Optimising a D Flip Flop Through Delay and Power Estimation Using an RC Model and Transistor Sizing

Manan Joshi, D.S. Chauhan and B.K. Kaushik

Uttarakhand Technical University, Dehradun, India
GLA University, Mathura, India
Indian Institute of Technology, Roorkee, India

Abstract: This paper deals with the designing of explicit pulsed dual edge triggered D Flip Flop. The flip flop is designed with the help of CMOS inverter and CMOS transmission gates using 180nm technology. Delay and power consumption are estimated using theoretical means and then compared with simulation results. It is found that the theoretical results of delay are within 1% of the simulation results. Theoretical calculations show that there is an alternative method of circuit optimisation when simulation based hit and trial method becomes too complicated in large digital circuits.

Key words: Dynamic power consumption • Lumped RC network • Delay model, power estimation model • Critical path • Circuit optimisation • Model file

INTRODUCTION

Flip Flops and latches consume a large portion of system power due to redundant transitions in the internal nodes [1]. Power budget of portable digital circuit is severely limited and thus requires lowering of power consumption. Due to tight timing budgets at high frequency operations latency of the flip flops has to be reduced [2]. Transistor sizing thus plays an important role in the optimisation of power and delay of a digital circuit.

Methods of calculating delay time easily are categorized into two groups. One is the look-up table method and the other is the delay time calculation method. A look-up table method has a trade-off between the cost of establishing the table and the accuracy [3-4]. Hence there is very little flexibility for various factors affecting delay time, such as input transition time, fan-out, output capacitance, sizes of MOSFETs, etc. The macro-modeling method [3] is well suited for delay calculations of regular structure VLSI circuits, as well as circuits designed from standard cell libraries. Timing models for both logic gate and transmission gate circuit forms are developed. A more accurate and efficient Gate Level Delay calculation [4] uses a delay model that includes an accurate representation of the waveform, a consistent and meaningful definition of delay, a consideration of waveform slope effects at both the input and output of a gate and an innovative approach for handling transmission gate circuits. The highly efficient delay characterization is accomplished through a fast timing simulation technique, a theorem that reduces a two-dimensional delay table into a scaled one-dimensional table and an incremental characterization process.

A systematic method [5] to reduce standard cell library characterization time significantly is to use a simple and physically reasonable logic gate delay model in which delay varies linearly with C and t. Determine its region of validity in the (C, t) space and express the delay model coefficients and its region of validity as a function of inverter (or logic gate) size. As device current/capacitance models are not used, the method is general enough to be used with scaling.

The methods that rely on calculation construct a special delay model and are divided into two groups \( i.e. \) the group that includes the properties of the circuits or MOSFETs [6] in the delay model and the other group that does not include [7] the properties of the circuits or MOSFETs in the delay model. The latter methods consider many factors affecting the delay time. The factors are not enough for more general cases where the sizes of MOSFETs, output capacitances, etc., are varied. The remodelling method [6], uses an RC model for a
MOSFET but this method also can not model exactly the non-linear property of MOSFET with linear resistor. Delay models for inverter and other logic gates are modelled with \( \alpha \)-power MOSFET model [6], but they considered only the case of very small input transition time, resulting in less accuracy as the input transition time increases. A \( \alpha \)-power-law MOS model [8] that includes the carrier velocity saturation effect, which becomes prominent in short-channel MOSFETs, is introduced. Since the model is simple, it is used to handle MOSFET circuits analytically and predicts the circuit behaviour well in the sub-micrometer region. Using the model, closed-form expressions for the delay, short-circuit power and transition voltage of CMOS inverters are derived. The delay expression includes input waveform slope effects and parasitic drain/source resistance effects. But the CMOS inverter delay becomes less sensitive to the input waveform slope and short-circuit dissipation increases as the carrier velocity saturation effect in short-channel MOSFETs gets more severe. This drawback is resolved in [7] by including many pre-simulation steps. A more precise model [9] that considers the velocity saturation effects and the gate capacitance of MOSFETs is used to present an expression for propagation delay. The propagation delay expression is for static CMOS logic gates considering short-circuit current and current flowing through gate capacitance and using the \( n^\alpha \) power law MOSFET model that considers velocity saturation effects. The short circuit current is represented by a piecewise linear function that enables detailed analysis of the transient behaviour of a CMOS inverter. These expressions are applied to logic gates made up of series-parallel connected MOSFETs by replacing the series-connected MOSFETs by an equivalent MOSFET. The influence of short-circuit power on delay, is modelled in the expression.

The output load is modelled with an RC model and differential equations are solved in [10]. The resistive-capacitive behaviour of long interconnects which are driven by CMOS gates are analysed. The analysis is based on the \( \delta \)-model of an RC load and is developed for submicron devices. Accurate and analytical expressions for the output voltage waveform, the propagation delay and the short circuit power dissipation are derived by solving the system of differential equations which describe the behaviour of the circuit. The effect of the coupling capacitance between input and output and that of short circuit current are also incorporated in the model.

A delay model for multiple delay simulation [11] for NMOS and CMOS logic circuits is given. For the simple inverter the rise or fall delay time is approximated by a product of polynomials of the input waveform slope, the output loading capacitance and the device configuration ratio. This approach is extended to the case of multiple-input transitions. A model that works at gate-level [12] with the modelling process that includes the characteristics of MOSFETs handles the delay variation according to the kind of gates, input transition time, output load (fan-out) and transistor sizes of the gate. A delay model for CMOS inverter is extracted first and then it is extended to other gates by converting them into an equivalent inverter. This model calculates the delay time regardless of the input transition time, output load (fan-out), or the size of MOSFETs. A systematic method [13] to reduce standard cell library characterization time significantly is to use a simple and physically reasonable logic gate delay model in which delay varies linearly with \( C_i \) and \( t_{ns} \), Determine its region of validity in the \((C_p, t_{ns})\) space and express the delay model coefficients and its region of validity as a function of inverter (or logic gate) size. As device current/capacitance models are not used, the method is general enough to be used with scaling. A simple analytical model [14] that allows the estimation of the propagation delay utilizes a closed-form model of an RC circuit with a linear input to evaluate the propagation delay of CMOS gates or wires in modern VLSI and ULSI processes.

On the basis of the level of abstraction power estimation techniques are broadly classified into [15].

- Low-level power estimation
- High-level power estimation

**Low-level Power Estimation Techniques Are Those That Operate at the Gate Level:** High-level power estimation techniques [16] estimate power dissipation from a design description at a high level of abstraction. The power estimation techniques at the gate level and lower levels of abstraction are broadly classified into [17]:

- Simulation based techniques
- Probabilistic techniques and
- Statistical techniques

Simulation based techniques are the earliest proposed techniques where the average power is calculated by monitoring both the supply voltage and current waveforms. These are too slow to handle very large circuits. In probabilistic techniques, user-supplied input signal probabilities are propagated into the circuit. Statistical techniques do not require any specialized models for the components. The circuit is simulated with randomly generated input vectors until power converges to the average power.
Power estimation techniques are also classified as either static or dynamic. An approach is called static when it is based on propagating a probability or activity measure directly through the logic, in order to estimate the average switching frequency and uses non simulation based methods. Dynamic techniques use traditional simulation models and simulate the circuit, using existing simulation capabilities, for a limited number of randomly selected input vectors (or vector blocks) while monitoring the power.

This paper describes optimisation of a digital CMOS VLSI circuit with respect to power and delay. The transistor level optimisation algorithm uses RC model that deviates from SPICE simulations by 4-5%. The general idea of transistor modelling is to characterize the delay and power contribution of a critical path so as to optimise PDP. To calculate data path delays, the model has used generalized input step voltage and then transformed the results for realistic input signal waveforms.

This paper is summarised as follows: Section II gives a brief introduction of the computation methodology. Section III consists of the proposed work and section IV compares the theoretical and simulation results. Finally conclusions are drawn in section V.

**Computation Methodology:** A fast methodology to calculate the approximate value of the delay and average dynamic power consumption for combinational logic circuits is used. The critical path that is defined as the path between an input and an output that has maximal delay [17]. Optimisation of the critical path helps in the performance improvement of the overall digital circuit. In large digital circuits a method that focuses on the critical path only requires less computational effort compared to the exhaustive simulation approaches. For calculating the propagation delay of D Flip Flop, delay model for a CMOS inverter and CMOS transmission gate is considered as the first step.

The D Flip Flop is modelled as an RC network and the propagation delay for a lumped RC network is given by

\[ \tau_{PH} = 0.69 R_{on} C_L \]

In the expression for propagation delay ‘\( R_{on} \)’ is the average “ON” resistance of the MOS transistor and \( C_L \) is the capacitive load.

An approximate expression for \( R_{on} \) in transmission gate using an average value of current equal to the saturation current of the PMOS (NMOS) transistor is given by [18].

\[ R_{on} = \frac{2V_{dd}}{k_n(V_{dd} - V_{on}) + k_p(V_{dd} - |V_{tp}|)} \]

Load capacitance is given by the expression below [19].

\[ C_L = C_{pl,n} + C_{pl,p} + C_{dh,n} + C_{dh,p} + C_{g,n} + C_{g,p} + C_{load} \]  

(1)

Here \( C_{pl,n} \) and \( C_{pl,p} \) are gate to drain capacitances of the NMOS and PMOS, \( C_{dh,n} \) and \( C_{dh,p} \) are the drain to base capacitances of the NMOS and PMOS and \( C_g \) is the thin oxide capacitance over the gate area.

The table below is used to calculate each of the above capacitance values [20].

<table>
<thead>
<tr>
<th>Capacitance</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( C_{pl,n} )</td>
<td>( 2CGD, W_n )</td>
</tr>
<tr>
<td>( C_{pl,p} )</td>
<td>( 2CGD, W_p )</td>
</tr>
<tr>
<td>( C_{dh,n} )</td>
<td>( K_{on}AdC + K_{pass}APD C JSW )</td>
</tr>
<tr>
<td>( C_{dh,p} )</td>
<td>( K_{on}AdPDC + K_{pass}APD C JSW )</td>
</tr>
<tr>
<td>( C_{g,n} )</td>
<td>( (CGDO + CGSO)W_n + C_{on}W_nL_n )</td>
</tr>
<tr>
<td>( C_{g,p} )</td>
<td>( (CGDO + CGSO)W_p + C_{on}W_nL_p )</td>
</tr>
</tbody>
</table>

The parameters taken from model file are \( V_{on}, |V_{tp}|, CGD, CGSO, CGBO, CGD, CGSO, CGBO, X_n, t_{on}, \mu_n, \mu_p, C_n, C_{MOS}, N_{ox}, N_{ox,n} \) and \( N_{ox,p} \).

\( K_{on} \) is calculated using the expression

\[ K_{on} = \frac{2\sqrt{e}}{V_{t,2} - V_{t,1}} \left( \sqrt{\frac{V_{on}}{V_{t,1}}} - \sqrt{\frac{V_{on}}{V_{t,2}}} \right) \]

where \( V_{t,1} = \frac{kT}{e} \left( \frac{N_A}{N_D} \right) \).

Likewise \( C_{on} = \frac{C_{on,p}}{C_{on,n}} \) and \( k_n = \mu_n C_{on,n} \text{ and } k_p = \mu_p C_{on,p} \).

The Low to High delay of an inverter is given by

\[ \tau_{PH} = \frac{C_{G}}{K_{on}(V_{dd} - |V_{tp}|)} \left[ \frac{2|V_{tp}|}{V_{dd} - |V_{tp}|} + \ln \left( \frac{4(V_{dd} - V_{tn})}{V_{dd} - |V_{tp}|} \right) \right] - c_{m}R_{i} \]

and the High to Low delay of an inverter is given by

\[ \tau_{PH} = \frac{C_{G}}{k_n(V_{dd} - V_{on})} \left[ \frac{2V_{dd}}{V_{dd} - V_{tn}} + \ln \left( \frac{4(V_{dd} - V_{tn})}{V_{dd} - V_{tn}} \right) \right] - c_{m}R_{i} \]
The rise time or fall time of signal through inverter is given by

\[ t_{\text{rf}} = \ln(\theta) R_{\text{on}} C_i \]

Total power consumption is the power consumed in the explicit pulse generator (and the drivers), the D Flip Flop (and the drivers).

The dynamic power consumed is proportional to the capacitance hence the same has been calculated for each MOSFET.

The average dynamic power required to charge and discharge a capacitance 'C_{load}' at a switching frequency 'f' and activity factor '\alpha' is given by [20].

Average Dynamic Power Consumption = \alpha C_{load} V^2 f.

For CMOS gates other than inverter, MOSFETs serial or parallel connected are combined into one equivalent MOSFET [21].

**Optimisation Through Delay and Power Calculations in D-Flip flop**: Delay and Power calculation of the D Flip Flop [22] is given below. It is an explicit pulsed dual edge triggered D Flip Flop. The Fig. 1 below gives the explicit pulse generator of the D Flip Flop.

The D Flip Flop shown in Fig. 2 consists of an input and output stage that is an inverter, whereas the middle stage is a transmission gate, that allows the input to pass through only at the clock edge. The two inverters invert the signal twice and hence the output is the same as input. Two clock pulses that are generated from pulse generator are applied to the MOSFETs (M_5 & M_6).

The circuits are simulated using two buffers (inverters) that drive the flip flop inputs (data and clock) and the output drives a capacitive load of 21fF. An additional capacitance of 3fF is placed after the clock driver. An input data with an activity factor of 25% to reflect the average power consumption is applied at the input. Power consumed in the data and clock drivers is included in our measurements. Clock frequency used is 125MHz. In the above circuit an inverter drives another inverter or a transmission gate which is its load.

The load capacitance at each stage is calculated using expression (1) and there after delay \( \tau_{\text{PHL}} \) and \( \tau_{\text{PLH}} \) are calculated using expressions (2) and (3).

Calculation of Delay

For Inverter 1 (buffer 1) the capacitances calculated are as under

<table>
<thead>
<tr>
<th>Transistor Capacitance</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>( C_{\text{pin}} )</td>
<td>( 3.4 \times 10^{-16} )</td>
</tr>
<tr>
<td>( C_{\text{plp}} )</td>
<td>( 8.01 \times 10^{-16} )</td>
</tr>
<tr>
<td>( C_{\text{dln}} )</td>
<td>( 2.46 \times 10^{-16} )</td>
</tr>
<tr>
<td>( C_{\text{dlp}} )</td>
<td>( 5.88 \times 10^{-16} )</td>
</tr>
<tr>
<td>( C_{\text{gln}} + C_{\text{glp}} )</td>
<td>( 11.4 \times 10^{-16} )</td>
</tr>
<tr>
<td>( C_L )</td>
<td>( 31.15 \times 10^{-16} )</td>
</tr>
</tbody>
</table>

Given the value of \( C_L = 31.15 \times 10^{-16} \) F the value of \( \tau_{\text{PHL}} = 11.37 \times 10^{-12} \) sec.

Initial rise or fall time was 10ps.

Hence actual delay time is 12.42ps.

\[ t_{\text{rf}} = 2.2 \tau_{\text{g}} \]

\[ t_{\text{rf}} = 27.32 \times 10^{-12} \text{sec} \]

\[ R_p = 3.65 \times 10^3 \text{ ohms} \]

For Inverter 2 (buffer 2) the value of \( C_L = 31.15 \times 10^{-16} \) F and \( R_p = 3.65 \times 10^3 \) ohms the value of \( \tau_{\text{PHL}} \) is 11.37 \( \times 10^{-12} \) sec and initial rise or fall time was 27.32ps. Hence actual delay time is 17.77ps.
Fig. 4: Delay model of inverter driving a transmission gate

New rise or fall time is $t_{\text{r/f}} = 39.1 \times 10^{-12} \text{sec}$.

For Inverter 3 and Transmission gate the value of $C_1 = 47.4 \times 10^{-16} \text{F}$ and $C_2 = 66.7 \times 10^{-16} \text{F}$ and $R_1$ is 3650 ohm and $R_2$ of transmission gate is calculated as 678 ohm.

The delay of inverter and transmission gate combination is given by

$\tau_{\text{in}+\text{out}} = R_1C_1 + (R_2 + R_2)C_2$

Propagation delay $\tau_{\text{PHL}} = 66.7 \times 10^{-12} \text{sec}$ and initial rise or fall time was 39.1ps. Hence actual delay time is 69.5ps and $t_{\text{r/f}} = 152.9 \times 10^{-12} \text{sec}$.

For Inverter 4 the value of $C_1 = 289.0 \times 10^{-16} \text{F}$ and $R_1 = 3.65 \times 10^3 \text{ohms}$ the value of $\tau_{\text{PHL}}$ is $105.5 \times 10^{-12} \text{sec}$ and initial rise or fall time was 152.9ps. Hence actual delay time is 130.4ps and $t_{\text{r/f}} = 286.9 \times 10^{-12} \text{sec}$.

Total Delay is the sum of the two delays and is equal to 199.9ps.

**Calculation of Power Consumption:** Total power consumption is the power consumed in the explicit pulse generator (and the drivers) and the D Flip Flop (and the drivers).

Power Consumption in a digital circuit is given by [1]

$P_{\text{total}} = I_{\text{leak}}V_{\text{dd}} + aC_1V_{\text{dd}}^2f + \frac{t_r + t_f}{2}V_{\text{dd}}I_{\text{peak}}f$

The power dissipation of a CMOS circuit is instead dominated by the dynamic power dissipation resulting from the charging and discharging of capacitances. The dynamic power consumed is proportional to the capacitance hence it has been calculated for each MOSFET.

To minimise the power consumption of a CMOS circuit the dynamic power consumption has to be reduced. It depends on the capacitances in the MOSFET and the load capacitances. The MOSFET capacitance depends on the W and L of the transistor. The average dynamic power required to charge and discharge a capacitance $C_{\text{load}}$ at a switching frequency $f$ and activity factor $\alpha$ is given by

$\frac{\alpha}{2}C_{\text{load}}V_{\text{dd}}^2f$

**Dynamic Power Consumption** $= \alpha C_{\text{load}}V_{\text{dd}}^2f$.

Here $f = 125\text{MHz}$ and $V_{\text{dd}} = 1.8\text{V}$ and $\alpha = 0.25$. Hence $V_{\text{dd}}^2f = (1.8)^2 \times 125 \times 10^6 = 4.05 \times 10^8$.

Total capacitance in explicit pulse generator and drivers is $426.01 \times 10^{-16} \text{F}$.

$= 37.52 + 87.45 + 54.56 + 60.44 + 40.55 + 60.44 + 37.96 + 47.09$

$= 418.05 + 37.96$

Capacitance in Flip Flop and drivers is $570.35 \times 10^{-16} \text{F}$.

Thus power consumed in D Flip Flop is

$= [405 \times 10^6 \times (418.05 + 2 \times 37.96) + 0.25 \times 405 \times 10^6 \times 570.35] \times 10^{-16} \text{W}$

$= 20.00 \times 10^4 + 5.78 \times 10^4 = 25.78\mu\text{W}$. 

Fig. 5: Flow chart of the optimisation technique
Table 1: Iterations leading to optimisation of D Flip Flop

<table>
<thead>
<tr>
<th>Iterations Performed</th>
<th>Power (µW)</th>
<th>Power (µW)</th>
<th>Delay (ps)</th>
<th>Delay (ps)</th>
<th>PDP (×10^-3 J)</th>
<th>PDP (×10^-3 J)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(Theoretical)</td>
<td>(Simulated)</td>
<td>(Theoretical)</td>
<td>(Simulated)</td>
<td>(Theoretical)</td>
<td>(Simulated)</td>
</tr>
<tr>
<td>1st</td>
<td>72.01</td>
<td>83.05</td>
<td>283.53</td>
<td>285.8</td>
<td>20.4</td>
<td>23.6</td>
</tr>
<tr>
<td>2nd</td>
<td>48.9</td>
<td>56.4</td>
<td>240.7</td>
<td>242.6</td>
<td>11.8</td>
<td>13.7</td>
</tr>
<tr>
<td>3rd</td>
<td>36.34</td>
<td>41.91</td>
<td>219.4</td>
<td>221.22</td>
<td>7.97</td>
<td>9.3</td>
</tr>
<tr>
<td>4th</td>
<td>28.31</td>
<td>32.65</td>
<td>207.8</td>
<td>209.5</td>
<td>5.9</td>
<td>6.9</td>
</tr>
<tr>
<td>Final</td>
<td>25.78</td>
<td>29.73</td>
<td>199.9</td>
<td>201.5</td>
<td>5.2</td>
<td>5.9</td>
</tr>
</tbody>
</table>

The output fall time is kept slightly larger than the input rise time where ever possible to reduce short circuit current. Difference between simulation values and computed values occurs due to leakage power and short circuit power.

**Circuit Optimisation:** In order to optimise the circuit in view of power and delay load capacitance of each transistor is calculated. Dynamic power and delay are then derived as shown above. Both the parameters are dependent upon capacitance which is in turn dependent upon the width of the transistor. The width of the transistors most likely to reduce power and delay are adjusted and theoretical computations and simulations are performed to ascertain the improvements. The flow chart below gives the procedure used to design the circuit.

The following table gives the iterations that lead to the optimisation of the D Flip Flop.

**Comparison of Results:** The parameter values after optimisation, from simulation for delay and power consumption are 201.5ps, 29.73µW. The theoretically derived values are 199.9ps and 25.78µW. The circuit optimisation techniques and algorithms used in VLSI circuit design must have low computational complexity. In a large digital VLSI circuit repeated transistor size adjustments to improve power and performance becomes too much time consuming. In these conditions theoretical calculation of power and delay in the critical path and transistor size adjustments therein can reduce design time significantly.

**CONCLUSIONS**

This paper discusses the theoretical calculation of power and delay in digital VLSI circuits. The theoretically calculated values are close to the simulated results. The theoretical calculations are used to optimise the D Flip Flop circuit. The width of the transistor most likely to reduce delay and power is adjusted till theoretical results gave minimum power and delay. Then simulation is used to validate the design. In a large digital VLSI circuit theoretical model that can give approximate results, can help reduce design time by optimisation of the critical path.

**REFERENCES**


