Fan-in Sensitive Low Power Dynamic Circuits Performance Statistical Characterization

Jinhui Wang[1], Na Gong[2], Wuchen Wu[1], Ligang Hou[1]
[1]VLSI and System Lab, Beijing University of Technology, Beijing, China
[2] Department of Computer Science and Engineering, SUNY at Buffalo, Buffalo, NY, USA
wangjinhui888@emails.bjut.edu.cn

ABSTRACT

Using neural networks, a highly reliable and precise system to display low power dynamic circuit performance statistical characterization is proposed in this paper. The proposed model successfully estimates the nonlinear changing of the leakage power, the active power and the delay of the different fan-in low power dynamic gates with the dual threshold voltage technique, the multiple-supply technique and the sleep transistor technique. At last, the precision priority of estimating system is obtained.

I. INTRODUCTION

Due of the superior speed and area characteristics, the dynamic gates (DG) have been extensively applied in critical part of modern high performance microprocessor and memory [1]-[4]. However, as technology aggressive downscales and clock frequency increases, DG typically consumes higher active and leakage power, which even degrades their performance. Therefore, the low power and error free operation of DG has become a crucial issue in the current CMOS technologies, especially as the growing demand for the portable and battery operated systems such as cell phones and laptop [5].

A number of approaches have been proposed to reduce power of DG, among them, the dual threshold voltage (V_t) technique (DVT) [6], the multiple-supply technique (MST) [7] and the sleep transistor technique (STT) [8] are three of the most popular approaches. However, although these techniques reduce the power consumption, they degrade speed of the circuits more or less. Hence, quickly and precisely estimating the power and delay changing of DG with DVT, MST and STT would help designers judge whether optimized DG meet the design constrains of both the power and the delay, and therefore save a huge amount of time and reduce iteration greatly.

Many methods have been proposed to realize this estimation, include the circuit simulation [9], [10], the probabilistic technique [11], [12], the macro-modeling approach [13] and the Monte Carlo analysis [14], [15]. However, although existing work has successfully implement estimation, there are some insufficiencies. Exhaustive simulation would cost much more time, sometimes up to several days. Probabilistic techniques are weakly pattern dependent. And they are fast and tractable but typically involve assumptions about joint distribution. Macro-modeling approach needs lower and upper bounds. Monte Carlo analysis is confidence interval based and as a result, it requires more iteration to converge [16].

As no explicit models are required, quick convergence and effective classification [17], [18], the neural networks have emerged to avoid insufficiency of above methods and provide a very appealing approach to estimate nonlinear changing of power and delay of DG. Especially, the wavelet neural networks (WNN) has been widely known and successfully used in most estimation systems. In this paper, based on WNN, a novel estimating system for forecasting power and delay of DG with DVT, MST and STT is proposed. And its accuracy is validated with the simulation test.

II. LOW POWER TECHNIQUES

A. DVT in DG

The DVT is realized through employing different V_t transistors in different paths of circuits. As shown in fig. 1 (b), the critical signal transitions determining the delay of DG occurring along the evaluation path. In DG with DVT, therefore, all of the transistors, activated during the evaluation phase, have a low V_t. Alternatively, the precharge phase transitions are not critical for the performance of DG, they have high V_t [6].

978-1-4244-6683-2/10/$26.00 ©2010 IEEE

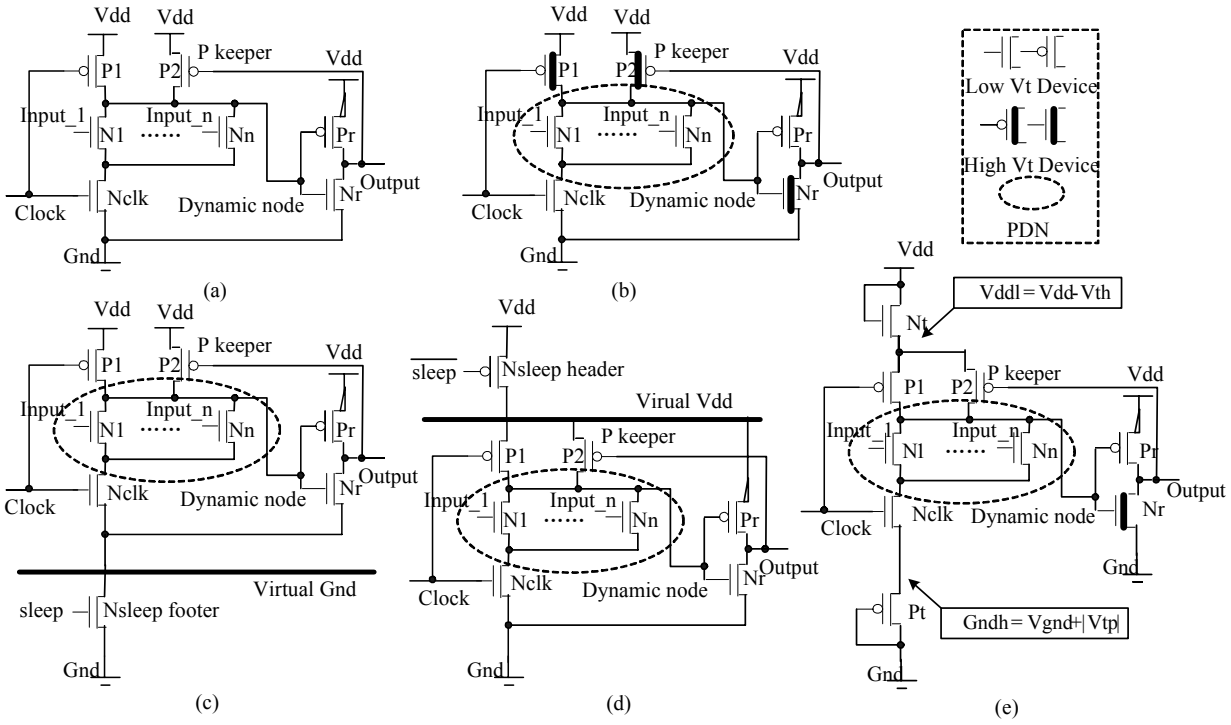

Fig. 1 (a) standard DG (b) DG with DVT (c) DG with footer (d) DG with header (e) DG with STT

B. STT in DG

Fig.1 (c) and (d) displays STT based on a current switch, which can be either a nMOSFET footer or a pMOSFET header. When DG is active (the footer and header are both turned on, sleep=1), the voltage of virtual ground/V_{dd} is close to ground/V_{dd}, which is determined by the size of the footer/header and the current that flows into the footer/header. Once setting idle (both the footer and the header are turned off, sleep=0), virtual ground/V_{dd} slowly goes up/down until it reaches a steady state potential, which is close to V_{dd}/ground. The steady state potential and the time that footer/header takes to reach are determined by the amount of current flows through the logic block and the footer/header. Thus, in idle mode, the leakage current path of DG is cut off by the sleep transistor, and the leakage power is suppressed greatly [8].

C. MST in DG

MST is proposed in [7], as shown in fig.1 (e). In DG with MST, lower supply (V_{ddl}) and higher ground (Gndh) are applied to reduce the supply swing (V_{swing}) from V_{dd}-Gnd to V_{ddl}-Gndh (Gndh $=V_{gnd}+|V_{tp}|$, V_{ddl} $=V_{dd}-V_{tn}$, V_{tn} and V_{tp} are the threshold voltages of Nt and Pt in fig.1 (e), respectively). In this paper, both V_{tp} and V_{tn} are considered as optimization parameters to minimize the power consumption ($|V_{tp}|$ = V_{tn} = 0.1V).Therefore, V_{ddl} = 0.7V, Gndh = 0.1V.

However, both PMOS and NMOS transistors in output inverter are simultaneously turned on and produce significant short circuit current which increases leakage power consumption. Furthermore, the delay is increased due to significantly degraded gate overdrive of output inverter. In order to suppress this short-circuit current, the high V_t transistor is employed in the output inverter.

D. Fan-in effect on power and delay

In a DG with DVT, MST and STT, the number of fan-in has a crucial impact on its power and speed characteristics. On the one hand, most of DG power is produced by the transistors in pull-down network (PDN) and the number of these transistors greatly influences the effectiveness of DTV, MST and STT on suppressing power. On the other hand, with the increasing fan-in, the more current paths are constructed from dynamic node to ground. These additional current paths decrease the delay of the evaluation stage and compensate the speed loss.

III. ESTIMATING SYSTEM BASED ON WNN

WNN is multi-layer feedback architecture with wavelet, allowing the minimum time to converge to its global maximum [19]. The WNN employs a wavelet base rather than a sigmoid function, which discriminates it from general back propagation neural networks.

The function of mapping can be expressed as:

$$f(x)=\sum_{o=1}^{h}\sum_{j=1}^{m}\omega_o\frac{1}{\sqrt{|a_o|}}\psi_{jo}\left(\frac{\sum_{i=1}^{n}x_{ij}-b_o}{a_o}\right) \quad (1)$$

where ω_0 (0=1, 2…, h) and ψ_{j0} are output of hidden layer neurons and the wavelet bases, respectively. Networks have three parameters to be trained: output weight ω, translation factors a and dilation factors b.

In this paper, a set of training samples with labels *D= {(y$_i$, x$_i$) / i=1,2…,N}.* And the Morlet wavelet is used as stimulation function of hidden layer

$$\psi(x)=\cos(1.75x)e^{-x^2/2}$$

The error performance function is given by:

$$J=\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{m}(y_{j,i}^{d}-y_{j,i})^2 \quad (2)$$

where N is the total number of training patterns, and $y_{j,i}^{d}$ and $y_{j,i}$ are the desired and real outputs, respectively. The training process of WNN is performed as following:

1) Create initial population of individuals according to the initiation strategy-output weight ω, translation factors a and dilation factors b are in (0, 1).

2) Calculate the fitness function by (2).

3) To minimize the fitness function in (2), the weights and coefficients *a* and *b* can be updated using the following formulas

$$\begin{aligned}\omega_j(k)&=\omega_j(k-1)+\eta(y^d(k)-y(k))h_j\\&+\alpha(\omega_j(k-1)-\omega_j(k-2))\end{aligned} \quad (3)$$

$$\Delta a_j=a_j(y^d(k)-y(k))\omega_j h_j\frac{(\sum_{i=1}^{n}x_{ij}-b_j)^2}{a_j^{3}} \quad (4)$$

$$\begin{aligned}a_j(k)&=a_j(k-1)+\eta\Delta a_j\\&+\alpha(a_j(k-1)-a_j(k-2))\end{aligned} \quad (5)$$

$$\Delta b_j(k)=(y^d(k)-y(k))\omega_j\frac{\sum_{i=1}^{n}x_{ij}-b_j}{a_j^{2}} \quad (6)$$

$$\begin{aligned}b_j(k)&=b_j(k-1)+\eta\Delta b_j\\&+\alpha(b_j(k-1)-b_j(k-2))\end{aligned} \quad (7)$$

$$h_j=\frac{1}{\sqrt{|a_j|}}\psi_j\left(\frac{\sum_{i=1}^{n}x_{ij}-b_j}{a_j}\right)\ (j=1,2,\cdots,m) \quad (8)$$

Where η*=0.01*, α *=0.05. j* is the number of hidden layer neurons.

4) Repeat step 2) to step 3) until some constraint condition is satisfied, then stop and the desired individuals are obtained.

IV. RESULTS AND ANALYSIS

As described in section II, the input of the WNN is the number of fan-in of DG with DTV, MST and STT and the outputs are power reduction percentage and delay increase percentage as compared to standard DG. The training data and testing data are collected from the simulation results with HSPICE tools based on 45nm BSIM4 model [20]. Device parameters are shown in Table 2. Each gate drives a capacitive load of 8fF and is turned to operate at 1GHz clock frequency. When simulating the leakage power, all of DG are set in CHIH (clock=1, In_1= In_2=… In_n =1) state, which can ensure every gates in the lowest leakage state [21]. In order to test the availability of the model, the testing data come from 2, 4, 6, 8, 16, 32, 40, 48-inputs dynamic OR and MUX gates, because these typical gates are usually utilized in practice. The network was trained within 2500 learning iterations. The largest error E or given precision is 0.0001.

As can be seen from table 3, the leakage and active power reduction and delay increase percentage of DG with DVT ranges about from 2%-18%, 2%-13% and 29%-59% as compared to standard DG. This is because the active power of DG consist of leakage and switching power, and the leakage power decreases with the increasing of V_t and it can be expressed as followed

Table 1 Testing Errors (%)
(LR: Leakage power reduction; AR: Active power reduction; DI: delay increase)

Technique		Fan-in	2	4	8	16	32	40	48	Ave.
DVT	OR	LR	2.7000	1.1597	1.0799	0.7202	0.0266	0.8043	0.0500	1.0929
		AR	4.8800	2.2112	4.6943	0.8647	0.0482	4.8988	2.5100	2.8298
		DI	2.9700	8.2315	7.2349	0.8163	1.9014	2.5517	0.6900	4.5625
	MUX	LR	3.5480	0.8246	0.2776	0.8841	0.0296	1.1080	0.0520	1.0709
		AR	8.2160	3.1271	4.7540	3.7221	0.1339	7.0133	0.4890	4.2991
		DI	9.6940	1.0243	10.6633	0.4807	0.4265	5.8971	2.2080	4.3158
MST	OR	LR	0.3440	0.3042	0.1075	0.2945	0.2079	1.4729	0.0940	0.2554
		AR	5.2570	2.2082	0.5316	3.5276	0.5112	2.9843	3.9780	2.2191
		DI	5.7050	2.2749	2.9334	4.3154	1.8777	3.5023	3.6180	3.4652
	MUX	LR	0.6200	0.4956	0.3822	0.2708	0.1834	1.5259	0.0790	0.3272
		AR	0.1350	0.3921	0.9525	6.4847	3.4804	1.9057	0.6460	1.6381
		DI	3.6880	2.8448	1.8892	10.6309	1.5223	4.0556	0.7470	3.1827
STT	Footer OR	LR	0.1064	0.0608	0.9805	0.0796	0.0586	0.0030	0.7121	3.6375
		AR	9.5449	4.6514	0.1494	0.5800	0.3004	2.4330	3.2894	3.1483
		DI	2.1044	7.4196	1.4263	2.3059	0.3079	0.3890	3.1988	1.4078
	Footer MUX	LR	0.1390	0.0589	0.8953	1.8169	1.6270	0.0010	2.4089	0.9924
		AR	12.5010	1.4655	9.9338	1.9495	2.5194	4.8380	4.9590	5.4523
		DI	1.6220	0.5052	9.2605	2.4229	2.6869	2.9727	2.8536	1.7450
	Header OR	LR	1.1980	1.1607	3.8974	0.2416	0.3288	0.7410	1.2469	0.3900
		AR	10.1170	1.2211	5.9616	0.7836	1.1507	11.4632	4.5598	4.4510
		DI	14.2549	3.5362	1.0832	2.7835	3.2329	0.7951	4.3174	5.3410
	Header MUX	LR	2.1130	1.6970	1.4918	2.1581	0.0187	5.7683	2.1348	0.3440
		AR	1.1840	4.4402	1.5804	4.2555	3.0187	12.5128	4.4903	2.3150
		DI	8.5020	3.8495	1.9221	1.0447	3.7049	1.9394	3.5446	1.5560

Table 2 Parameters of devices

Supply		Gnd voltage	
V_{dd}	V_{ddl}	*Gnd*	*Gndh*
0.8V	0.7V	0V	0.1V
Low-V_t NMOS		*Low-V_t PMOS*	
0.22V		-0.22V	
High-V_t NMOS		*High -V_t PMOS*	
0.35V		-0.35V	

Table 3 Power and Delay change (%)

Tec.		Range	Tec.		Range
DVT	OR LR	2-18	STT	Footer OR LR	98-99
	OR AR	8-10		Footer OR AR	0-10
	OR DI	32-58		Footer OR DI	0-3
	MUX LR	2-15		Footer MUX LR	20-40
	MUX AR	2-13		Footer MUX AR	0-10
	MUX DI	29-56		Footer MUX DI	38-68
MST	OR LR	49-55		Header OR LR	98-99
	OR AR	15-32		Header OR AR	0-25
	OR DI	8-20		Header OR DI	0-5
	MUX LR	48-54		Header MUX LR	15-37
	MUX AR	20-31		Header MUX AR	5-23
	MUX DI	4-15		Header MUX DI	17-48

$$I_{sub} = \frac{W_{eff}}{L_{eff}} u \cdot \sqrt{\frac{q\varepsilon_{si}N_{ch}}{2\Phi_s}V_T^2 \exp\left(\frac{V_{gs}-V_t}{nV_T}\right)} \cdot \left(1-\exp\left(-\frac{V_{ds}}{V_T}\right)\right) \quad (9)$$

where I_{sub}, L_{eff} and W_{eff} are the sub-threshold leakage current, the effective channel length and width respectively, and other parameters have their usual meanings.

However, high V_t transistors in circuits will degrade the speed (v) of DG, which are given as

$$v \propto \frac{V_{dd}^{0.3}\left(1-\frac{V_t}{V_{dd}}\right)^{1.3}}{t_{ox}^{0.5}} \quad (10)$$

As can be seen from table 3, the leakage and active power reduction and delay increase percentage of DG with MST ranges about from 48%-55%, 15%-32% and 4%-20% as compared to standard DG. And in DG with MST, the total power P_{total} as expressed

$$\begin{aligned} P_{total} &= P_{leak} + P_{switching} \\ &= \alpha f C_L (V_{ddl} - Gndh)^2 + I_{leak} V_{ddl} \\ &= \alpha f C_L V_{switching} + I_{leak} V_{ddl} \end{aligned} \quad (11)$$

where α, f, I_{leak}, $P_{switching}$, P_{leak} C_L and V_{swing} are the switching activity factor, the clock frequency, the leakage current, the switching power, the leakage power, the capacitive load at the dynamic node, and the dynamic node voltage swing (V_{swing} = V_{ddl}-Gndh), respectively. The switching power consumed for charging/discharging the dynamic node has a quadratic dependence on the dynamic node voltage swing. The leakage power has a linear dependence on the dynamic node voltage swing. Hence, the power saving of DG with MST is significantly higher as compared to DG with the signal supply. But, as shown in (10), the speed (v) is also reduced with lower V_{dd}.

When the STT is applied in DG, the power characteristic is complex, as can be seen from table 3. Leakage power of DG with footer and header transistor is decreased up to 98% as compared to that of standard DG. What's more, the leakage power saving offset active power increase that is introduced by the additional parasitic capacitance of the sleep transistor. Therefore, the active power of dynamic power is reduced 0-25%. Due of simple structures of PDN of OR gates, the number of fin-in have little effect to speed. But as to MUX gates, the delay is determined by the stack effect and the relative contribution of the capacitance between the capacitance of P1+P2 (in fig.1 (c), (d)) and in PDN. When fan-in is 40, in DG with footer, the capacitance of P1+P2 mismatches with that of PDN+N_{clock}+N_{sleep}; in DG with header, the capaci-

tance of P1+P2+N_{sleep} mismatches with that of PDN+N_{clock}, which decelerate the circuits greatly, the maximum delay penalty is up to 68%, as shown in Table 3.

The testing errors are listed in table 1. And the power and delay estimating errors both are less than 15%, and average errors less than 5%. Obviously, the estimating system possesses the high estimation accuracy. Therefore, this system based on wavelet neural network has strong stability and it can be embedded in EDA tools as power and delay estimation program module. What's more, it also can be seen from table 1, in DG with DVT, STT and MST, the average estimating errors of the leakage power decrease are smaller than that of the active power decrease and the delay increase. As for DG with DVT and MST, the average estimating errors of the active power decrease are smaller than that of the delay increase. As for DG with STT, the average estimating errors of delay increase are smaller than that of the active power decrease. Thus, the precision priority of estimating system is as followed

DVT: leakage power > active power >delay
MST: leakage power > active power >delay
STT: leakage power > delay > active power

V. CONCLUSION

A fast and precise system based on wavelet neural networks for estimation is proposed to help designers judge whether the application of DVT, MST and MST in DG meet the design constrains of the power and the delay, thereby saving considerable time and energy. The simulation results for verification indicate that the system can be well applied in VLSI design process with accuracy ratio of more than 85%. And three techniques realize the power optimization with different optimized efficiencies and speed loss. At last, the precision priority of estimating system is discussed

REFERENCES

1. H. Kaul, et al, "A 320 mV 56 μW411 GOPS/Watt Ultra-Low Voltage Motion Estimation Accelerator in 65 nm CMOS" *IEEE Journal of Solid-State Circuits,* Vol. 44, pp.107-114, Feb. 2009.
2. B.Stackhouse, et al, "A 65 nm 2-Billion Transistor Quad-Core Itanium Processor" *IEEE Journal of Solid-State Circuits,* Vol. 44, pp.18-31, Jan. 2009.
3. R. Ratnayake, A. Kavcic, and G. Wei, "A High-Throughput Maximum a Posteriori Probability Detector", *IEEE Journal of Solid-State Circuits*, Vol. 43, pp.1846-1858, Aug. 2008.
4. S. B. Wijeratne, N. Siddaiah, et al, "A 9-GHz 65-nm Intel® Pentium 4 Processor Integer Execution Unit" *IEEE Journal of Solid-State Circuits*, Vol. 42, pp.26-37, 2007.
5. N. Gong, et al, 'Analysis and Optimization of Leakage Current Characteristics in Sub-65nm Dual Vt Footed Domino Circuits', *Microelectronics Journal*, Vol. 39, pp.1149-1155, Sep. 2008.
6. V. Kursun and E. G. Friedman, "Domino logic with variable threshold voltage keeper" *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Vol. 11, pp.1080-1093, Jun. 2003.
7. H. Kim and Y. Shin, "Analysis and Optimization of Gate Leakage Current of Power Gating Circuits", Asia and South Pacific Conference on Design Automation, pp.565-569, 2006.
8. R. Mader and I. Kourtev, "Reduced Dynamic Swing Domino Logic," in Proceedings of the ACM/SIGDA Great Lakes Symposium on VLSI, pp.33-35, 2003.
9. X. Li, et al, "A new SPICE reliability simulation method for deep submicrometer CMOS VLSI circuits" *IEEE Transactions on Very large Scale Integration (VLSI) Systems*, Vol. 6, pp.247-256, Jan. 2006.
10. L.W. Nagel, "SPICE2: A Computer Program to Simulate Semiconductor Circuits", ERL Memo #ERL-M520, Electronics Research Laboratory, University of Califomia, Bcrkeley, CA, May, 1975.
11. F. N. Najm, "Transition density: A new measure of activity in digital circuits", *IEEE Transactions on Computer-Aided Design,* Vol.12, pp.310-323, Jun. 1993.
12. S. Bhanja and N. Ranganathan, "Dependency preserving probabilistic modeling of switching activity using Bayesian networks" Design Automation Conference, pp.209-214, 2001
13. S. Gupta and F. N. Najm, "Power modeling for high-level estimation", *IEEE Transactions on Very large Scale Integration (VLSI) Systems*, Vol. 8, pp.18-29, Jun. 2000.
14. R. Burch, F. N. Najm, P. Yang, and T, "Trick, A Monte-Carlo approach for power estimation", *IEEE Transactions on Very large Scale Integration (VLSI) Systems,* Vol. 1, pp.63-71, Apr. 1993.
15. R. Burch, F. N. Najm, P. Yang, and T. Trick, "McPOWER: A Monte Carlo approach to power estimation" International Conference on Computer-Aided Design, pp.90-97, 1992.
16. A. K. Murugavel, N. Ranganathan, R. Chandramouli, and S. Chavali, "Least-Square Estimation of Average Power in Digital CMOS Circuits", *IEEE Transactions on Very Large Scale Integration (VLSI) Systems,* Vol. 10, pp.55-58, Feb. 2002.
17. F. Aminian, M. Aminian, and H. W. Jr. Collins, Analog Fault Diagnosis of Actual Circuits using Neural Networks, *IEEE Transaction on Instrumentation and Measurement*, Vol. 51, pp.544-550,
18. A. Oonsivilai and M. E. El-Hawary, "Wavelet neural network based short term load forecasting of electric power system commercial load," in Proceeding of 1999 IEEE Canadian Conference on Electrical and Computer Engineering, New York: IEEE, 1999, 1223-1228.
19. Q. H. Zhang and A. Benveniste, "Wavelet Network". *IEEE Transaction on Neural Network*, Vol. 3, pp.889-898, Jun, 1992.
20. Predictive Technology Model(PTM), http://www.eas.asu.edu/~ptm
21. J. T. Kao and A. P. Chandrakasan "Dual-threshold voltage techniques for low-power digital circuits" *IEEE Journal of solid-state Circuits*, Vol. 35, pp.1009-1018, Jul. 2000.