

A Capacitive Boosted Buffer Technique for High-Speed Process-Variation-Tolerant Interconnect in UDVS application

Saihua Lin, Yu Wang, Rong Luo, Huazhong Yang
 EE Dept. of Tsinghua University, Beijing, China, 100084
linsh@mails.tsinghua.edu.cn, {yu-wang, luorong, yanghz}@tsinghua.edu.cn

Abstract*

In this paper, we propose a new capacitive boosted buffer technique that can be used in high speed interconnect for ultra-dynamic voltage scaling (UDVS) application with the process variation effect mitigated. The circuit is simple and fully compatible with digital CMOS technology. Implemented in a standard 0.18 μm CMOS technology, the circuit is shown applicable for both sub-threshold circuit and above threshold circuit without the problem of short current. Simulation results demonstrate the conclusion that the proposed new buffer is more robust to load, process, voltage, and temperature (PVT) variations. When applied to a simple H-tree clock network, the proposed buffer can reduce the skew by 5.5X when compared to that of the traditional buffer.

1. Introduction

Power and process variations are two challenges that prevent us from integrating more and more transistors together on a chip and from ensuring their function properly across the wafer. For low power circuit design which is important for portable devices and sensor networks, researchers have proposed many kinds of methods to reduce power, such as the sub-threshold circuit techniques [1], MTCMOS techniques [2], multiple supply voltages (MSV) techniques [3], power gating & sleeping transistor techniques [4] and so on. Among these techniques, the sub-threshold circuit design techniques can achieve significant power reduction. However, since the circuit operates in the sub-threshold region where the current is exponentially dependent on the gate-source voltage and the threshold voltage, any small variation in the gate-source voltage and the threshold voltage can cause large drain current variation. As a result, the delay varies greatly and we have to keep enough delay margins when designing sub-threshold circuits.

Previously, there have been several methods to cope with the process variation problems, including the transistor sizing techniques [5], substrate biasing techniques [6], boosting techniques [7] and so on. For

interconnect design, the boosting technique proposed in [7] can effectively mitigate the process variation problems in sub-threshold interconnect. However, the circuit proposed in [7] is not applicable for above threshold circuit design and therefore is not suitable for ultra-dynamic voltage scaling (UDVS) applications. As the supply voltage increases, the short current increases so fast that the performance gains in delay and variation tolerance are compromised by the increase of power [7].

Recently, the authors in [8] proposed self-timed regenerators (STR) for high-speed and low-power interconnect. However 28 transistors are used in a STR and it self can not reduce power. The reason why power is reduced is simply because less STR are used than traditional buffer when two paths have the same delay. Furthermore, MTCMOS techniques are used which will increase the design cost.

In this paper, we propose a capacitive boosted buffer technique for high speed interconnect design with the tolerance to load variation and PVT variations. It is compatible with digital CMOS technology and can be operated in both sub-threshold region and above threshold region without the problems of short current. As a result, it is suitable for UDVS applications [1]. Simulation results show that when applied to an H-tree clock distribution network, the skew of the new buffer network is reduced by 5.5X compared to the traditional buffer network.

The rest of the paper is organized as follows. Section 2 describes the proposed new buffer circuit. Section 3 gives the verification results. Finally in Section 4, we make a conclusion of this paper.

2. Proposed New Buffer Circuit

2.1. Circuit Principle

Fig. 1(a) shows the boosted buffer proposed in [7]. Carefully analyze this circuit, we can find two problems: 1, the circuit is not suitable for above threshold interconnects because of short current. 2, the circuit occupies a large number of transistors.

From Fig. 1(a) we can see that the maximum gate-source voltage of the output PMOS transistor P1 is $-V_{DD}$ and the minimum gate-source voltage of the output NMOS transistor N1 is V_{DD} . If the circuit is operated in the sub-threshold region, V_{DD} is smaller

*This project is supported by NSFC 90707002, 60506010. It is also partly supported by Basic Research Foundation of Tsinghua SIST (No SIST2017).

than V_{TH} , and then a direct path always exists between VDD and GND, which will cause extra power consumption. Here V_{TH} is the absolute value of NMOS threshold voltage V_{THN} or PMOS threshold voltage V_{THP} (suppose $V_{THN} = -V_{THP}$). Although this effect can be minimized by using minimum sized P1 and N1, it can not be handled in above threshold circuit operation. Since $-V_{DD}$ is smaller than $-V_{THP}$ and VDD is larger than V_{THN} , large short current will exist in Fig. 1(a). Furthermore, it has totally 18 devices, much more than the traditional buffer which has only 4 devices. As a result, the power consumption is much larger in above threshold circuit operation.

Fig. 1 (b) shows the proposed new boosted buffer in this paper. Here, the boosted circuits are inserted into the sources of MOS transistors rather than the gates. Therefore, the drain of P2, N2 transistors has a $3 \times V_{DD}$ voltage swing rather than the VDD voltage swing. After a normal inverter, the high voltage swing signal is converted back to the normal voltage swing signal. When designing the boosting circuit, we have to make sure that it can prevent the boosted voltages from feeding back into the power supply and ground.

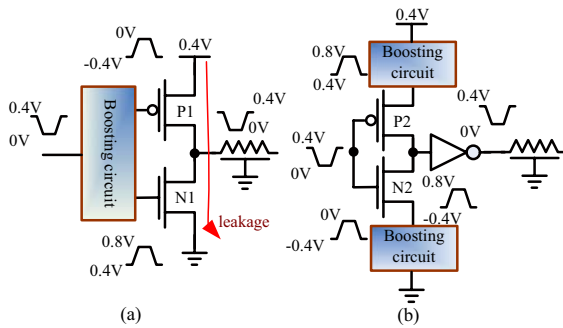


Figure 1. Circuit principle.

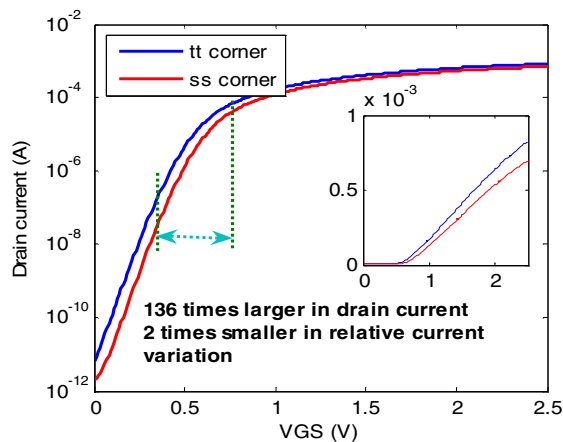


Figure 2. Drain current vs. VGS for a NMOS transistor.

Fig.2 shows a NMOS transistor drain current versus VGS relationship plotted in both log axis and in linear

axis. The threshold voltage of the NMOS transistor is about 0.5V in HJTC 0.18 μm CMOS technology. When the VGS voltage changes from 0.4 V to 0.8 V, the drain current changes from 0.6107 μA to 83.14 μA , which is about 136 times larger than the former in tt corner case. When operated in ss corner, the drain current at 0.4 V varies to 0.1308 μA , a change of 78.58% compared to the current in the tt corner. However, if we boost the VGS to 0.8 V, then the drain current at 0.8 V is 52.23 μA in ss corner, a change of 37.18%. Therefore, the relative current variation is about 2 times smaller if we use boost technique.

Similarly, if the circuit does not operate in the sub-threshold region i.e. we boost the VGS from 1V to 2V, then the relative current variation is reduced from 27.48% to 16.42%. Although the reduction is not as significant as that of the sub-threshold case, the 40.25% variation tolerance improvement is still very promising compared to any other techniques.

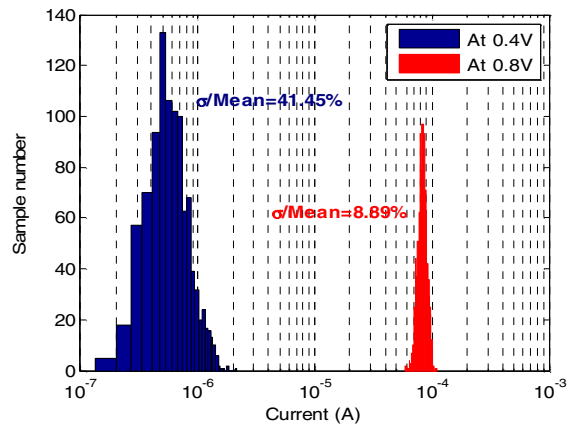


Figure 3. Current variation with different gate voltages.

The corner analysis can be seen as an inter-die analysis where the variation in a die is global. However, as technology continues to advance, the intra-die variation is becoming significant. We then compare the drain current variation by Monte Carlo simulation when the VGS voltage is set to 0.4 V and 0.8 V respectively in Fig. 3. It is assumed here that the NMOS transistor has a 15% inter-die variation and 10% intra-die variation. We can see that for 0.4 V case, the standard deviation of current divided by the average current is about 41.45%. However, if we boost the VGS to 0.8 V, the standard deviation of current divided by the average current is reduced to 8.89%, about 4.66X of reduction.

2.2. Circuit Architecture and Operation

The circuit is implemented in HJTC 0.18 μm CMOS technology as shown in Fig. 4. The threshold voltage for NMOS (PMOS) transistors is about 0.5 V (-0.5 V). Fig. 5 shows its operating waveform where VDD = 0.4

V. Boosted VDDH = 0.52 V, and boosted VSSH = -0.13 V. The clock frequency is 2 MHz for sub-threshold circuit. Since in digital CMOS technology capacitor is always not available, therefore in Fig. 4 P5 and P6 are connected to form two boost capacitors.

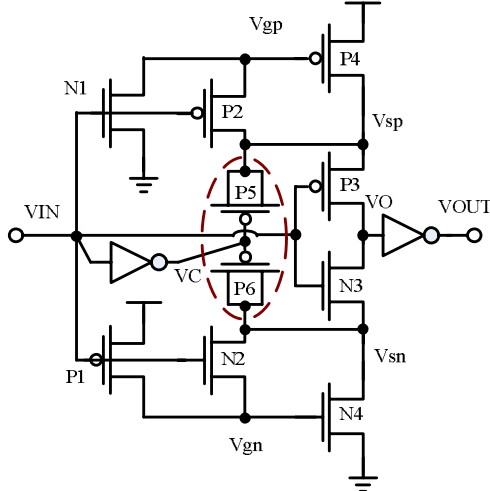


Figure 4. Proposed new buffer circuit.

The circuit is operated as follows. If initially VIN is 0 V then VC is 0.4 V. Since P1 is turned on and N2 is cut off, Vgn is 0.4 V and Vsn is 0 V because N4 is turned on. Suppose VIN transits from 0 V to 0.4 V, VC becomes 0 V. Since the P6 capacitor voltage can not change immediately, Vsn is pulled down to -0.13 V because of parasitic capacitances. Since N2 and N3 are turned on by VIN now, therefore Vgn and the gate voltage of the output inverter VO are both -0.13 V now. As a result, the VOUT is pulled up to 0.4 V quickly. Similar analysis can be applied to the reverse transition case of VIN.

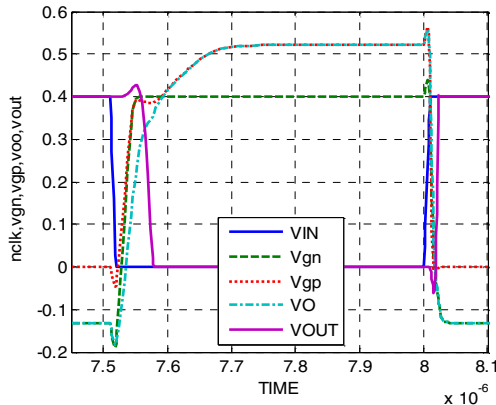


Figure 5. Circuit waveform.

2.3. Design Consideration for Ultra-Dynamic Voltage Scaling (UDVS) System

Dynamic voltage scaling (DVS) has become a standard approach for reducing power when performance requirements vary. In [1], the authors

proposed an ultra-dynamic voltage scaling (UDVS) technique that can provide a practical method for extending DVS into sub-threshold region. As previously discussed, the proposed new boosted buffer is very suitable for UDVS application. However, when using this buffer in the UDVS system, we should take some considerations.

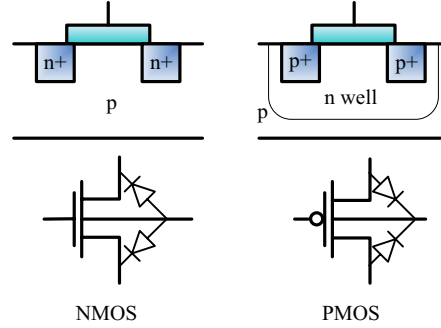


Figure 6. NMOS and PMOS transistor.

For sub-threshold application, the maximum supply voltage is V_{TH} . Therefore, maximum $\{V_{sp}\} = 2 \times V_{TH}$ and minimum $\{V_{sn}\} = -V_{TH}$. As a result, the parasitic substrate-drain or substrate-source diode will not be turned on if V_{TH} is smaller than diode threshold voltage V_{THD} , which is about 0.6 V. However, if the circuit is operated in the above threshold region, these diodes can be turned on and as a result the maximum or minimum boosted voltages are limited. For example, V_{sn} is limited to $-V_{THD}$ if the N4 substrate is connected to ground 0. V_{sp} is limited to $V_{DD} + V_{THD}$ if the n-well of P4 is connected to VDD. Otherwise, large junction leakage current will occur.

One of the solutions is to connect the n-well of PMOS transistor to the high voltage terminal, i.e. V_{sp} terminal for P4, and to connect the separated NMOS substrate to the low voltage terminal, i.e. V_{sn} terminal for N4 in Fig. 4. Although it is easy for PMOS transistor to do this modification, it is not easy for NMOS transistor because generally in CMOS technology, all the NMOS substrates are connected together. As a result, triple-well technology is needed. Otherwise for some other NMOS transistors due to decreased VGS, V_{THN} will be increased.

Another solution is simply not to boost the node voltage too much. This is more practical because the boosted voltage is:

$$\Delta V = \frac{C_{boost}}{C_{boost} + C_{parasitic}} V_{dd} \quad (1)$$

where C_{boost} is the boost capacitance, $C_{parasitic}$ is the parasitic node capacitance. For MOS capacitor, C_{boost} is not very large and therefore, the boosted voltage can not reach $2 \times V_{DD}$ even for sub-threshold circuit. For

example, in the previous section, for 0.4 V supply voltage, the high boosted voltage is only 0.52 V. Another virtue of this solution is that the transistor reliability is not affected much because the gate voltage is not much too high.

3. Circuit Verification

3.1. Buffer Sensitivity to Load Capacitance

We first optimize the proposed buffer and the traditional buffer under 0.4 V supply voltage so that they have the similar rise/fall delay. Then we add load capacitances to evaluate the buffer delay sensitivity to the load. The channel area for new buffer includes the transistor area $0.5112 \mu\text{m}^2$ and the MOS capacitance area $3.6 \mu\text{m}^2$ due to the low boost efficiency of MOS capacitance. The channel area for the traditional buffer is $0.396 \mu\text{m}^2$. The boosted high voltage is 0.503V and the boosted low voltage is -1.14 V. Fig. 7 shows the simulation results when the supply voltage is 0.4 V and 1 V. Here the average delay means the average of the rise and fall delays.

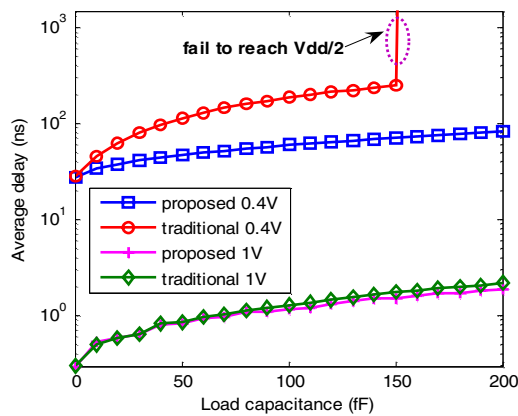


Figure 7. Buffer sensitivity to load capacitance.

From Fig. 7 we can see that when operated in sub-threshold region, the average delay of the traditional buffer increases very fast versus load capacitance. When the load capacitance is larger than 160 fF, the traditional buffer fails to rise to $V_{dd}/2$ point (hence delay is infinite). Therefore, we have to size up the traditional buffer. However, for the new proposed buffer the delay increases more smoothly and when the load capacitance is 200 fF, the average delay is only 80.86 ns.

We also evaluate the performance when the circuit operates in the above threshold region. We set the supply voltage to 1 V and the results are also provided in Fig. 7. We can see that the new buffer is still superior to the traditional buffer although the performance gain is not as significant as in the sub-threshold case. For example, when the load capacitance is 200 fF, the delays of the new buffer and the traditional buffer are 1.799 ns and 2.081 ns

respectively. Therefore, the delay reduction is only 13.55%.

The reason inherited is that if we boost the circuit from subthreshold area to the above threshold area, the delay variation changes from exponentially dependent form to quadratically (saturate region) or linearly (triode region) dependent form. However, if the circuit already operates in the above threshold region, the boost technique only reduces the quadratic part and increase the linear part delay variation. Therefore, the technique efficiency is reduced when operated in above threshold region.

3.2. Buffer Sensitivity to PVT Variation

In addition to the superior load sensitivity, we also perform PVT variation tolerance comparison between these two buffers. We first assume NMOS transistor V_{TH0} has an inter-die variation of 15%, PMOS transistor V_{TH0} have a 15% inter-die variation and 10% intra-die variation. Then we do 1000 Monte Carlo analysis and the result is shown in Fig. 8. The load capacitance is set to 20 fF.

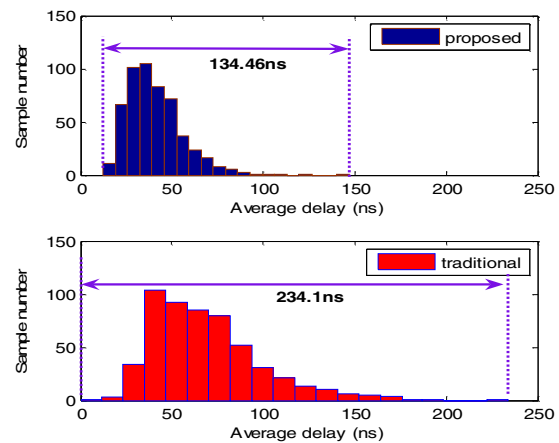


Figure 8. Buffer sensitivity to VTH variation.

From Fig. 8 we can see that with the same process variation, the average delay variation of the traditional buffer can be as large as 234.1 ns while for the proposed buffer, the variation is only 134.46 ns. Therefore, the boost technique can effectively suppress the process variation effect.

We then assume that the supply voltage has a 3σ 5% variation in 0.4 V case. The load capacitance is set to 20 fF. Again we perform 1000 Monte Carlo analysis and the result is shown in Fig. 9. The standard deviation of traditional buffer is 8.32 ns while for the proposed buffer, it is only 4.91 ns. The average delays of the traditional buffer and new buffer are 63.08 ns and 37.74 ns respectively.

We further analyze the temperature variation effect on the circuit performance. We assume the normal temperature is 27°C and has a 3σ variation of 15

°C. We perform 1000 Monte Carlo analysis and present the result in Fig. 10.

We can find the standard deviation of the traditional buffer is 7.61 ns while the standard deviation of the new buffer is 4.07 ns. The average delays of the traditional buffer and new buffer are 60.06 ns and 36.13 ns respectively.

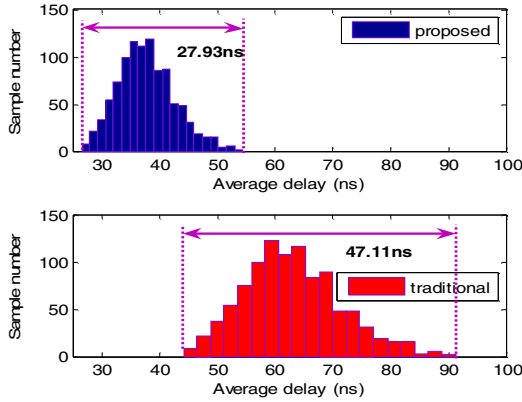


Figure 9. Buffer sensitivity to supply voltage.

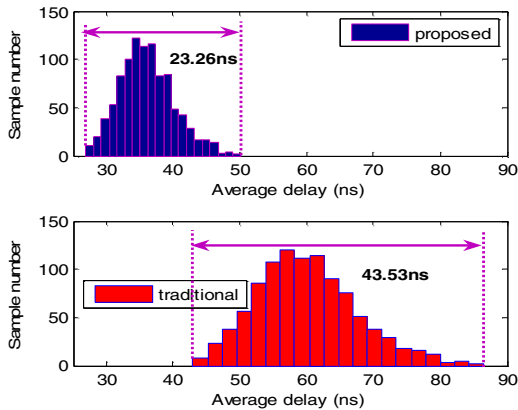


Figure 10. Buffer sensitivity to temperature.

3.3. Buffer Usage in Clock Distribution

Fig. 11 shows a typical H-tree used for clock distribution. Here the load is simplified as four parallel connected inverters. The interconnect length is shown in Fig. 11. We use the PTM model to calculate the interconnect parameters [9]. The r, c parameters for the global interconnect line are $22 \Omega/\text{mm}$, and $243.768 \text{ fF}/\text{mm}$. The l parameters for 1 mm, 500 μm line, and 200 μm line are $1.476 \text{ nH}/\text{mm}$, $1.338 \text{ nH}/\text{mm}$, and $1.155 \text{ nH}/\text{mm}$.

Fig. 12 shows the T1 node waveform comparison between the proposed buffer network and the traditional buffer network when the input clock frequency is 1 MHz.

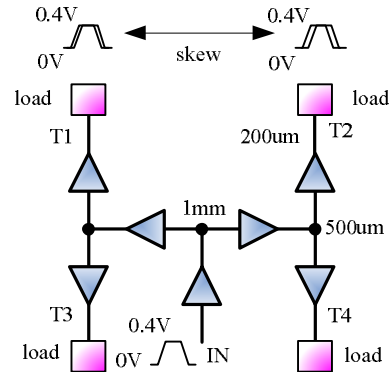


Figure 11. A simple H-tree clock network.

From Fig. 12, we can observe that the rise delay of the traditional buffer is very large compared to that of the new buffer. However, the fall delay is not increased as fast as the rise delay thereby decreasing the pulse width. In order to compensate this effect, we have to size up the circuit. Simulation results show that when the traditional buffer is sized up to 5.4X, or when the channel area is $2.1312 \mu\text{m}^2$, the delay is equal to the proposed new one.

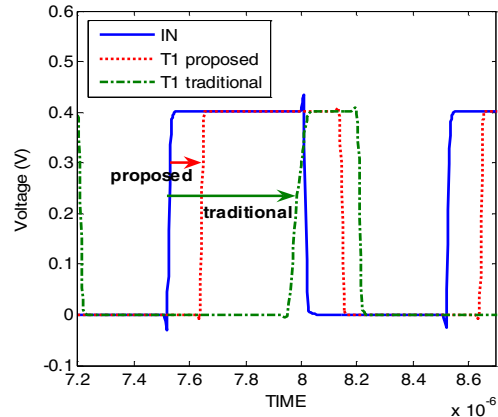


Figure 12. Clock waveform comparison between two buffers.

In addition to the speed benefits, we also compare the skew variation between node T1 and node T2. We assume NMOS transistor V_{TH0} has an inter-die variation of 10%, PMOS transistor V_{TH0} have a 10% inter-die variation and 5% intra-die variation. We then assume that the supply voltage has a 3σ 5% variation. At last we assume the normal temperature is 27°C and has a 3σ variation of 5°C . Under this combined PVT variation condition, we perform 500 Monte Carlo simulations and give the result in Fig. 13.

In Fig. 13, the standard deviation of the skew of the traditional buffer network is 44.67 ns while for the new buffer network, it is only 7.86 ns. The average skew of the traditional buffer network is 2.08 ns while for the new buffer network it is -0.85 ns. The skew change due

to PVT variation is reduced by 5.5X from 336.3 ns to 61.2ns. Even when the traditional buffer is sized up to 5.4X, the skew is still 36.81% larger than the new one.

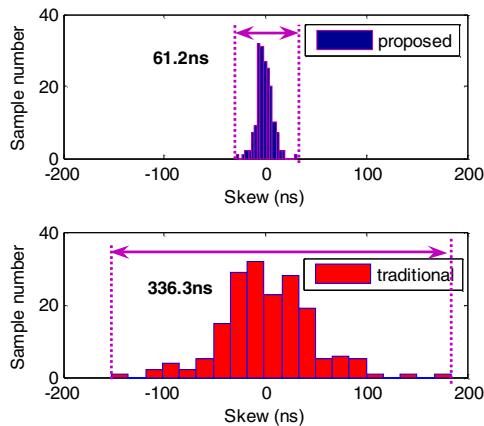


Figure 13. Skew variation with different buffers.
3.4. Power Analysis of the New Buffer

At the last of the paper, we compare the power consumption of the two buffers. As shown in Fig. 14, the proposed buffer consumes more power than the traditional buffer. It is reasonable because from Fig. 4 we can see that there are 14 devices, more than that of the traditional buffer which is 4. From this figure, we can see that as the supply voltage increases above the threshold voltage, the power increases very fast, from 47% in the sub-threshold region to 70% when VDD is 0.9 V. However, considering the delay improvement and the PVT variation tolerance enhancement, which are the main considerations in the sub-threshold circuit design and the future nanometer regime circuit design, the proposed boosted buffer is still very promising. Furthermore, since the delay of the traditional buffer is much smaller, if given a time spec, the total buffer number of the traditional buffer network is larger than the proposed buffer network. As a result, the total power and area consumed by the new buffer network are still smaller similar to [8].

4. Conclusion

In this paper, we propose a new capacitive boosted buffer technique for high speed process variation tolerant interconnect. It is compatible with digital CMOS technology and can be operated in both sub-threshold region and above threshold region without

the problem of short current. Simulation results show that compared to the traditional buffer, the new buffer is more robust to load variation and PVT variation. When applied to an H-tree clock distribution network, the skew of the new buffer network is reduced by 5.5X compared to the traditional buffer network.

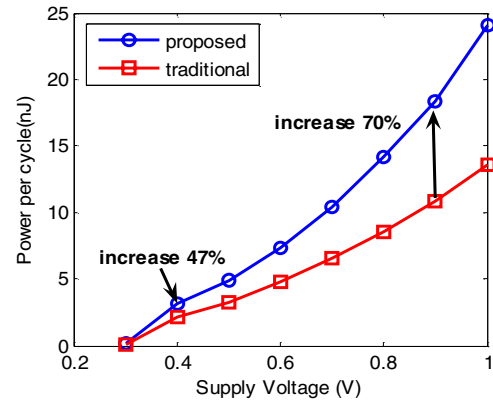


Figure 14. Power analysis of the two buffers.

5. References

- [1] B. H. Calhoun, et al, "Ultra-dynamic voltage scaling (UDVS) using sub-threshold operation and local voltage dithering", *JSSC*, 41 (1): 238-245, 2006.
- [2] B. H. Calhoun, et al, "A leakage reduction methodology for distributed MTCMOS", *JSSC*, 39: 818-826, 2004.
- [3] A. Srivastava, et al, "Power minimization using simultaneous gate sizing, dual-Vdd and dual-Vth assignment," in DAC, pp.783-787, 2004.
- [4] Choi, Kyu-Won, et al, "Optimal zigzag (OZ): An effective yet feasible power-gating scheme achieving two orders of magnitude lower standby leakage" in Symposium on VLSI, 2005, pp. 312-315.
- [5] Dinesh Patil, et al, "A new method for design of robust digital circuits," in ISQED, 2005, pp. 676-681.
- [6] C. Neau, et al, "Optimal body bias selection for leakage improvement and process compensation over different technology generations," in ISLPED, 2003, pp. 116-121.
- [7] Jonggab Kil, et al, "A high-speed variation-tolerant interconnect technique for sub-threshold circuits using capacitive boosting," in ISLPED, 2006, pp. 67-72.
- [8] Jae-sun Seo, et al, "Self-timed regenerators for high-speed and low-power interconnect" in ISQED, 2007, pp. 621-626.
- [9] PTM model, <http://www.eas.asu.edu/~ptm/>.