Effects of buffer insertion on the average/peak power ratio in CMOS VLSI digital circuits

Antonio J. Acosta, José M. Mora, Javier Castro, and Pilar Parra

Instituto de Microelectrónica de Sevilla-CNM-CSIC/Universidad de Sevilla, Spain

ABSTRACT

The buffer insertion has been a mechanism widely used to increase the performances of advanced VLSI digital circuits and systems. The driver or repeater used to this purpose has effect on the timing characteristics on the signal on the wire, as propagation delay, signal integrity, transition time, among others. The power concerns related to buffer insertion have also received much attention, because of the low power requirements of modern integrated systems. In the same way, the buffer insertion has strong impact on the reliability of synchronous systems, since the suited distribution of clock requires reduced or controlled clock-skew, being the buffer and wire sizing, a crucial aspect. In a different way, buffer insertion has been also used to reduce noise generation, especially in heavily loaded nets, since the inclusion of buffers help to desynchronize signal transitions. However, the inclusion of buffers of inverters to improve one or more of these characteristics have often negative effect on another parameters, as it happens in the average and peak of supply current. Mainly, the inclusion of a buffer to reduce noise (peak power), via desynchronizing transitions, could introduce more dynamic consumption, but reducing the short-circuit current because of the increment of signal slope. Thus, the average/peak current optimization can be considered a design trade-off. In this paper, the mechanism to obtain an average/peak power optimization procedure are presented. Selected examples show the feasibility of minimizing switching noise with negligible impact on average power consumption.

Keywords: Switching noise generation, submicron CMOS VLSI, mixed analog/digital circuits, buffer insertion, buffer repeater.

1. INTRODUCTION

As the progress of deep sub-micron process technology, the noise margins are considerable reduced whereas the power supply fluctuation is exacerbated. The presence of large current peaks on the power and ground lines is a serious concern for designers of synchronous digital circuits [1-5]. Although a lot of research efforts have been made in the minimization of total power consumption of circuits [6-9], there are not many works that pay attention in minimization of the peak current of them [10-13]. Peak current is a primary concern in the design of power distribution networks. In state-of-the-art VLSI systems, power and ground lines must be over-dimensioned in order to account for large current peaks. Such peaks determine the maximum voltage drop and the probability of failure due to electro-migration. Peak currents are caused by simultaneous switching of highly loaded clock lines and by the signal propagation through the sequential logic elements. So, for current circuits, the control of power supply noise becomes very important in terms of circuit performance and reliability, demanding a reduction of peak power (current) as well as power consumption [3].

The basic idea analyzed in this paper is the inclusion of buffers (inverters) to desynchronize signal transitions. This inclusion has effect on both the signal slope and the propagation delay, in such a way that, for peak current we expect a reduction because of signals do not arrive at the same time to the load, and for average power, the short-circuit current is expected to be reduced due to the increment of signal slope, but dynamic consumption could be increased. As a consequence, a design trade-off can be found to convert the buffer insertion technique in a good power and noise saving tool. Our contribution does not provide a global and optimum solution to buffer insertion but aims to show the effects of buffer
insertion on the average/peak power ratio. A current peak reduction of around a 40% has been observed, without any increase of power consumption.

The organization of the paper is as follows: Section 2 shows an introduction to the CMOS inverter. Section 3 will focus on the low power techniques, the effect on switching noise and the methodology used in the process. In Section 4 the proposal is validated with HSPICE simulations. Finally, Section 5 includes the summary of results and the conclusions.

2. CMOS INVERTER/REPEATER

Much work has been reported in literature about CMOS inverters. The CMOS inverter can be considered as the simplest buffer or repeater in VLSI circuits. The three components of power consumption in a CMOS inverter are dynamic, short-circuit and leakage. To toggle between logic zero and logic one, capacities have to be discharged and charged. The electric current that flows during this process causes a power dissipation, this is dynamic power consumption. The current is dependent on the capacitive output load and the supply voltage. CMOS inverter consist of a pull-up (PMOS) and pull-down (NMOS), which have a finite input fall/rise time larger than zero. During this short time interval, when the PMOS and NMOS transistors are conducting both, a current flows from power to ground. This current is called short-circuit-current. Dynamic and short-circuit power depend on switching activity, so no power should be lost during idleness of a CMOS circuit. But the existence of leakage currents shows another piece of reality. These are due mainly to currents flowing through the reverse biased diodes that are formed between the diffusion regions and the substrate, and flows through transistors that are non-conducting.

Figure 1 shows a CMOS buffer driving a load and its equivalent symbolic representation. In figure 1a the input change from logic one to logic zero (fall transition), then the load capacitance is charged and dynamic current flows through the PMOS transistor, while in this case through the NMOS transistor only flows the short-circuit current (short circuit path between supply rails during switching). There is also leakage currents due to leaking diodes and transistors, but this current is not considered in this analysis. In figure 1b the input change from logic zero to logic one (rise transition), then the load capacitor is discharged through NMOS (dynamic power consumption) and through PMOS only flows short-circuit current.

The charge and discharge of the load capacitance requires current from the supply source. As transition time increases, the peak in supply current will be reduced, since the same amount of charge will be supplied during more time. Some HSPICE results showing this dependence for a 0.13 μm CMOS inverter are depicted in figure 2. Results in table of figure 2 show the increment in propagation delay (tpLH), fall time (tfall) and short-circuit current (I_p) as input slope.

![Figure 1](http://proceedings.spiedigitallibrary.org/)

Figure 1. Power consumption in an inverter, a) in a fall transition, where current flows through PMOS to charge the load (dynamic power consumption), while through NMOS only flows short-circuit current. b) Rise transition, where current flows through NMOS to discharge the load (dynamic power consumption), while through PMOS only flows short-circuit current, and c) inverter equivalent symbolic representation.
decreases. In a smooth way, a reduction in the peak and an increment in the average of total current \( I_N \) is also noticed. Related to average power consumption, short-circuit power increases when transition time increases, due to NMOS and PMOS simultaneous operation in linear and saturation mode, respectively for the output falling transition. Managing these effect, a trade-off between peak current reduction and average power-timing performances degradation can be found if a suited value for input slope is considered.

3. LOW POWER/LOW SWITCHING NOISE AND METHODOLOGY

First, an introduction of the importance of low power consumption techniques and their effects on switching noise is presented. Next, the methodology followed to carry out the simulations to show the effects of buffer insertion is explained.

3.1. Low power consumption and switching noise generation

When designers recognized power consumption as a design constraint, simple models were created. Power per MHz is still a commonly used representation of a component. With a closer look at power dissipation, it becomes obvious that the subject is not that simple. Electric current is not constant during operation, and peak power is an important concern too, because of the device would malfunction due to electro-migration and voltage drops even if the average power consumption is low. So, peak power is an important parameter to take into account in the simulations as well as average power.

Reduction of the supply voltage has a quadratic effect on the reduction of power consumption but it also increases the delay. It furthermore reduces the switching noise generated by single gates, but noise margins are reduced too. To balance the delay increase produced by the supply voltage reduction, some authors suggest increasing parallelism or using pipeline structures. Both techniques increase simultaneous activity and noise. There is a trade-off between low voltage-induced switching noise reduction and the increase in simultaneous activity that has to be dealt with.

To sum up, it is not clear that all low power techniques necessary reduce switching noise. Some of them could even increase this (like power-down modes). The reduction of the circuit activity does not necessarily reduce maximum simultaneous activity by the same amount, and this question indeed requires further analysis [3].
3.2. Schematic disposal

To check the feasibility of buffer insertion with regard to switching noise, two schematic disposal have been proposed. First, a schematic with inverters that simulate a circuit and a heavy load is used (figure 3a). In a UMC 0.13 µm technology, an inverter with PMOS width $W_p=50 \mu m$ represent a circuit (note that in all cases $W_n=W_p/2$, and transistor channel $L=0.12 \mu m$), while an inverter with $W_p=1000 \mu m$ represent a heavy load. In this situation we can measure the original circuit characteristics, such as power consumption and peak power supply current. Then we modify the schematic structure with the inclusion of a buffer (figure 3b). As it is well known, when a signal arrives at a circuit with an equivalent heavy load, a lot of power is consumed and a large current peak is generated. We can divide the arrival of the signals to the circuit load with the introduction of a buffer in part of the load, obtaining a desynchronization in signal transitions. Then we can simulate the average current (dynamic and short circuit) and peak current in both cases, with and without buffer insertion, in order to compare the values. A reduction in peak power supply current and short circuit current are expected due the buffer insertion, while a penalty to pay is an increase in propagation time (delay).

For the simulation we use HSPICE and a periodic input signal “In” ($T=10 \text{ ns}$, $t_f=t_r=100 \text{ ps}$, $V_{dd}=1.2 \text{ v}$, where $T$ is the period, $t_f$ and $t_r$ are the fall and rise time respectively, and $V_{dd}$ is the power supply).

4. SIMULATION RESULTS

We obtain measurements of average and peak currents for the fall and rise transition. We measure the current through the PMOS and NMOS transistors in every signal transition, and in this way we can identify the current due to short circuit and dynamic. In the average current we notice that due to dynamic power, short circuit and total average current per transition, for both cases, the one without buffer insertion and the one with buffer insertion, as can be seen in table 1. We also put in parentheses, in the case with buffer insertion, the ratio between both cases, so we can see the improvement achieved with regard to the original case without buffer insertion. As can be seen in table 1, the average current is reduced when the buffer is inserted, around a 20% in dynamic power, and a 50% in the case of short-circuit current (this is due to the higher signal slope with buffer insertion, and as a consequence the short circuit current must be less than in the case without buffer insertion). The total power consumption is reduced around a 25% since the dynamic power is the dominant.

In the case of peak power supply current, we did measurements of peak at each transition (“In” fall and “In” rise) and we measure on $V_{dd}$ line, ground line, and total peak current (see table 2). Again we annotate the ratio between the case with and without buffer insertion, so that we can see how the peak current is reduced. Results show that this reduction is around a 40%, which makes the buffer insertion a good technique to reduce switching noise without any increase in power consumption, even a reduction is obtained as we saw in results in table 1, and then, a good trade-off between both parameters is achieved. In figure 4 you can see the current waveform in fall and rise transition for both, the case without buffer insertion and the case with buffer insertion. In the current waveform in the case with buffer insertion we notice that
there are two peaks for each transition, due to the desynchronization of the signal arrival, but these peaks are more or less inside the current waveform obtained without buffer insertion. That is why power consumption is not increased. And regarding to propagation time, in the case without buffer insertion is 0.14 ns, and with buffer insertion is 0.18 ns. So, as a penalty, an increase of around a 25% in delay occur with the introduction of the buffer. Then, we will focus on these two parameters, peak current and delay, and we will look for a trade-off between them.

Table 1: Average current (mA), in the cases without and with buffer insertion, for fall and rise time, dynamic and short circuit (SC) current, and total current for each transition

<table>
<thead>
<tr>
<th></th>
<th>Without buffer insertion</th>
<th>With buffer insertion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Fall</td>
<td>Rise</td>
</tr>
<tr>
<td>Dynamic</td>
<td>2.02 (0.91)</td>
<td>1.693</td>
</tr>
<tr>
<td>SC</td>
<td>0.912</td>
<td>0.584</td>
</tr>
<tr>
<td>Total (mA)</td>
<td>2.932</td>
<td>2.277</td>
</tr>
</tbody>
</table>

Table 2: Peak current (mA) in power supply line, ground line, and total, for both cases, without and with buffer insertion.

<table>
<thead>
<tr>
<th></th>
<th>vdd</th>
<th>gnd</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Without buffer</td>
<td>With buffer</td>
<td>Without buffer</td>
</tr>
<tr>
<td>peak current (mA) - fall</td>
<td>64.95</td>
<td>40.15 (0.62)</td>
<td>64.91</td>
</tr>
<tr>
<td>peak current (mA) - rise</td>
<td>58.7</td>
<td>36.82 (0.63)</td>
<td>58.81</td>
</tr>
</tbody>
</table>

Figure 4. HSPICE simulation results of current waveform for fall and rise transition of input signal and for the two configurations, with two inverter and with five inverters.

there are two peaks for each transition, due to the desynchronization of the signal arrival, but these peaks are more or less inside the current waveform obtained without buffer insertion. That is why power consumption is not increased. And regarding to propagation time, in the case without buffer insertion is 0.14 ns, and with buffer insertion is 0.18 ns. So, as a penalty, an increase of around a 25% in delay occur with the introduction of the buffer. Then, we will focus on these two parameters, peak current and delay, and we will look for a trade-off between them.
In order to complete the results, some modifications were made on the dimensions of the inverters (figure 5) used as buffer insertion for simulation, depending on the size of the buffer. We change the PMOS width $W_p$ between 10 $\mu$m and 60 $\mu$m, being the NMOS width $W_n = W_p/2$.

With these new simulations we can see which buffer dimension is the best in terms of peak current and delay.

<table>
<thead>
<tr>
<th>buffer $W_p/W_n$</th>
<th>10/5</th>
<th>20/10</th>
<th>30/15</th>
<th>40/20</th>
<th>50/25</th>
<th>60/30</th>
</tr>
</thead>
<tbody>
<tr>
<td>peak current (mA) -fall</td>
<td>36</td>
<td>38.1</td>
<td>38.49</td>
<td>38.82</td>
<td>40.15</td>
<td>41.61</td>
</tr>
<tr>
<td>peak current (mA) -rise</td>
<td>31</td>
<td>32.8</td>
<td>33.56</td>
<td>34.31</td>
<td>36.82</td>
<td>39.25</td>
</tr>
<tr>
<td>$t_p$ (ns)</td>
<td>0.40</td>
<td>0.26</td>
<td>0.22</td>
<td>0.19</td>
<td>0.18</td>
<td>0.17</td>
</tr>
<tr>
<td>noise-delay product (pJ/v)</td>
<td>14.4</td>
<td>9.9</td>
<td>8.5</td>
<td>7.4</td>
<td>7.2</td>
<td>7.1</td>
</tr>
</tbody>
</table>

Figure 5. Modifications on buffer insertion dimensions. There are six different cases of buffer insertion for simulation, depending on the size of the buffer. We change the PMOS width $W_p$ between 10 $\mu$m and 60 $\mu$m, being the NMOS width $W_n = W_p/2$.

Table 3: Peak current (fall and rise transition), propagation time ($t_p$), and noise-delay product with modifications on buffer insertion widths.

In order to complete the results, some modifications were made on the dimensions of the inverters (figure 5) used as buffer insertion [14-16], and we measure the peak current in fall and rise transitions, and also the propagation time. In table 3 are the results of these simulations. As can be seen from the results, as the dimension of the buffer sizing is decreased, the peak current is reduced (figure 6a), but the propagation delay is increased (figure 6b). If we look for a trade-off between peak current and propagation delay, we can see figure 6c, where the noise-delay product is represented. As can be seen, the lowest values in buffer sizing give the worst values of the peak current and propagation time trade-off.

5. SUMMARY AND CONCLUSIONS

In this paper, we have introduced an efficient way for reducing peak current without significant increase in other parameters, such as average power consumption or propagation delay. In fact, a reduction in power consumption is also achieved. This simple methodology shows that the buffer insertion is a good option in terms of low power consumption and switching noise generation. Measurements, using HSPICE simulations, of both parameters were done in an UMC 0.13 $\mu$m technology, and also several buffers with different sizes were used in order to obtain the one with the best noise-delay product. Experimental results show around a 40% reduction in peak current (switching noise) and a 30% reduction in average power consumption, while the propagation time got worse around a 20%.

ACKNOWLEDGEMENTS

This work has been sponsored by the Spanish MEC TEC2004-01509 DOC and the Junta de Andalucía TIC2006-635 Projects.
REFERENCES
6. J.L. Neves and E.G. Friedman, “Minimizing Power Dissipation in Non-Zero Skew-Based Clock Distribution Net-

Figure 6. Modifications on buffer insertion width. Measurements of a) peak current, b) propagation time (delay), and c) noise-delay product.