helpful to express delay in a process-independent form so that circuits can be

compared based on topology. A process-independent measure for delay was observed when we carried out the simulation of FO4 – Inverter delay, and we could get a reasonable difference in power dissipation. Simulations were carried out on virtuso platform of cadence VLSI design tool of FO4 – Inverter for 180 nm, 90 nm and 45 nm technology nodes. From the results of our simulations we could substantiate the fact that there is an increase in power dissipation with scaling and that is due to increase in the fringing capacitance as the technology is scaled down from 180 nm to 45 nm.

D. Simulations

The schematic of FO4 Inverter was simulated for three different technology nodes of 180 nm, 90 nm and 45 nm to extract the delay and power.

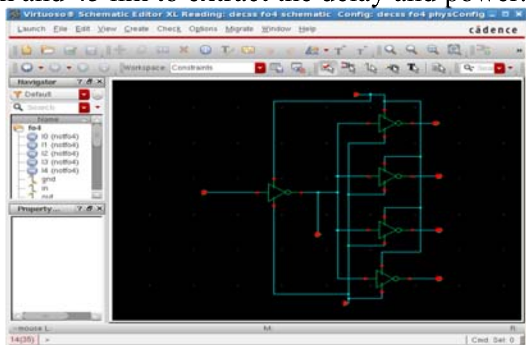

Fig. 5. Schematic of FO4 – Inverter

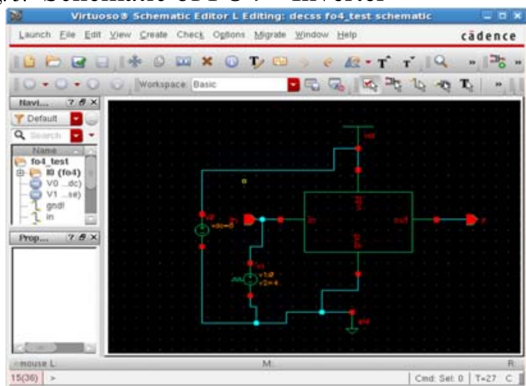

Fig. 6. Test - Bench for FO4 - Inverter

Fig. 7. Waveforms of FO4 – Inverter Response

TABLE I.

Technology Node	Dynamic Power	Static Power
180 nm	302.5 μW	467.7 μW
90 nm	3.462 mW	6.514 mW
45 nm	1.675 mW	3.451 mW

Fig. 8. Power dissipated in FO4 – Inverter simulation in three different Technology Nodes, with same delay of 10.06 nsec in all three Technology Node

E. Conclusion

From the results of simulations it is clear that as the technology node is moved from 180 nm to 90 nm the power dissipated has increased from μW to mW giving a clear indication that the fringing capacitance component has contributed for the increase in power. All three technology nodes used for simulations were with the same power supply magnitude. Only possible cause in the increase in power is capacitance. In that capacitance the component of device capacitance and the resistance which gives RC – component remains constant because of which the delay remains the same as such the delay for all the three simulations is 10.06 nsec. This confirms the fact that the increase in power is because of the fringing capacitance. This is a simple method of extracting the fringing capacitance. An estimation of fringing capacitance has been carried out by taking an FO4 - Inverter and finding the power dissipation for two different technologies the same scheme can be used for any circuit.

REFERENCES:

[1] Effect of Fringing Capacitance in Sub 100nm MOSFET with high – k dielectrics – Nihar. K Mohapatra, A Datta, M P Desai and V Ramgopal Rao, Department of Eleectrical Engg., IIT Bombay, 2000 IEEE

[2] Extraction procedure for MOS Structure fringing Gate Capacitance components – ROPEC 2015 – Advanced Material and Devices ; J. C. Tinoco, A. G. Martinez-Lopez, G. Lezama, M. Estrada, A. Cerdeira;2015 IEEE

[3] CMOS VLSI Design – A Circuits and Systems Perspective; NEIL H E WESTE, DAVID HARRIES, AYAN BANERJEE; 4th Edition; PEARSON Education