TU Delft

Delft University of Technology

System-level sub-20 nm planar and FinFET CMOS delay modelling for supply and threshold voltage scaling under process variation

Majzoub, Sohaib; Taouil, Mottaqiallah; Hamdioui, Said

DOI
10.1166/jolpe.2019.1590

Publication date
2019

Document Version
Final published version

Published in
Journal of Low Power Electronics

Citation (APA)
Majzoub, S., Taouil, M., & Hamdioui, S. (2019). System-level sub-20 nm planar and FinFET CMOS delay modelling for supply and threshold voltage scaling under process variation. *Journal of Low Power Electronics*, *15*(1), 1-10. https://doi.org/10.1166/jolpe.2019.1590

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.
For technical reasons the number of authors shown on this cover page is limited to a maximum of 10.

Copyright © 2019 American Scientific Publishers
All rights reserved
Printed in the United States of America

Journal of
Low Power Electronics
Vol. 15, 1–10, 2019

System-Level Sub-20 nm Planar and FinFET CMOS Delay Modelling for Supply and Threshold Voltage Scaling Under Process Variation

Sohaib Majzoub[1,*], Mottaqiallah Taouil[2], and Said Hamdioui[2]

[1] *Department of Electrical and Computer Engineering, University of Sharjah, Sharjah, 27272, UAE*
[2] *Laboratory of Computer Engineering, Delft University of Technology, 2628, The Netherlands*

(Received: 4 November 2018; Accepted: 7 January 2019)

Standard low power design utilizes a variety of approaches for supply and threshold control to reduce dynamic and idle power. At a very early stage of the design cycle, the V_{dd} and V_{th} values are estimated, based on the power budget, and then used to scale the delay and estimate the design performance. Furthermore, process variation in sub-20 nm feature technologies introduces a substantial impact on speed and power. Thus, the impact of such variation on the scaled delay has to also be considered in the performance estimation. In this paper, we propose a system-level model to estimate this delay, taking into consideration voltage scaling under within-die process variation for both planar and FinFET CMOS transistors in the sub-20 nm regime. The model is simple, has acceptable accuracy and is particularly useful for architectural-level simulations for low-power design exploration at an early stage in the design space exploration. The proposed model estimates the delay in different supply voltage and threshold voltage ranges. The model uses a modified alpha-power equation to measure the delay of the critical path of a computational logic core. The targeted technology nodes are 14 nm, 10 nm, and 7 nm for FinFETs, and 22 nm, and 16 nm for planar CMOS. Within-die process variation is assumed to be lumped in with the threshold voltage and the transistor channel length and width to simplify its impact on delay. For the given technology nodes, the average percentage error numbers of theproposed delay equation compared to hSpice are between 0.5% to 14%.

Keywords: Low-Power Design, System-Level Modelling, Planar CMOS, FinFET, Process Variation, Within-Die Variation, Alpha-Power Model, Multi-V_{dd}, Voltage Scaling.

1. INTRODUCTION

Over the last decade, the main advantage of technology scaling was thought to be its ability to increase transistor density. The other traditional gains, such as increasing the switching speed and lowering the supply voltage to improve power consumption, are no longer sustainable.[1–4] Beyond 32 nm, the conventional planar CMOS transistors also suffered from high variability and performance degradation.[5] During the process of attempting to improve transistor performance, the double-gated transistors showed good potential towards improving the switching strength and hence the performance of the transistor.[6–11] Despite this improvement, the new structure introduced different types of dimensional variability. Thus, process variation continued to be one of the main challenges for design reliability. At the system level, process variation results in an asymmetric speed distribution among cores in today's multicore and manycore processors.[5]

Given today's chip sizes, billions of transistors are fabricated on the same chip to make thousands of processing and memory cores. Running such chips is only possible if the power consumption is within budget.[12] Initially, a system-level power optimization simulation environment is utilized to analyze different optimization techniques using estimated power and performance values as figures of merit. In such cases, this analysis is considered to be the initial stage in the design process. Such a simulation can be realized only with an effective delay model.[13–15] Thus, a simple and relatively accurate system-level delay model which can be used in a multi supply voltage design

*Author to whom correspondence should be addressed.
Email: smajzoub@sharjah.ac.ae

optimization is needed. This delay model can be used to estimate the core performance under different simulation conditions.

Within-die variation comes as a result of the CMOS manufacturing process. Its impact on the transistor's features, which comes as systematic and random effects, is usually modelled using statistical methods.[16,17] Such impact affects the transistor's threshold voltage and its channel's length and width. Changes in the mentioned transistor's parameters impact the transistor's switching delay. Thus, these changes in the delay have to be estimated while taking such variability into consideration. Thus, a delay equation that accurately measures the performance of a core in a multi-core processor, while considering the impact of such variations, is needed.

This paper is organized as follows: Section 2 explains the motivation behind the proposed work. Section 3 covers previous work. Section 4 discusses the proposed delay model. Finally, Section 5 presents the experimental setup and then the analysis of the results.

2. MOTIVATION

System-level simulation is a common practice in VLSI design.[13,18,19] The main challenge is that system-level simulation can be very complex. Analysis must be carried out from the application perspective while considering system, circuit, and device level issues.[20,21] For instance, a typical application running on a manycore platform is expected to include thousands of tasks with communication load between dependent tasks. Such a platform has a heterogeneous set of cores which are carefully located in order to reduce traffic and provide locality for the resources needed by a given application. Implementing a low power processor with a multi-V_{dd}/multi-frequency configuration also requires a proper distribution of supply voltages and frequencies that are based on the workload to reach minimum power consumption.[23] Moreover, process, voltage and temperature (PVT) variations can add another level of complexity to the design problem. Process variation is very important, especially when a platform has thousands of cores, which can only be realized at smaller process technologies. A simulation environment that includes all of these factors is used to design low-power and variation-aware processors. Modeling and estimating the delay in such a complex simulation environment is critical. The delay model should be accurate enough to produce realistic numbers and simple enough to keep simulation time reasonably fast.[24–26]

The alpha-power delay model is one of the most popular delay equations due to its simplicity and relative accuracy. It is used at the system-level to estimate the delay under different supply voltage or threshold voltage conditions.[19,27,28] The model is used in the literature for estimating the delay when evaluating different network-on-chip policies or routing algorithms,[13,18,21] multicore and manycore low power optimizations,[14,29,30] or low power memory designs at the system-level.[19] Furthermore, the model is also used to evaluate the delay in near-threshold operations.[31] It is also used to estimate the impact of process variation on the system-level delay.[32–36] Usually, the critical path delay of a core block is modelled as an inverter chain. This is a common practice at the early stage of the design cycle.[20,31–33,37–39] At this point, the design process is focused on a specific goal using different optimization methods. Thus, the alpha-power model is sufficient to study different optimization options.

In this work, we attempt to exploit the simplicity of the alpha-power model while improving its accuracy. Such a conventional delay model might not be applicable in its simple form for such a complex system-level simulation environment. As a demonstration of the problem, consider the threshold voltage (V_{th}), which is a very important parameter to accurately measure the delay. Usually, the threshold voltage is assumed to be a constant number produced using statistical methods that capture process variation impact. However, the threshold voltage is always affected by other transistor parameters. Thus, the effective threshold voltage is a result of the target V_{th}, V_{dd}, along with the L and W. Figure 1 shows the effective V_{th} considering the previously mentioned parameters. Each of the planes shows the effective V_{th} value compared to the assumed constant V_{th} (labeled with an arrow pointing to the plane). This figure shows how the effective V_{th} changes with respect to other parameters (W considered to be nominal for the given plots for simplicity). These numbers are produced for 7 nm technology using hSpice.

In this paper, we propose a modified alpha-power model used to estimate the delay at the system-level for FinFET and planer CMOS transistors for multi-V_{dd} designs under process variation. We did extensive hSpice simulations to

IP: 131.180.130.242 On: Tue, 01 Oct 2019 09:17:06
Copyright: American Scientific Publishers
Delivered by ingenta

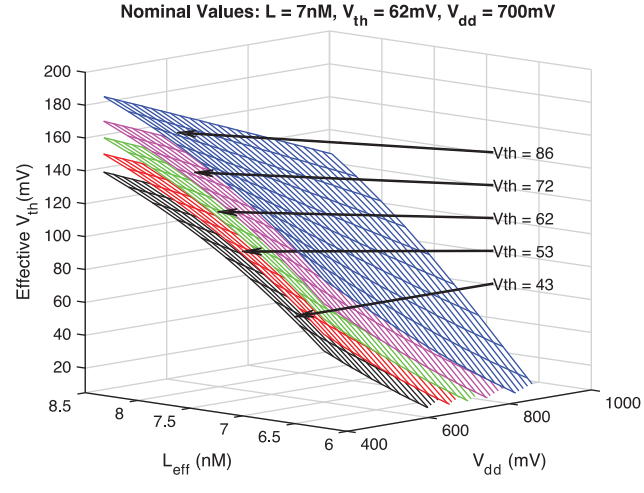


Fig. 1. Effective V_{th} under different conditions for the 7 nm technology.

cover different process variation ranges for different supply voltage regimes. Our proposed delay model brings the hSpice accuracy, i.e., circuit level accuracy, up to the system level. Finally, we compare the delay numbers produced by the proposed delay model to the hSpice delay numbers.

3. RELATED WORK

Many papers have addressed delay and power modelling of FinFET transistors under process variations. Most of the papers addressed this type of variation and its impact on the transistor parameters measure in statistical forms, such as mean and standard deviation for each of these parameters. Stillmaker et al. presented in Ref. [4] scaling equations for all CMOS technologies starting from 180 nm down to 7 nm. They did extensive spice simulations and produced equations to calculate the delay of any node based on the delay of the previous node using their scaling equation. They also produced energy and power factors as well. Khatamifard et al.[17] proposed a modular architecture-level model of parametric variation to characterize variation-incurred unpredictability at an early design stage. This paper examines the case of FinFETs in particular. Tang et al.[40] presented a detailed delay/power estimation of FinFET at the circuit level. They used statistical modelling to estimate the process variations impact on delay by comparing the process variation impact for the 22 nm and 14 nm technologies. In Ref. [41], Tang et al. propose McPAT-PVT to address the PVT impact at the architectural level. They focus on the memory, network-on-chip and cores in a multicore platform. They provide a simulation platform to estimate PVT impact on a specific multicore design. In Ref. [42], Baravelli et al. studied the impact of line-edge-roughness on FinFETs for sub-45 nm technologies. The authors used a hydrodynamic model with density-gradient approximation for carrier transport, along with work-function calibration. Patel et al. explored the impact of line-edge-roughness on 13 nm technology FinFETs.[43] They presented a model to estimate the performance of double-gate devices considering the impact of line-edge-roughness.

In the literature, most of the work used lengthy and exhaustive simulations to estimate the delay. Such methods are prohibitive at large-scale system-level architecture design due to simulation time. In this work we propose a simple yet effective delay equation to be used at the architectural level. Statistical methods can be used to estimate the process variation impact and then the proposed delay equation is used to calculate the delay. The delay numbers are then compared to hSpice for validation.

4. DELAY MODELLING

In this section, we discuss the proposed model used to estimate the transistor delay, taking voltage scaling under process variation into consideration.

4.1. FinFET Structure and Operation

FinFETs are multi-gate transistors that offer better gate-channel control in small feature technologies. They show higher performance and lower leakage compared to the CMOS planar.[44] As shown in Figure 2, the gate is made of a thin fin that connects the source and the drain together to make the channel. The channel is sandwiched between two side gates on two opposite sides. The dimensions of the gate are measured through the gate length, oxide thickness, fin width, and fin height. The operation of the FinFET transistor is very similar to the CMOS planar. The performance improvement comes from the extended gate-channel border area. The effective width of the FinFET is calculated in terms of the fin height and width, i.e., $W_{\text{eff}} = 2 * H_{\text{fin}} + W_{\text{fin}}$.[45]

4.2. Process Variation Modelling

Process variation affects FinFET parameters such as its dimensions and threshold voltage. For instance, the fin height and width, affect the gate's effective width, length, oxide thickness, etc. The variation of the transistor dimensions impacts the concentration of dopants as well, which also affects the threshold voltage.[17]

Process variation is usually captured by the threshold voltage, the gate length, and the fin height.[17,26] For every mentioned parameter, the impact of process variation is estimated using two components: systematic and random. The systematic component is calculated using a multivariate normal distribution.[17] Thus, the chip is divided into small fragments. Each fragment is given a normal distribution of the given parameter.

The systematic variation of V_{th}, is calculated using a distance dependent model with spherical correlation shape function, namely $\rho(r)$, where r is the distance between fragment X and Y. This type of modelling is simple enough to be used for system-level architectures and it is known to match empirical data from silicon measurements.[3,17,26,46] The spatial correlation resulting from systematic effects is captured using the following function:

$$\rho(r) \begin{cases} 1 - \dfrac{3r}{2\varphi} + \dfrac{r^3}{2\varphi^3}, & r \le \varphi \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

The random variation is represented by the standard deviation σ_{rand} and is due to uncorrelated random effects. Consequently, the random and systematic components are

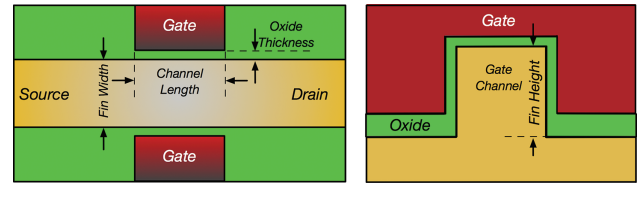


Fig. 2. Top and side section view of FinFET structure.

IP: 151.100.179.242 On: Tue, 01 Oct 2019 08:17:04
Copyright: American Scientific Publishers
Delivered by Ingenta

considered to have equal impact on the total variation. Thus, after adding the random component, the total V_{th} standard deviation is:

$$\sigma_{total} = \sqrt{\sigma_{rand}^2 + \sigma_{sys}^2} \quad (2)$$

Examples of the transistor parameters generated using the statistical methods that have been discussed are shown in Figures 3 and 4. Figure 3 shows the V_{th} sample distribution assuming the variation to be $\sigma_{total} = 25\%$. Figure 4 shows the L sample distribution assuming the variation to be $\sigma_{total} = 12.5\%$. Both figures are assumed for the 7 nm technology with the nominal V_{th} to be 62.4 mV. These variation ranges are used later in the hSpice to generate accurate delay values that makes the proposed delay equation.

4.3. Proposed Delay Model for Voltage Scaling Under Process Variation

The alpha-power delay equation is a conventional delay equation used to estimate the delay in CMOS gates. The delay equation incorporates transistor's process and operational features, as shown in (3). Some of these features such as the gate width (which includes fin height and width for FinFETs) and gate length, threshold voltage and oxide thickness, are subjected to variability due to low precision in the fabrication process in the sub-20 nm range. Thus, this equation can directly evaluate the delay with respect to any change in the parameters we have mentioned. The equation also includes the supply voltage which can also be used for voltage scaling calculations.

$$D = \frac{C_{eff}}{\mu \cdot C_{ox} \cdot (W/L)} \times \frac{V_{dd}/2}{(V_{dd} - V_{th})^{\alpha}} \quad (3)$$

given that C_{eff} is the effective switching capacitance of the transistor, W and L are the effective width and length,

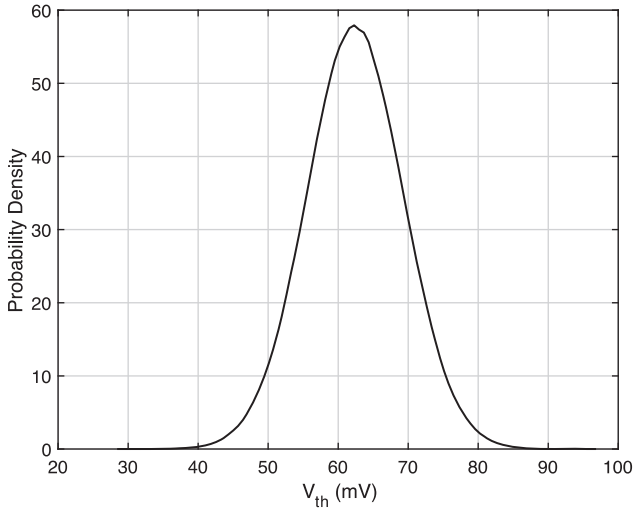


Fig. 3. V_{th} probability density function generated for 7 nm technology with nominal $V_{th} = 62.4$ mV, and $\sigma_{total} = 25\%$.

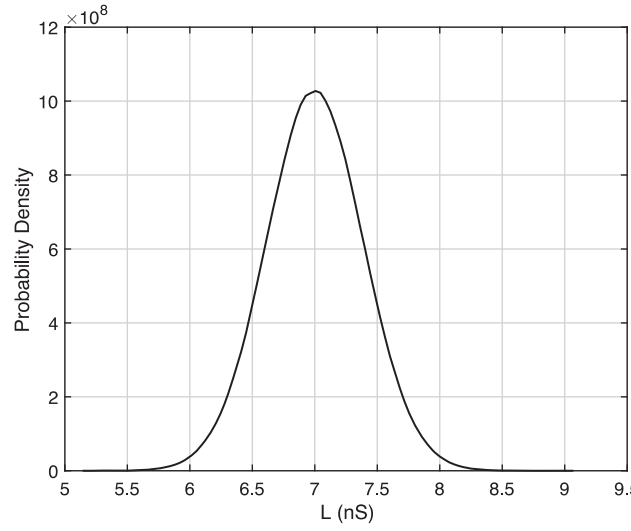


Fig. 4. L probability density function generated for 7 nm technology, and $\sigma_{total} = 12.5\%$.

V_{dd} is the supply voltage, V_{th} is the threshold voltage, and α, μ and C_{ox} are technology dependent parameters.

The alpha-power model can be extended to evaluate the delay at the system-level. Although, Eq. (3) is considered for a single transistor, it can be used to evaluate the critical path delay of a logic block. The critical path can be modelled as a chain of inverters.[3, 7, 47] The effective capacitance in the equation would then represent the capacitance of all switching transistors along the inverter chain. Thus, it can be used to estimate the total delay of a critical path of a logic block.

IP: 13.1.180.196 On: Tue, 01 Oct 2019 02:17:14
Copyright: American Scientific Publishers
Delivered by Ingenta

The switching delay of the critical path is calculated using the alpha-power model. However, this delay equation might be too simple for FinFET devices, and particularly for small feature technologies beyond 20 nm. Thus, the model has to be adjusted to handle such small technologies with acceptable accuracy. In this paper, the alpha-power model is modified such that it is suitable for sub-20 nm technologies for both planar and FinFET CMOS. The proposed model is shown in Eq. (4), where b_1, b_2, b_3, and b_4 are fitting parameters.

$$D = b_1 \frac{C_{eff}}{\mu \cdot C_{ox} \cdot (W/L)} \times \frac{V_{dd}/2}{(b_2 \cdot V_{dd} - b_3 \cdot V_{th})^{b_4}} \quad (4)$$

The proposed equation is used to evaluate the critical path delay of a processing core in different supply voltage regimes. The impact of process variation is quantified through transistor features, such as the threshold voltage, the channel length, the fin height and width (or the effective device width). The fitting parameters are calculated for different gate length L_{eff}, effective gate width W_{eff} (which includes the fin height and width), and threshold voltage V_{th} values of a given technology, representing process variations. The fitting parameters are obtained by fitting the given equation to the delay from hSpice.

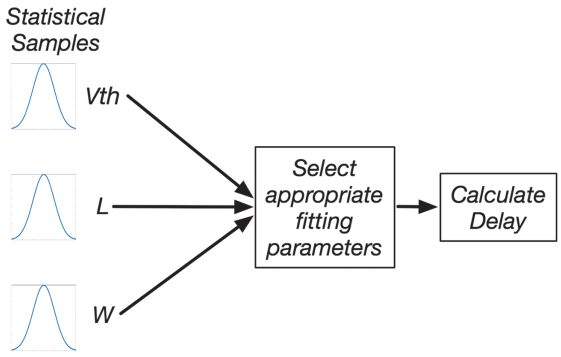


Fig. 5. Proposed delay calculation.

Figure 5 shows the proposed delay calculations. Transistor parameters such as V_{th}, L, and W are generated using statistical methods as discussed in Section 4.2 (samples shown in Figs. 3 and 4). The numbers are then used to select the proper fitting parameters obtained from this study and the delay can be calculated using Eq. (4).

5. RESULTS AND ANALYSIS

In this section, we discuss the experimental setup, results, and the analysis of these results.

5.1. Experimental Setup

Predictive Technology Models (PTM) of Ref. [48] are used to create inverter chains implemented and run in hSpice to produce real delay numbers.[49] An inverter chain with FO4 transistor sizing is built and simulated in hSpice. The value of the supply voltage, gate length, fin height and width, and threshold voltage are changed within a predefined range to measure the inverter chain delay. The delay values due to voltage scaling for a given L_{eff}, W_{eff} and V_{th} are then fitted into the proposed delay Eq. (4) and the fitting parameters b_1, b_2, b_3, and b_4 are obtained. A nonlinear least-square regression using matlab is used to fit the modelled equation into hSpice delay.[50]

Table I shows the nominal values considered for different process technologies in this work. The parameters: L, and H_{fin} are varied by +/−20%, while V_{dd}, and V_{th}, are varied by +/−40%. These ranges are set to cover process variation impact while considering voltage scaling and near-threshold operation. The standard deviation used in the literature for small feature technologies ranges from 15% to 30%[28,51] depending on the supply voltage; the lower the V_{dd} the higher the variation. Near-threshold operations claim a reduction in the supply voltage by 30% from the nominal voltage.[52] Such voltage reduction increases when process variation is considered. These ranges are also justified by the probability distributions given in Figures 3 and 4. As shown in Figure 3, the range for the V_{th} can vary from 40 mV to 85 mV (i.e., −36% to +36%) given that the nominal V_{th} is 62.4 mV and standard deviation is 25. Similarly in Figure 4, the range for L_{eff} goes from about 5.75 nm to 8.25 nm (i.e., −17% to +17%) given that the nominal L_{eff} is 7 nm and the standard deviation is 12.5. Note that some corner cases resulted in failed hSpice simulation. Within the given range, 25 points are considered during the simulation for each of the given parameters. The inverter chain length was assumed to be 20 gates.

Table I. Nominal values of considered technologies.

Tech	L (nm)	W (nm)	H_{fin} (nm)	W_{fin} (nm)	T_{ox} (nm)	V_{dd} (mV)	V_{th} (mV)
FinFET	7	42.5	18	6.5	1.15	700	62.4
	10	50	21	8	1.2	750	142
	14	56	23	10	1.3	800	196.5
Planar	16	32	–	–	0.95	700	48
	22	44	–	–	1.05	800	50

5.2. Results and Analysis

The delay is first simulated in hSpice for each of the parameter's values. Then matlab function *nlinfit*[50] is used to fit the model Eq. (4) into the hSpice delay numbers to obtain the fitting parameters. Note that the number of simulated cases is the combination of 25 points for L, 25 points for H_{fin} (for finFETs only), 25 points for V_{th}, and 25 points for V_{dd}, with total of 390,625 hSpice simulations for each technology listed in Table I. In all the figures shown, we selected only a subset of the simulated cases for clarity purposes.

Figure 6 shows the delay numbers obtained from hSpice simulation and the delay equation. The figure plots the delay while varying the effective threshold voltage. Each curve in the figure represents the delay-V_{th} relationship for a given gate-length value. It also shows the delay value

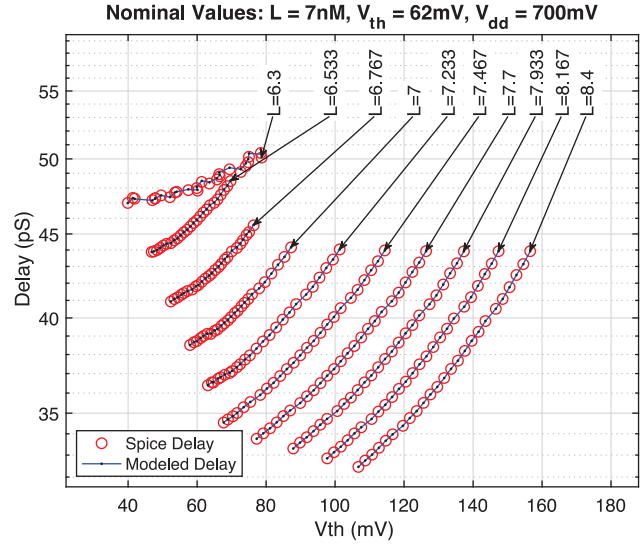


Fig. 6. Delay versus threshold for different gate-length values, 7 nm FinFET technology.

IP: 71.19.144.174 On: Tue, 01 Oct 2019 09:17:04
Copyright: American Scientific Publishers
Delivered by Ingenta

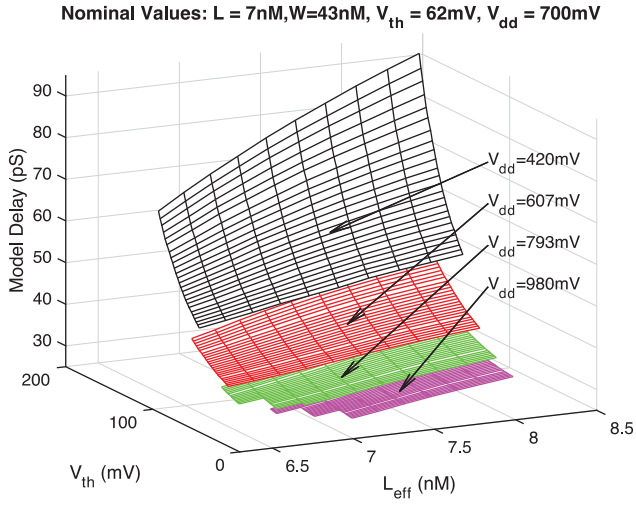


Fig. 7. Modelled delay of selected simulation points.

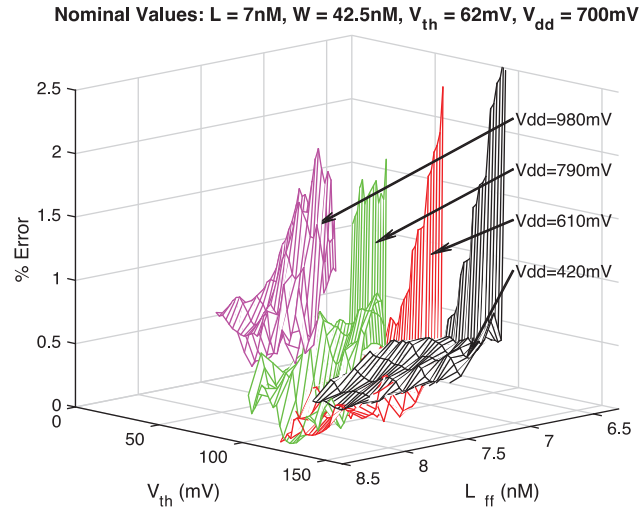


Fig. 9. Error percentage averaged across W_{eff}, w.r.t. V_{th}, L_{eff} and V_{dd}.

obtained from hspice and the delay obtained from the proposed model. The covered points go from $L_{\text{eff}} = 6.3$ nm to 8.4 nm. Note that each point on the curve represents the average delay across different V_{dd} and W_{eff} values, and for a given L_{eff} and V_{th} values. In general, the delay curves shown take a smooth trajectory. As the L_{eff} reaches 6.3 nm and V_{th} less than 80 mV, the curve starts to show some irregularities. Figure 7, on the other hand, shows the modelled delay plotted versus L_{eff} and V_{th}, where each plane is calculated under different supply voltages.

Figure 8 shows the values of the fitting parameters b_1, b_2, and b_4 (b_3 is always equal to 1, so it is eliminated from the plot). The values shown are obtained for different (L_{eff}, W_{eff}) and V_{th}. In this figure, the W_{eff} value changed along with L_{eff} to simplify the demonstration. For each (L_{eff}, W_{eff}) value, three curves are plotted representing b_1, b_2, and b_4 against changing the threshold voltage V_{th} on the x-axis. The curves do not cover the full range of V_{th} due to hSpice/PTM simulation fail. Also, the L_{eff} value starts at 6.3 nm, which is less than the nominal amount by only 10% due to the same problem. As shown, b_2 and b_4 converge into 0.65 and 1.15, while b_1 is more spread out from 0.7 to 0.35 (not including the first case $L = 6.3$, where the curve goes from 0.7 all the way to 3.9). Note that not all the (L_{eff}, W_{eff}) values are labeled in the figure to improve clarity.

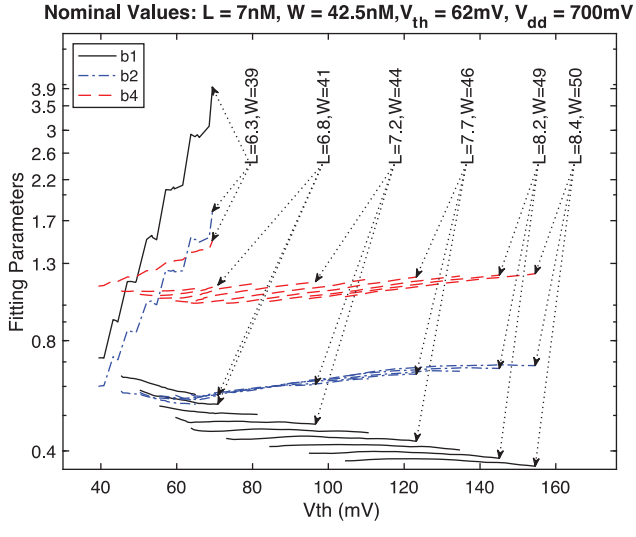


Fig. 8. Fitting parameters b_1, b_2, and b_4 ($b_3 = 1$).

Figure 9 shows the error percentage averaged across W_{eff} range plotted with respect to changing the V_{th}, L_{eff}, and V_{dd}. The error shown is calculated for the 7 nm technology. As shown, the error increases above 2% as the L_{eff} gets beyond 6 nm and V_{th} above 100 mV. A higher L_{eff} keeps the error rate at less than 1%. This is an acceptable approximation for delay calculations at the system level.

Figures 10(a) and (b) show the fitting parameters for the FinFET 10 nm and 14 nm technologies. As stated earlier, in order to simplify the demonstration, we did not show all simulation points. The figure shows a similar trend to the one for the 7 nm technology. For both 10 nm and 14 nm technologies, b_2 and b_4 are limited to a confined range. The b_4 parameter converges into 1.1 for both technologies (similar to 7 nm as well). Note that b_4 represents the α in the original alpha-power model as shown in Eq. (3). So, it is safe to assume that α (or b_4) is around 1.1 for all three technologies. Parameter b_2, which is the fitting parameter multiplied by the V_{dd} in Eq. (4), converges towards 0.65 for the 7 nm, 0.8 for the 10 nm, and 0.9 for the 14 nm. Thus, it is clear that b_2 decreases as the technology decreases. Finally, the b_1 parameter is more spread out from 0.5 to 1.5 for the 10 nm, and from 0.6 to 1.7 for the 14 nm. Parameter b_1 is the fitting parameter proportional to the delay value. In all cases, the trend is that the increase in the transistor features, such as V_{th}, L_{eff} and W_{eff}, seems for force b_1 to decrease.

Figures 11(a) and (b) show the fitting parameters for the planar technologies 16 nm and 22 nm. Parameter b_4

IP: 131.180.130.242 On: Tue, 01 Oct 2019 15:17:04
Copyright: American Scientific Publishers
Delivered by Ingenta

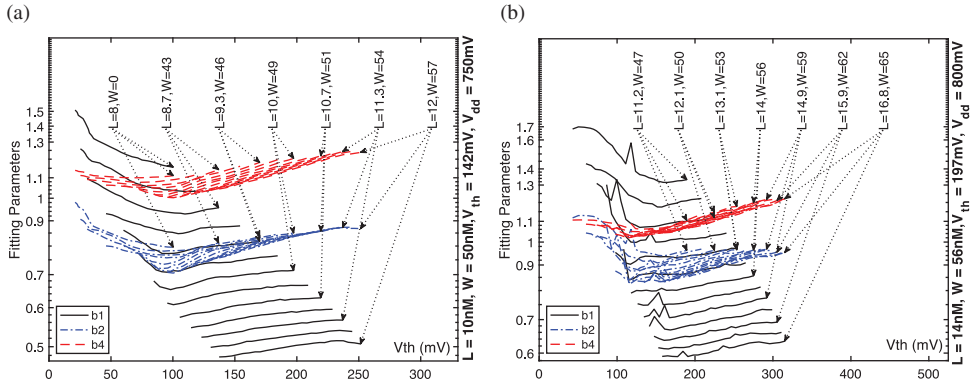


Fig. 10. Fitting parameters for the FinFET (a) 10 nm, and (b) technologies.

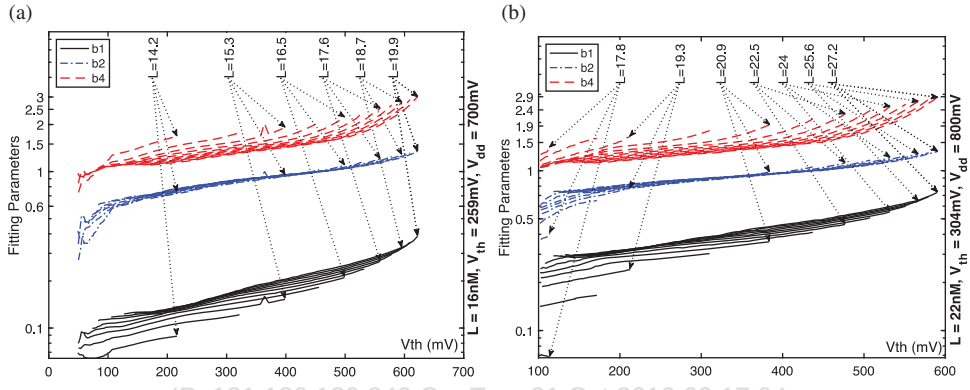


Fig. 11. Fitting parameters for the planar (a) 16 nm and (b) 22 nm technologies.

IP: 131.180.130.242 On: Tue, 01 Oct 2019 09:17:04
Copyright: American Scientific Publishers
Delivered by Ingenta

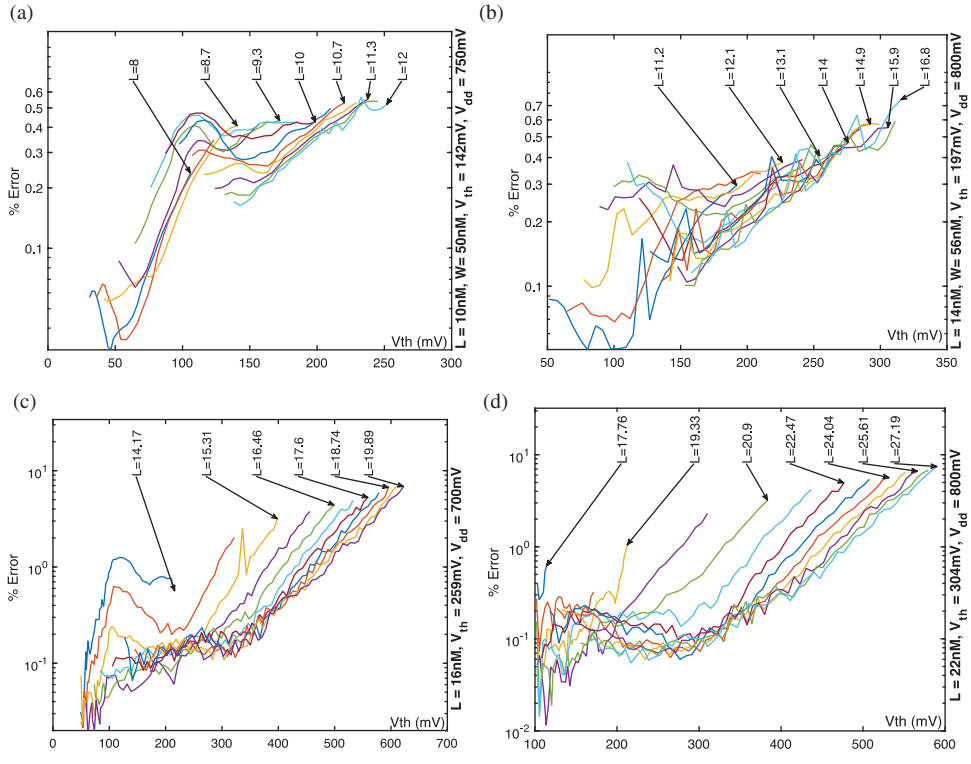


Fig. 12. Error percentage of the proposed delay equation compared to hSpice for the FinFET (a) 10 nm, (b) 14 nm, and planar (c) 16 nm and (d) 22 nm technologies.

Table II. Total average % error compared to hSpice results.

Technology (nm)	Total average % error (%)
FinFET	
7	0.43
10	0.29
14	0.26
Planar	
16	4.9
22	13.44

is around 1.3 or slightly higher than that of the FinFETs. Fitting parameter b_2 is around 0.75 for both technologies. Parameter b_1, on the other hand, is around 0.1 to 0.3, and it is more concise compared to the FinFET technologies. The channel width, W_{eff}, was not included in the simulation of the planar technologies, because usually the W_{eff} variation is lumped into the L_{eff} variation for planar CMOS.[26]

Figure 12 shows an error in the delay calculated using the proposed model of Eq. (4) compared to hSpice simulation. The plots cover all four technologies, 10 nm and 14 nm FinFETs and 16 nm and 22 nm Planar CMOS. The percentage of error that is shown is the average across supply voltage and channel width values plotted for different gate length values against the threshold voltage.

It should be noted that only selected simulation cases are plotted to simplify the figure. The trend is that errors increase as the threshold voltage increases. Table II shows the percentage of error across all simulations for all technologies under investigation. As shown, the numbers are very reasonable. The planar technologies show higher error numbers compared to the FinFETs.

The proposed delay model along with the fitting parameters numbers can be used to estimate the delay under process variation in different supply voltage regimes at the system-level with reasonable accuracy. As stated earlier, samples of L_{eff}, W_{eff}, and V_{th} are generated using statistical methods. Thus, the numbers of L_{eff}, W_{eff}, and V_{th} used in this work might not match those generated by these methods. In such cases, interpolation can be used to pick the corresponding fitting parameters to calculate the delay.

6. CONCLUSION

In this paper, we proposed a simple yet effective delay equation that can be used at the architectural level. The proposed delay equation is used to calculate the delay under process variation considering multi supply voltage designs. The target technologies are FinFETs with 7 nm, 10 nm, and 14 nm, and planar with 16 nm and 22 nm gate lengths. The results obtained here showed an average percentage error between 0.5% and 14% compared to hSpice simulation.

IP: 131.180.130.242 On: Tue, 01 Oct 2019 09:17:04
Copyright: American Scientific Publishers
Delivered by Ingenta

References

1. T. Cui, J. Li, Y. Wang, S. Nazarian, and M. Pedram, An exploration of applying gate-length-biasing techniques to deeply-scaled FinFETs operating in multiple voltage regimes. *IEEE Trans. Emerg. Top. Comput.* 6, 172 **(2018)**.
2. K. Asanovic, B. C. Catanzaro, D. A. Patterson, and K. A. Yelick, The landscape of parallel computing research: A view from berkeley. *EECS Dep. Univ. Calif. Berkeley Tech. Rep. UCBEECS2006183* 18, 19 **(2006)**.
3. S. S. Majzoub, R. A. Saleh, S. J. E. Wilton, and R. K. Ward, Energy optimization for many-core platforms: Communication and PVT aware voltage-island formation and voltage selection algorithm. *IEEE Trans. Comput. Des. Integr. Circuits Syst.* 29, 816 **(2010)**.
4. A. Stillmaker and B. Baas, Scaling equations for the accurate prediction of CMOS device performance from 180 nm to 7 nm. *Integration* 58, 74 **(2017)**.
5. Y. Li, C. Hwang, S. Member, and T. Li, Process-variation effect, metal-gate work-function fluctuation, and random-dopant fluctuation in emerging CMOS technologies. *IEEE Trans. Electron Devices* 57, 437 **(2010)**.
6. A. Arasteh, M. Hossein Moaiyeri, M. Taheri, K. Navi, and N. Bagherzadeh, An energy and area efficient 4:2 compressor based on FinFETs. *Integr. VLSI J.* 60, 224 **(2018)**.
7. X. Lin, Y. Wang, and M. Pedram, Joint sizing and adaptive independent gate control for FinFET circuits operating in multiple voltage regimes using the logical effort method, *2013 IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*, **(2013)**, pp. 444–449.
8. Q. Xie, X. Lin, Y. Wang, M. J. Dousti, A. Shafaei, M. Ghasemi-Gol, and M. Pedram, 5 nm FinFET standard cell library optimization and circuit synthesis in near-and super-threshold voltage regimes, *2014 IEEE Computer Society Annual Symposium on VLSI* **(2014)**, pp. 424–429.
9. K. Xu, R. Patel, P. Raghavan, and E. G. Friedman, Exploratory design of on-chip power delivery for 14, 10, and 7 nm and beyond FinFET ICs. *Integration* 61, 11 **(2018)**.
10. M. Bahadori, M. Kamal, A. Afzali-Kusha, and M. Pedram, A comparative study on performance and reliability of 32-bit binary adders. *Integr. VLSI J.* 53, 54 **(2016)**.
11. P. Singh, B. S. Reniwal, V. Vijayvargiya, V. Sharma, and S. K. Vishvakarma, Ultra low power-high stability, positive feedback controlled (PFC) 10T SRAM cell for look up table (LUT) design. *Integration* 62, 1 **(2018)**.
12. I. Stamelakos, S. Xydis, G. Palermo, and C. Silvano, Workload- and process-variation aware voltage/frequency tuning for energy efficient performance sustainability of NTC manycores. *Integration* **(2018)**, In Press.
13. J. Alshraiedeh and A. Kodi, An adaptive routing algorithm to improve lifetime reliability in NoCs architecture, *IEEE International Symposium on Defect and Fault Tolerance in VLSI and Nanotechnology Systems (DFT)* **(2016)**, pp. 127–130.
14. H. Nejatollahi, M. Ersali, and S. Nasab, Reliability-aware voltage scaling of multicore processors in dark silicon era. *Big Data and HPC: Advances in Parallel Computing Ecosystem and Convergence* **(2018)**, Vol. 33, pp.242–262.
15. H. Nejatollahi and M. E. Salehi, Voltage scaling and dark silicon in symmetric multicore processors. *J. Supercomput.* 71, 3958 **(2015)**.
16. Z. Hao, S. X.-D. Tan, and G. Shi, Statistical full-chip total power estimation considering spatially correlated process variations. *Integr. VLSI J.* 46, 80 **(2013)**.
17. S. K. Khatamifard, M. Resch, N. S. Kim, and U. R. Karpuzcu, VARIUS-TC: A modular architecture-level model of parametric variation for thin-channel switches. *Proc. 34th IEEE Int. Conf. Comput. Des. ICCD 2016* **(2016)**, pp.654–661.
18. A. Shafaei, Y. Wang, L. Chen, S. Chen, and M. Pedram, Maximizing the performance of NoC-based MPSoCs under total power and

power density constraints, *17th International Symposium on Quality Electronic Design (ISQED)* **(2016)**, pp. 49–56.

19. J. Lim, H. Kim, H. Oh, and S. Kang, Dynamic voltage frequency scaling-aware refresh management for 3D DRAM over processor architecture. *Electron. Lett.* 53, 910 **(2017)**.

20. N. Azizi, M. M. Khellah, V. K. De, and F. N. Najm, Variations-aware low-power design and block clustering with voltage scaling. *IEEE Trans. Very Large Scale Integr. Syst.* 15, 746 **(2007)**.

21. R. Kim, G. Liu, P. Wettin, R. Marculescu, D. Marculescu, and P. P. Pande, Energy-efficient VFI-partitioned multicore design using wireless NoC architectures, *Proceedings of the 2014 International Conference on Compilers, Architecture and Synthesis for Embedded Systems-CASES '14* **(2014)**, pp. 1–9.

22. A. R. Silva, W. Jose, H. Neto, and M. Vestias, Modeling and simulation of a many-core architecture using systemC. *Conf. Electron. Telecommun. Comput.* 17, 146 **(2014)**.

23. I. Uddin, R. Poss, and C. Jesshope, Cache-based high-level simulation of microthreaded many-core architectures. *J. Syst. Archit.* 60, 529 **(2014)**.

24. D. J. R. Ferreira, Analysis of many-core CPUs simulators. Tech. Rep., Instituto Superior Tecnico, Universidade de Lisboa, pp. 1–10.

25. S. Borkar, Thousand core chips—A technology perspective, *DAC* **(2007)**, pp.749–754.

26. S. R. Sarangi, B. Greskamp, R. Teodorescu, J. Nakano, A. Tiwari, and J. Torrellas, VARIUS: A model of process variation and resulting timing errors for microarchitects. *IEEE Trans. Semicond. Manuf.* 21, 3 **(2008)**.

27. Z. Wang, S. Kanwal, L. Wang, and A. Chattopadhyay, Automated high-level modeling of power, temperature and timing variation for microprocessor. *KMUTNB IJAST* 10, 163 **(2017)**.

28. S. Kalra and A. Bhattacharyya, Scalable α-power law based MOSFET model for characterization of ultra deep submicron digital integrated circuit design, *AEU-Int. J. Electron. Commun.* 83, 180 **(2018)**.

29. R. Jevtić, H. P. Le, M. Blagojević, S. Bailey, K. Asanović, E. Alon, and B. Nikolić, Per-core DVFS with switched-capacitor converters for energy efficiency in manycore processors. *IEEE Trans. Very Large Scale Integr. Syst.* 23, 723 **(2015)**.

30. E. Cai, D. Stamoulis, and D. Marculescu, Exploring aging deceleration in FinFET-based multi-core systems, *Proceedings of the 35th International Conference on Computer-Aided Design—ICCAD '16*, **(2016)**, pp. 1–8.

31. S. K. Samal, Y. Peng, M. Pathak, and S. K. Lim, Ultralow power circuit design with subthreshold/near-threshold 3-D IC technologies. *IEEE Trans. Components, Packag. Manuf. Technol.* 5, 980 **(2015)**.

32. S. Seo, R. G. Dreslinski, M. Woh, Y. Park, C. Charkrabari, S. Mahlke, D. Blaauw, and T. Mudge, Process variation in near-threshold wide SIMD architectures. *Proc. 49th Annu. Des. Autom. Conf.-DAC '12* **(2012)**, p. 980.

33. K. A. Bowman, A. R. Alameldeen, S. T. Srinivasan, and C. B. Wilkerson, Impact of die-to-die and within-die parameter variations on the throughput distribution of multi-core processors, *Proceedings of the 2007 International Symposium on Low Power Electronics and Design—ISLPED '07* **(2007)**, pp. 50–55.

34. N. Pinckney, S. Jeloka, R. Dreslinski, T. Mudge, D. Sylvester, D. Blaauw, L. Shifren, B. Cline, and S. Sinha, Impact of FinFET on near-threshold voltage scalability. *IEEE Des. Test* 34, 31 **(2017)**.

35. S. Saurabh and V. Vikash, Assessing the impact of temperature and supply voltage variations in near-threshold circuits using an analytical model, *Proceedings of the 2018 on Great Lakes Symposium on VLSI-GLSVLSI '18* **(2018)**, pp. 93–98.

36. S. Kalra and A. B. Bhattacharyya, Variability Study Using α-Power-Based MOSFET Model for Ultradeep Submicron Digital Circuit Design, Springer, Singapore **(2019)**, pp. 601–610.

37. J. Lee and N. S. Kim, Analyzing potential throughput improvement of power- and thermal-constrained multicore processors by exploiting DVFS and PCPG. *IEEE Trans. Very Large Scale Integr. Syst.* 20, 225 **(2012)**.

38. S. K. Samal, G. Chen, and S. K. Lim, Improving performance under process and voltage variations in near-threshold computing using 3D ICs. *ACM J. Emerg. Technol. Comput. Syst.* 13, 1 **(2017)**.

39. S. Majzoub, Instruction-based voltage scaling for power reduction in SIMD-based MPSoCs. *Journal of Low Power Electronics (JOLPE)* 7, 141 **(2011)**.

40. A. Tang, X. Gao, L.-Y. Chen, and N. K. Jha, Delay/power modeling and optimization of FinFET circuit modules under PVT variations: Observing the trends between the 22 nm and 14 nm technology nodes. *ACM J. Emerg. Technol. Comput. Syst.* 12, 42 **(2016)**.

41. A. Tang, Y. Yang, C. Y. Lee, and N. K. Jha, McPAT-PVT: Delay and power modeling framework for FinFET processor architectures under PVT variations. *IEEE Trans. Very Large Scale Integr. Syst.* 23, 1616 **(2015)**.

42. E. Baravelli, A. Dixit, R. Rooyackers, M. Jurczak, N. Speciale, and K. De Meyer, Impact of line-edge roughness on FinFET matching performance. *IEEE Trans. Electron Devices* 54, 2466 **(2007)**.

43. K. Patel, T. J. K. Liu, and C. J. Spanos, Gate line edge roughness model for estimation of FinFET performance variability. *IEEE Trans. Electron Devices* 56, 3055 **(2009)**.

44. M. Jurczak, N. Collaert, A. Veloso, T. Hoffmann, and S. Biesemans, Review of FINFET technology, *Proc.-IEEE Int. SOI Conf.* **(2009)**.

45. J. P. Colinge, FinFETs and Other Multi-Gate Transistors, Springer-Verlag, New York **(2008)**.

46. S. Majzoub, Reducing random-dopant fluctuation impact using footer transistors in many-core systems. *Integr. VLSI J.* 48, 46 **(2015)**.

47. T. Koyagi, S. Majzoub, M. Fukui, and R. Saleh, RTL delay macro-modeling with V_t and V_{dd} variability, *International Design and Test Workshop* **(2011)**.

48. Y. Cao, W. Zhao, E. Wang, W. Wang, J. Velamala, A. Balijepalli, S. Sinha, and S. Sinha, Predictive Technology Model, Nanoscale Integration and Modeling (NIMO) Group. [Online]. Available: http://ptm.asu.edu/. [Accessed: 03-Sep-2018].

49. S. Sinha, G. Yeric, V. Chandra, B. Cline, and Y. Cao, Exploring sub-20 nm FinFET design with predictive technology models, *Proceedings of the 49th Annual Design Automation Conference on—DAC '12* **(2012)**, p. 283.

50. Nonlinear regression-MATLAB nlinfit. [Online]. Available: https://www.mathworks.com/help/stats/nlinfit.html. [Accessed: 05-Sep-2018].

51. R. T. Possignolo, E. Ebrahimi, E. K. Ardestani, A. Sankaranarayanan, J. L. Briz, and J. Renau, GPU NTC process variation compensation with voltage stacking. *IEEE Trans. Very Large Scale Integr. Syst.* 26, 1713 **(2018)**.

52. U. R. Karpuzcu, K. B. Kolluru, N. S. Kim, and J. Torrellas, VARIUS-NTV: A microarchitectural model to capture the increased sensitivity of manycores to process variations at near-threshold voltages, *IEEE/IFIP International Conference on Dependable Systems and Networks (DSN 2012)* **(2012)**, pp. 1–11.

Downloaded from Scientific.Net at Tue, 05/03/2022 04:04:04
Copyright (c) Trans Tech Publications Inc. All rights reserved.
Content may be downloaded by Ingenta

Sohaib Majzoub

Sohaib Majzoub *completed his B.E. in Electrical Engineering, Computer Section at BAU 2000, and his M.E. degree from AUB, 2003 Lebanon. Then he worked for one year at the Processor Architecture Lab at the Swiss Federal Institute of Technology, Lausanne Switzerland. In 2010, he finished his Ph.D. working at the System-on-Chip research Lab, University of British Columbia, Canada. He worked for two years as assistant professor at American University in Dubai, Dubai, UAE. He then joined King Saud University in 2012, KSA. In 2015, he joined University of Sharjah, UAE, as a faculty in the electrical and computer engineering department. His research field is delay/power modeling at the system level. He is IEEE, Computer Society, Solid State Society, and Circuits and Systems member.*

Mottaqiallah Taouil

Mottaqiallah Taouil *received the M.Sc. and Ph.D. degrees (both with Hons.) in computer engineering from the Delft University of Technology, Delft, The Netherlands. He is currently a Post-Doctoral Researcher with the Dependable Nano-Computing Group, Delft University of Technology. His current research interests include reconfigurable computing, embedded systems, very large scale integration design and test, built-in-self-test, and 3-D stacked integrated circuits, architectures, design for testability, yield analysis, and memory test structures.*

Said Hamdioui

Said Hamdioui *is currently a Chair Professor on Dependable and Emerging Computer Technologies at the Computer Engineering Laboratory of the Delft University of Technology (TUDelft), the Netherlands. Prior to joining TUDelft, Hamdioui worked for Intel Corporation (Californina, USA), Philips Semiconductors R&D (Crolles, France) and for Philips/NXP Semiconductors (Nijmegen, The Netherlands). His research focuses on two domains: Dependable CMOS nano-computing (including Reliability, Testability, Hardware Security) and emerging technologies and computing paradigms (including 3D stacked ICs, memristors for logic and storage, in-memory-computing). He owns one patent and has published one book and co-authored over 170 conference and journal papers. He delivered dozens of keynote speeches, distinguished lectures, and invited presentations and tutorial at major international forums/conferences/schools and at leading semiconductor companies. Hamdioui is a Senior member of the IEEE, Associate Editor of IEEE Transactions on VLSI Systems (TVLSI), and he serves on the editorial board of IEEE Design and Test, and of the Journal of Electronic Testing: Theory and Applications (JETTA). He is also member of AENEAS/ENIAC Scientific Committee Council (AENEAS = Association for European NanoElectronics Activities).*

IP: 131.180.130.242 On: Tue, 01 Oct 2019 09:17:04
Copyright: American Scientific Publishers
Delivered by Ingenta