Evolving On-Chip Power Delivery through Particle Swarm Optimization

Divya Pathak
Department of Electrical and Computer Engineering
Drexel University
Philadelphia, PA
Email: divya.pathak@drexel.edu

Ioannis Savidis
Department of Electrical and Computer Engineering
Drexel University
Philadelphia, PA
Email: isavidis@coe.drexel.edu

***Abstract*—An evolving on-chip power delivery method is developed for the adaptive voltage assignment of a given voltage domain. The reference voltages of the on-chip voltage regulators (OCVRs) are determined and set through particle swarm optimization (PSO) to negate the effects of transistor aging, process variation, and power supply noise induced variation in the load circuit, OCVRs, and on-chip timing sensors. The on-line learning of the optimum voltages that evolve with time reduces the static voltage guard-band added during the design of the power delivery network for worst case process, temperature, and aging induced timing variation of a digital circuit. Applying the proposed PSO based on-chip power management technique ensures a minimal voltage assignment without incurring any timing violations on the critical paths, which also evolve with time. The run-time adaptive voltage delivery technique is applicable to any processor architecture as demonstrated through simulation of a four core multi-processor implemented in a 7 nm PTM FinFET technology. Results indicate an average reduction of 35% and 38% in, respectively, the dynamic power consumption and transistor aging.**

***Index Terms*—particle swarm optimization, machine learning, power supply noise, on-chip voltage regulators, transistor aging.**

I. Introduction

The advent of many-core platforms with heterogeneous architectures has introduced new challenges for dynamic power management. A heterogeneous architecture offers higher energy efficiency for performance intensive data centers executing deep learning workloads. However, the power delivery to heterogeneous many-core circuits through on-chip voltage regulators (OCVR) is an ongoing and complex research topic [1]. The optimal energy efficiency of a system is achieved with distributed and heterogeneous OCVRs that are interconnected through the power delivery network (PDN). The large and sustained current demand of the cores subjects the OCVRs to increased aging effects and a higher susceptibility to process variation and noise. In addition, due to the finite latency of the OCVRs to react to changes in the load current, large dI/dt events in the cores result in power supply noise. Therefore, advanced *run-time* power management techniques are needed to mitigate timing errors due to process, voltage, and temperature variation and aging.

The existing power management strategies for multi-core systems employ a central power management unit (PMU) that controls the operating voltage (and frequency for DVFS) of the cores and the corresponding interconnect fabric. The decision to scale the voltage and frequency is executed by the operating system. The PMU provides the data gathered by the on-chip sensors (operating voltage, current, and/or temperature) to the operating system governor. In addition to the data provided by the physical sensors, the activity counters implemented in the architecture of the processor provide additional guidance on the appropriate operating voltage and frequency to the governor. A centralized power management strategy does not scale well as the number of cores and/or accelerators within an integrated circuit (IC) increases. As described in [2], the energy savings from the implementation of an active voltage guard-band in an IBM POWER7 diminishes as the number of active cores increases.

The design of the power distribution network is based on the optimization of a static voltage margin. A fixed voltage margin or guard-band is added to the power supply voltage to compensate for noise induced by the current drawn through the parasitic impedance of the PDN of the integrated circuit, package, and board in addition to the finite *IR* drop due to the resistive path from the voltage regulator to the load circuits. In sub-20 nm nodes, the increase in process variation, the complex power-thermal interactions, and, most importantly, the reduced voltage margin between the transistor operating voltage and threshold voltage requires novel methodologies for the effective implementation of the PDN.

In addition to technology scaling, FinFET based circuits produce higher current densities, which leads to an increase in the dynamic power consumption, voltage noise, and thermal density, as well as greater susceptibility to electromigration within the interconnects of the PDN [3]. Power management techniques such as power gating increase the inrush current, and therefore, the power supply noise. Optimization of decoupling capacitors alone does not effectively address the increased power supply noise [4]. Due to the complex interdependence of the various components of the power delivery network of sub-20 nm technology nodes, a robust and cost effective PDN using existing design techniques is difficult to implement. Techniques based on vector-less peak power result in an overcompensated and expensive PDN [5]. Vector or stimulus based PDN design with analysis of the peak power consumption provided through emulators is computationally expensive.

978-1-7281-5758-0/19/$31.00 ©2019 IEEE

Machine learning (ML) based techniques have recently been applied to the design of the PDN [6] to minimize the required routing resources while meeting the constraints for *IR* drop and electromigration. Another approach that utilizes ML algorithms for the design of the PDN leverages the power profile from a prior completed circuit to implement the PDN of a current circuit. The two ICs differ in functionality, but learning algorithms capture the physical characteristics of the circuits to effectively execute an informed decision that optimizes the PDN. With the rapid advancement in ML techniques, the objective is to replace decisions in the design of an integrated circuit based on human experience with ML algorithms. However, ML techniques applied to the PDN during the design phase do not mitigate the impact of circuit aging and power supply noise at run-time.

To effectively address the challenges faced in the delivery of power to circuits in advanced technology nodes, a *run-time machine learning based technique* is proposed for reliable and energy efficient power delivery. In this paper, an evolving on-chip voltage assignment is implemented with distributed OCVRs. The on-chip PDN self-learns and regulates the local voltages intelligently and autonomously to minimize the voltage guard-band without inducing any timing failures. Latched tap delay lines are integrated to sense changes in the timing margin of the critical paths, with the collected data provided to an on-line learning algorithm. The data from the circuit-level sensors is utilized by the run-time algorithm to reduce the latency of executing power modes including DVFS and to improve the accuracy of the prediction of the optimal voltage and frequency for a given workload.

The remainder of the paper is organized as follows. The proposed particle swarm optimization based power management methodology is described in Section II. Simulation results of an implementation of the PSO are provided in Section III. Concluding remarks are provided in Section IV.

II. Learning algorithm for an evolvable PDN

Classical methods of optimization including linear programming, non-linear programming, Newton's method, quadratic programming, and sequential unconstrained minimization assume the optimization of continuous variables, which yields local optimum solutions [7]. The on-chip power delivery system implemented with OCVRs contains both discrete and continuous control variables. Applying techniques for continuous variables to discrete variables results in both an increase of the objective function and violations of inequality constraints. Evolutionary programming methods including simulated annealing, genetic algorithms (GA), tabu search, and particle swarm optimization are better suited for discrete variables and non-differentiable objective functions [8].

Genetic algorithms are based on Darwinian theories of evolution and implement processes analogous to genetic recombination and mutation to promote the *evolution* of a population that best satisfies a predefined objective. A genetic algorithm was used for post-silicon tuning of the clock delay [9] to mitigate timing errors due to power supply noise and temperature. The primary disadvantage of applying a GA for run-time applications is an increase in the latency to optimize both circuit and architectural parameters as the search space increases. The searching latency and the convergence to a globally sub-optimal result when proceeding through the solution space are limitations that apply to other evolutionary techniques including simulated annealing and ant-colony optimization.

The particle swarm optimizer (PSO), however, offers a robust and simple implementation that produces superior results as compared to other evolutionary algorithms. Prior research has shown that the PSO offers different routes through the problem hyperspace as compared to GA and other optimization algorithms [10]. The PSO algorithm provides an ideal solution for run-time control of the on-chip power supply voltages as the overhead to store results after each execution of the algorithm is small and the circuit implementation of the algorithm is not complex.

A. Problem formulation

The proposed power management methodology applies particle swarm optimization to a many-core system that includes distributed OCVRs and distributed timing and aging sensors. The aging sensors measure the timing margin at the operating frequency of the circuit and transmit any violation(s) to the online PSO algorithm controlling the voltage reference to each OCVR. Timing sensors that capture clock edges that penetrate different logical depths of a buffer chain are implemented in critical paths to characterize local temporospatial variations in process, temperature, and voltage. The reference voltage of each OCVR is independently modulated based on the optimized values provided by the online PSO algorithm. Alternatively, to reduce the complexity of implementing the circuit that generates the reference voltage of each OCVR, the PSO provides a global optimum voltage applied as a common reference to all OCVRs.

The particle swarm optimizer operates on a set of *particles* (OCVRs), where the position of each particle x_i represents the reference voltage applied to the i^{th} OCVR and is a potential solution to the optimization problem. The distributed OCVRs operate as a swarm, locally optimizing the operating voltage of the IC to compensate for aging related degradation in both the load and OCVR circuits while applying the smallest needed guard-band that prevents timing violations on the local critical paths. Each execution of the PSO algorithm produces a local best reference voltage (P_{best}) for each OCVR and the global best voltage (G_{best}) for the voltage domain supplied by the swarm of OCVRs. The P_{best} of an OCVR is a function of the characterized time delay by the nearest timing sensor. The P_{best} values, as given by (1), are the local optimal voltages set to account for local circuit characteristics including *IR* drop, d*I*/d*t* noise, process and aging induced V_{th} degradation, and hotspots. The best position among the entire particle population is recorded as G_{best}. The G_{best} for the swarm of OCVRs, as given by (2), is the maximum P_{best} value obtained across all OCVRs. The P_{best} and G_{best} are functions of time and set the lowest power supply voltage to the circuit that results in no timing violations. The position (reference voltage) of a particle (OCVR) constantly changes based on

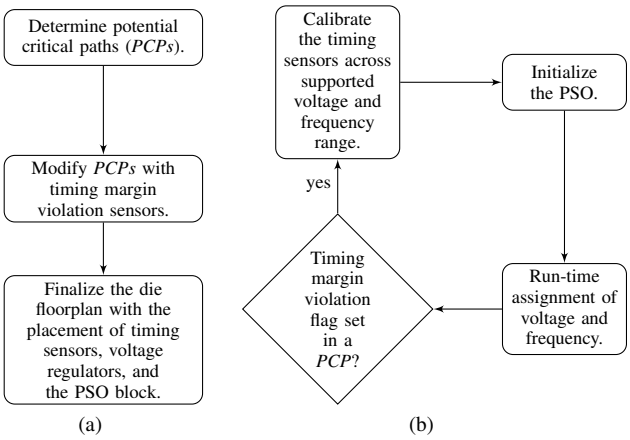


Fig. 1: Methodology that implements run-time adaptive voltage assignment through execution of the PSO algorithm as described by (a) pre-silicon design steps, and (b) post-silicon process steps.

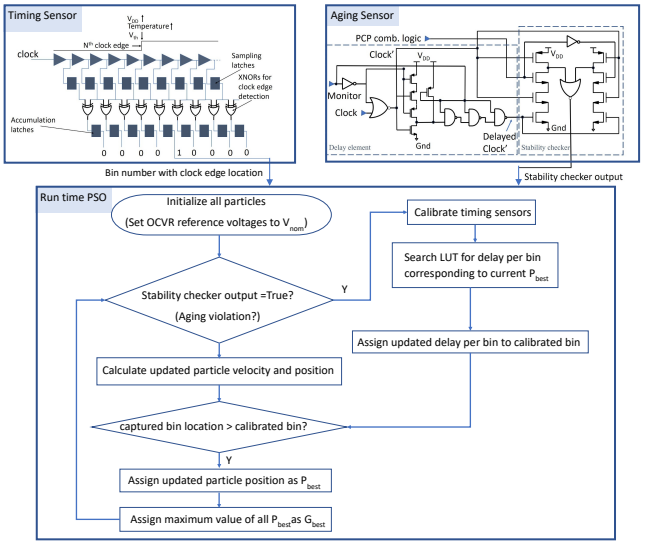


Fig. 2: Data flow graph of the execution of the PSO with inputs from aging sensors [11] and timing sensors (latched tapped delay lines [12]) implemented in a 7 nm PTM technology [13].

the experiences of both a given particle and the remaining particles in the swarm.

$$P_{best_i} = f(Temp(t), V_{noise}(t), V_{th\_aging}(t), W_{load}(t)) \quad (1)$$

$$G_{best} = max(P_{best_i}) \quad (2)$$

B. Hardware implementation

The steps required to implement the PSO are shown in Fig. 1(a) for a system with multiple processing cores. Existing tools for statistical static timing analysis are applied to each voltage domain to determine the set of timing paths that exhibit high delay or are statistically likely to evolve to become the paths with the maximum delay as the circuit ages. The set of likely timing paths are termed as potential critical paths (*PCPs*). An aging sensor [11] is integrated within each *PCP* to monitor violations of the timing margin. In addition, a timing sensor, such as a latched tapped delay line [12], is placed in close proximity to each OCVR to monitor local process, voltage, and temperature variations. The placement of the distributed voltage regulators and time based sensors is completed with respect to the distribution of the *PCPs* across the IC.

The post-silicon procedure that applies the PSO is depicted in Fig. 1(b). During the first power up of the circuit, the distributed timing sensors are calibrated across the supported voltage and frequency ranges of the IC while executing a known workload that produces the lowest variation in the activity factor. The calibrated timing bins for each sensor at each frequency are stored in an on-chip look up table (LUT). Due to intra-die process variation, the calibrated bins vary for each timing sensor. The LUT is accessed by the on-line PSO, which compares the latest output from a given timing sensor with the calibrated data stored in the corresponding LUT. As long as there is no violation of the timing margin of any of the *PCP* and the captured timing bins are lower than the calibrated bins, the P_{best} position is updated. In the case of a detected violation in the timing margin of any of the *PCPs* by the aging sensors, a recalibration of the time based sensors is performed.

The P_{best} and G_{best} values evolve with time as well as with changes in environmental conditions, operating temperature, and circuit aging. The combined effect of aging and environmental conditions on the load circuits, voltage regulators, and the time based sensors is compensated for by the computed voltages generated by the PSO algorithm, which are applied to the distributed regulators. With no timing violations, the assigned voltage(s) mitigate the formation of thermal hotspots. As the system ages based on the executed workloads and environmental conditions, a number of re-calibrations of the sensors are performed until there are no further possible modifications (reductions or increases) to the supply voltage at a given frequency.

The circuit schematics of the timing and aging sensors that provide real-time data to the online PSO are shown in Fig. 2. A latched tapped delay line is implemented in a 7 nm FinFET PTM [13] process and is utilized as a timing sensor to characterize the delay of the *PCPs* distributed across the IC by monitoring the binned location of a clock edge, with results provided to the on-line PSO. The local clock signal for the given voltage domain in which the delay line is placed is applied to a buffer chain. The buffer (or bin number) at which the clock edge is captured provides a measure of the local physical and electrical characteristics of the circuit in the vicinity of the delay line. For instance, the clock edges propagate through more buffers when the delay per bin (*Delayperbin*) decreases with an increase in supply voltage or temperature. The number of bins the clock edge propagates through decreases as V_{th} increases due to process variation and aging. The captured bin location of the fifth clock edge is provided as input to the PSO.

A technique to predict circuit failure is developed in [11], where, based on the monitoring of the transition of the output

signal of a critical path, the detection of any transitions within the set timing interval of the guard-band is flagged. A signal transition detected in the guard-band interval implies that for the given input to the combinational logic, the critical path has slowed due to circuit aging and is close to generating a timing fault. An aging sensor is embedded into the output latch of each critical path. An output flag of a stability checker is raised when there is an aging violation, where the flag is provided as input to the PSO.

III. Simulated Results

The feasibility of run-time voltage assignment through the PSO is analyzed for cores in a chip multi-processor (CMP). The run time PSO is, however, applicable to any circuit with on-chip distributed voltage regulators. The PSO algorithm is implemented in Verilog. The Verilog model is used in SPICE simulations along with Verilog-A models of the LDOs, timing sensors, aging sensors, and the on-chip power distribution network.

A. Simulation setup

A four core chip multi-processor (CMP), which includes a Verilog model of the PSO, is simulated in SPICE. The architectural parameters of the CMP are listed in Table I. A set of 20 applications from the SPEC CPU2006 benchmark suite is used to determine the per cycle power dissipation [14] of a core implemented in a 7 nm FinFET PTM process [13]. The power traces are represented as piece-wise linear waveforms and are used to drive the current loads on the on-chip power distribution network for SPICE simulation. The current loads are connected to a circuit model of the (PDN), with each branch of the PDN designed as a series connection of a resistor and inductor. A total of 5 µF of distributed on-chip capacitance is implemented. The 16 Verilog-A models of the LDO are distributed across the PDN such that there are four LDOs per core. A Verilog model of the latched tapped delay line is developed that accounts for the variation in delay due to changes in temperature, V_{DD}, and V_{th}. A total of 16 delay lines are distributed in close proximity to the LDO models. The model of the circuit used for SPICE simulation is shown in Fig. 3. The inputs to the PSO block include the clock edge locations captured by the n timing sensors each with m bit precision and the set of stability checker flags from each aging sensor θ. The variation in V_{th} for each core is determined using VARIUS [15], while assuming a coefficient of variation (σ/μ ratio) of 0.9% and a spatial correlation parameter ϕ of 0.2. The map of the variation in V_{th} across the four core CMP, as shown in Fig. 4, is provided as input to the PSO model developed in MATLAB.

B. Voltage assignment in the presence of power supply noise and process variation

Simulations are performed to characterize the execution of the PSO algorithm on a circuit with spatial and temporal variations in the power supply voltage and threshold voltage. Results from simulation of a voltage domain with 16 OCVRs are shown in Fig. 5. The voltage domain consists of transistors fabricated in a 7 nm process that operate at a nominal voltage

TABLE I: Architectural parameters of each core of a four core CMP.

Parameter	Value
Core clock frequency	4 GHz
Power supply voltage (V_{dd})	0.7 V
Issue, commit width	2
INT and FP instruction queue	16 entries
Load and Store Queue	16 entries
INT and FP Register File	48 entries
ROB size	48
L1 cache	32KB, 4-way
L2 cache	256KB

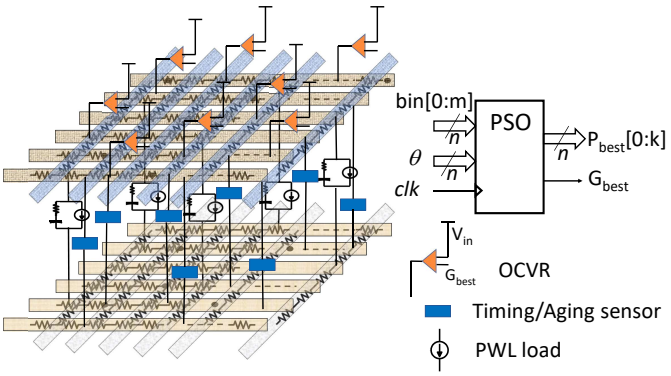


Fig. 3: Circuit model of a voltage domain used for SPICE simulation that includes a Verilog-A circuit implementation of the PSO algorithm that assigns voltages to distributed OCVRs.

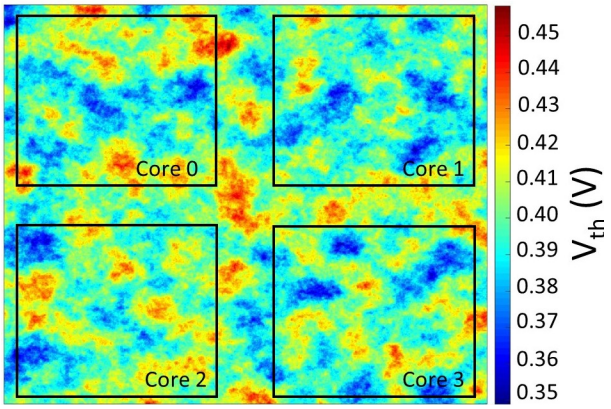
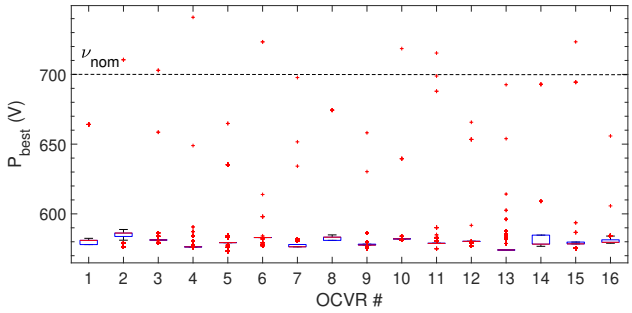
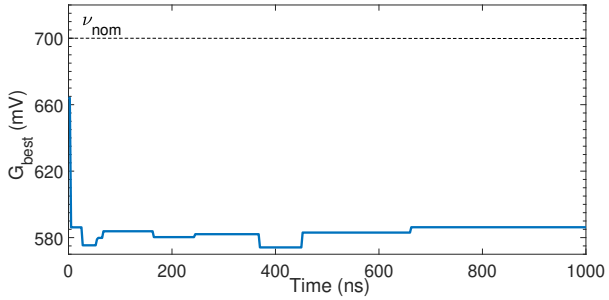


Fig. 4: Floorplan of a four core CMP implemented in a 7 nm PTM technology overlaid on a map of 0.9% σ/μ variation in V_{th}.

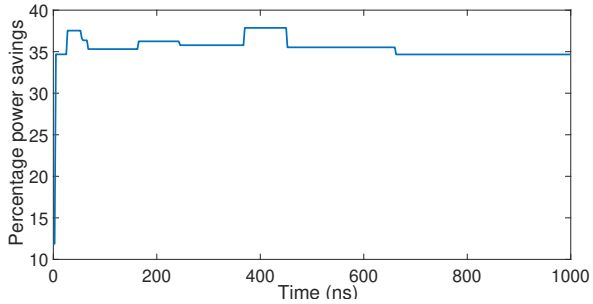
ν_{nom} of 0.7 V. The nominal *Delayperbin*, simulated at a V_{DD} of ν_{nom} and temperature of 25°C is 2.02 ps. A 20 ps timing margin is added to account for the latency of the aging sensor that monitors the timing of a critical path, where the initial delay of the paths at the start of life is 200 ps. A σ/μ ratio of 0.9% is assumed across the voltage domain to characterize process related variation. The clock frequency is set to 4 GHz, and a uniform random power supply noise of 10% of V_{DD} is applied at a temperature of 25°C. The variation in the P_{best} values computed by the PSO for the 16 OCVRs for a time of execution of 1 µs is shown in Fig. 5(a). The evolving G_{best} value computed by the PSO for the given voltage domain is shown in Fig. 5(b). Even when accounting for voltage noise of



(a)



(b)



(c)

Fig. 5: Results from characterization of the PSO algorithm that assigns power supply voltages to 16 OCVRs implemented in a single voltage domain. The nominal voltage ν_{nom} for the transistors is 0.7 V. Results of the characterization include (a) the best voltage assignment per OCVR, (b) the global best voltage with time, and (c) the percentage reduction in power consumption when applying the adaptive global voltage to the given voltage domain as compared to an assignment of ν_{nom}.

$\pm 5\%$ of a ν_{nom} of 0.7 V, the PSO converges to a G_{best} value of 0.59 V without any timing violations in the critical path(s). The percentage reduction in the combined dynamic and static power consumption when applying the adaptive global G_{best} voltage assigned to the domain as compared to an assignment of ν_{nom} is shown in Fig. 5(c).

C. Improved transistor aging characteristics

The aging induced degradation of the threshold voltage V_{th} of the four core CMP is analyzed. A circuit topology that includes four off-chip voltage regulators supplying current to each of the four cores is considered as the baseline for comparison with the proposed technique implementing the run-time voltage assignment of sixteen distributed OCVRs through the PSO algorithm. The selected aging model is validated through wafer measurements on a sub-20 nm FinFET-based ring oscillator [16]. The process and aging induced shift in the threshold voltage ΔV_{th}(t) is mathematically expressed as given by (3). The ΔV_{th}(t) has a normal distribution with a mean $\langle \Delta V_{thA}(t) \rangle$, which is the average shift in the threshold voltage attributed to bias temperature instability (BTI) and is expressed by the power law given by (4). The technology and fitting parameters, A, κ, α, β, and γ in (4), are taken from prior work described in [17] and are provided in Table II. The environmental and physical parameters required to calculate (4) include the total stress time in seconds t, the temperature θ in Kelvin (K), the duty factor df of the applied stress signal (the executing workload), and the electric field across the gate oxide E_{OX} in V/m. The time dependant variance in the threshold voltage $\sigma^2_{\Delta V_{th}}(t)$ due to process variation and aging is given by (5), where $\sigma^2_{\Delta V_{th0}}$ is the variance due to process variation at the beginning of life of the CMP. The CMP is assumed to consist of an equal number of PMOS and NMOS transistors. The process induced variation in V_{th} is assumed identical for the baseline and the PSO based CMP throughout the lifetime of the circuit.

TABLE II: Technology and power-law fitting parameters used in the aging model [17].

Parameter	Description	NMOS	PMOS
γ	power-law exponent	5.2	3
α	power-law exponent	0.158	0.173
A	fitting constant	3.12e-2	2.02e-2
β	fitting constant	0.667	0.667
κ	fitting constant	50	50

$$\Delta V_{th}(t) = \mathcal{N}(\langle \Delta V_{thA}(t) \rangle, \sigma_{\Delta V_{th}}(t)) \tag{3}$$

$$\langle \Delta V_{thA}(t) \rangle \cong A e^{-\kappa/\theta} t^{\alpha} E_{OX}^{\gamma} df^{\beta} \tag{4}$$

$$\sigma^2_{\Delta V_{th}}(t) = \left(1 + \frac{\langle \Delta V_{thA}(t) \rangle}{0.1\ V}\right) \sigma^2_{\Delta V_{th0}} \tag{5}$$

The distributed OCVRs are implemented as LDOs. The aging of the PMOS header of the LDO is considered when determining the effect on V_{th} due to the aging of the CMP. As the output voltage from the LDOs is modulated by the PSO, variations occur in both the operating temperature and the electric field across the gate oxide (E_{OX}) of the load circuits, which includes the distributed timing sensors. The updated temperatures across the CMP are determined using Hotspot. The rate of aging for the baseline CMP and the CMP with PSO voltage assignment is calculated using the model given by (4) and (5) for the same stress time t and duty factor df. The variation in V_{th} at the start of life and at an end of life (EOL) of 10 years for both the baseline CMP and a CMP with adaptive voltage assignment by the PSO are shown in Fig. 6. The analysis considers applying the G_{best} reference voltage to the 16 OCVRs of the four core CMP. The cumulative effect with time of the adaptive power supply voltage significantly reduces the rate of transistor aging, with a mean reduction in $\langle \Delta V_{thA}(t) \rangle$ of 38%. The improved aging characteristics of the circuit is due to the reduction in the magnitude of the applied electric field across the oxide E_{OX} of the transistors as compared to the baseline CMP. The degradation in V_{th}

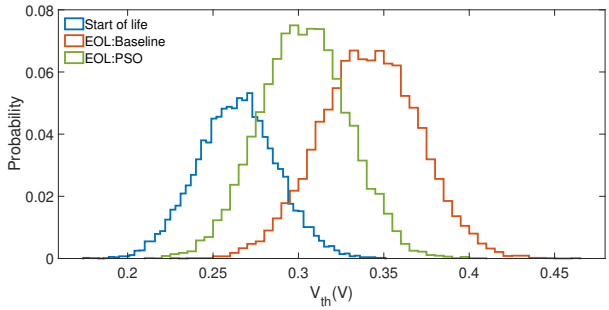


Fig. 6: Reduction in the aging induced degradation of V_{th} through the adaptive voltage assignment of 16 OCVRs integrated in a four core CMP in a low-V_{th} 7 nm FinFET process. The end of life (EOL) of the transistors is ten years.

TABLE III: Circuit and computation overhead to implement the run-time PSO.

Parameter	Value
On-chip memory for LUT	512 B
Execution time of the *PSO* procedure	20 clock cycles per OCVR
Aging sensors [11] in 12 *PCPs*	2K transistors
16 latched tapped delay lines [12]	6K transistors

due to aging is reduced as shown for low-V_{th} 7 nm FinFET transistors in Fig. 6.

D. Hardware overhead to implement the run-time PSO

The overhead in area and compute cycles of implementing the on-line PSO algorithm that adaptively assigns voltages to on-chip regulators is estimated for the circuit shown in Fig 3, with the costs summarized in Table III. The circuit and computational overheads to implement and execute the PSO are determined for a voltage domain with 16 distributed OCVRs supporting six distinct reference voltage levels, 16 delay line based timing sensors, and 12 *PCPs*. The computation time to determine a new P_{best} and G_{best} for each OCVR is twenty clock cycles. The required size of the LUT increases with the operating age of the IC. As a means to reduce the size of the LUT, the allocated memory is reused after every two years of storing the calibration data. The fastest degradation in the threshold voltage of the OCVRs and load circuits occurs in the first two years of the operating life of the IC. Beyond the first two years, the operating temperature becomes a more critical parameter than the total stress time of the load circuit. Through the adaptive voltage assignment of the PSO, the reductions in power supply voltage from the design time nominal value occur less frequently near the end of life of the IC. Therefore, the calibration results during the start of life of the IC do not need to be retained in the LUT.

IV. Conclusion

An evolving voltage assignment to distributed on-chip voltage regulators is proposed. The reference voltage to each regulator is obtained through an evolutionary algorithm, specifically a particle swarm optimizer. The distributed voltage regulators serving a voltage domain operate as a swarm and utilize the local data collected by timing and aging sensors to compensate for the effects of process variation, transistor aging in both the load circuit and the voltage regulators, and variations in the temperature across the die. The variation in the delay of the timing sensors is dependent on process, aging, temperature, and power supply noise, which provides optimal properties that allow for the monitoring of the reliability and performance of the circuit. Through simulation on a circuit implementing the PSO, OCVRs, and integrated timing and aging sensors in a 7 nm FinFET technology, an average reduction of 35% in the power consumption of the voltage domains was observed. In addition, the end of life of the circuit increases due to an average reduction of 38% in the aging induced variation in V_{th}.

References

[1] D. Pathak, M. Hajkazemi, M. Tavana, H. Homayoun, and I. Savidis, "Energy Efficient On-Chip Power Delivery with Run-Time Voltage Regulator Clustering," *Proceedings of the IEEE International Symposium on Circuits and Systems*, pp. 1210–1213, May 2016.

[2] Y. Zu, C. R. Lefurgy, J. Leng, M. Halpern, M. S. Floyd, and V. J. Reddi, "Adaptive guardband scheduling to improve system-level efficiency of the POWER7+," *Proceedings of the International Symposium on Microarchitecture*, pp. 308–321, December 2015.

[3] W. Lee, Y. Wang, T. Cui, S. Nazarian, and M. Pedram, "Dynamic thermal management for FinFET-based circuits exploiting the temperature effect inversion phenomenon," *Proceedings of the International Symposium on Low Power Electronics and Design*, pp. 105–110, August 2014.

[4] D. Pathak and I. Savidis, "On-chip power supply noise suppression through hyperabrupt junction varactors," *IEEE Transactions on Very Large Scale Integration Systems*, Vol. 26, No. 11, pp. 2230–2240, November 2018.

[5] Sanjay Pant, *Design and analysis of power distribution networks in VLSI circuits*, Ph.D. Thesis, University of Michigan, 2008.

[6] W. H. Chang, C. H. Lin, S. P. Mu, L. D. Chen, C. H. Tsai, Y. C. Chiu, and M. Chao, "Generating routing-driven power distribution networks with machine-learning technique," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 36, No. 8, pp. 1237–1250, January 2017.

[7] S.C. Chapra and R.P. Canale, *Numerical methods for engineers*, Boston: McGraw-Hill Higher Education, 2010.

[8] A. P. Da Silva and P. Abrão, "Applications of evolutionary computation in electric power systems," *Proceedings of the Congress on Evolutionary Computation*, Vol. 2, pp. 1057–1062, May 2002.

[9] T. Susa, M. Murakawa, E. Takahashi, T. Furuya, and T. Higuchi, "Post-fabrication clock-timing adjustment for digital LSIs ensuring operational timing margins," *Proceedings of the Hybrid Intelligent Systems conference*, pp. 907–910, March 2008.

[10] D. W. Boeringer and D. H. Werner, "Particle swarm optimization versus genetic algorithms for phased array synthesis," *IEEE Transactions on Antennas and Propagation*, Vol. 52, No. 3, pp. 771–779, April 2004.

[11] M. Agarwal, B. C. Paul, M. Zhang, and S. Mitra, "Circuit failure prediction and its application to transistor aging," *Proceedings of the IEEE VLSI Test Symposium*, pp. 277–286, May 2007.

[12] R. Franch, P. Restle, N. James, W. Huott, J. Friedrich, R. Dixon, S. Weitzel, K. Van Goor, and G. Salem, "On-chip timing uncertainty measurements on IBM microprocessors," *Proceedings of the IEEE International Test Conference*, pp. 1–7, October 2007.

[13] Nanoscale Integration & Modeling Group at Arizona State University, "Predictive Technology Model (PTM)," http://ptm.asu.edu/.

[14] D. Pathak, H. Homayoun, and I. Savidis, "SMART GRID ON CHIP: Work load balanced on-chip power delivery," *IEEE Transactions on Very Large Scale Integration Systems*, Vol. 25, No. 9, pp. 2538–2551, September 2017.

[15] 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 Transactions on Semiconductor Manufacturing*, Vol. 21, No. 1, pp. 3–13, February 2008.

[16] P. Weckx, B. Kaczer, M. Toledano-Luque, P. Raghavan, J. Franco, P. J. Roussel, G. Groeseneken, and F. Catthoor, "Implications of BTI-induced time-dependent statistics on yield estimation of digital circuits," *IEEE Transactions on Electron Devices*, Vol. 61, No. 3, pp. 666–673, January 2014.

[17] E. Cai, D. Stamoulis, and D. Marculescu, "Exploring aging deceleration in FinFET-based multi-core systems," *Proceedings of the International Conference on Computer-Aided Design*, pp. 111:1–111:8, November 2016.