Exploiting Algorithmic Noise Tolerance for Scalable On-Chip Voltage Regulation

Longfei Wang, S. Karen Khatamifard, Ulya R. Karpuzcu, and Selçuk Köse, *Member, IEEE*

***Abstract*— With the advent of on-chip digital low-dropout (DLDO) regulators, distributed on-chip voltage regulation has become increasingly promising. Environmental and operating conditions have been demonstrated to degrade DLDO performance, which directly affects execution accuracy. The area overhead (OH) needed to compensate aging-induced voltage noise degradation can be significant. Accordingly, in this paper, the algorithmic noise tolerance of certain processor components is exploited as an area–quality control knob to trade the program output quality for area OH. Furthermore, efficient and lightweight techniques utilizing a unidirectional shift register and reduced clock pulsewidth triggering are proposed to realize a novel aging-aware (AA) DLDO to achieve a better area and quality tradeoff. Owing to the large number and distributed nature of voltage regulators, with the proposed design, both the number of regulators utilized in the system and the size of each local regulator are scalable to satisfy the needs of different applications and processor components with varying algorithmic noise tolerance. It is demonstrated through simulation of an IBM POWER8 like processor that the proposed AA design can achieve up to, respectively, 43.2% and 3× transient and steady-state performance improvement. Additionally, more than 10% area OH saving can be achieved over a 5-year period.**

***Index Terms*— Algorithmic noise tolerance, digital low-dropout (DLDO) regulator, reliability, scalable on-chip voltage regulation, voltage noise.**

I. Introduction

DISTRIBUTED on-chip voltage regulation [1], [2] in fine temporal and spatial granularity enables fast and timely control of the operating point. Thereby, the operating voltage and frequency can better match the needs of the workload to maximize energy efficiency. As a function of the workload, throughout the execution time, different components of a processor chip exhibit different microarchitectural activities, which translates into different demands for current to be pulled from the respective regulators. Different components of the processor chip also show different degrees of tolerance to errors, which may result from deviation of design parameters from their target values due to device wearout, voltage noise, temperature, or process variations. For example, it has been observed that the emerging recognition, mining, and synthesis [3] applications can tolerate errors in the data flow but not in control [4].

Heterogeneous distributed on-chip voltage regulation has been explored to best capture spatiotemporal variations in current demand of different processor components, where the regulator operating regimes are tailored to the activity range of the respective load (processor component). Such tailoring can be achieved by: 1) keeping the regulator design constant across chip but making each regulator reconfigurable or 2) by designing each regulator from the groundup to match different load conditions. A promising direction that has not been explored is how such heterogeneity can help in trading the program output quality for area overhead (OH) by, e.g., assigning error-prone (i.e., slower and/or less accurate) regulators to feed processor components in charge of data flow which can tolerate errors. Control heavy components, on the other hand, should not be permitted to leave the error-free zone to avoid catastrophic program termination or excessive loss in program output quality even if the program does not crash.

To this end, we must understand the type and impact of errors that voltage regulators can introduce to the system, such that we can assess to what extent such regulator-induced errors can be masked by their respective loads (i.e., data flow heavy processor components) and how regulator-induced errors interact with load-induced potential errors in determining the final computation accuracy. In this paper, we will try to shed light into this question by quantifying the impact of one of the most prevalent reliability concerns, aging, on regulator robustness, without loss of generality.

The major transistor aging mechanisms include bias temperature instability (BTI), hot carrier injection, time-dependent dielectric breakdown, and electromigration, among which BTI is the dominant reliability concern for nanometer integrated circuits design [5]–[7]. BTI can induce threshold voltage increase and consequent circuit-level performance degradation. Positive BTI (PBTI) induces aging of nMOS transistors while negative BTI (NBTI) causes aging of pMOS transistors [6]. The impact of BTI aging mechanism is a strong function of temperature, electrical stress, and time.

As an essential part of a processor chip, on-chip voltage regulators need to be active most of the time to provide the

Manuscript received April 11, 2018; revised July 19, 2018; accepted August 27, 2018. Date of publication November 9, 2018; date of current version December 28, 2018. This work was supported in part by the National Science Foundation CAREER Award under Grant CCF-1350451, in part by the National Science Foundation Award under Grant CCF-1421988, and in part by the Cisco Research Award. *(Corresponding author: Longfei Wang.)*

L. Wang and S. Köse are with the Department of Electrical Engineering, University of South Florida, Tampa, FL 33620 USA (e-mail: longfei@mail.usf.edu; kose@usf.edu).

S. K. Khatamifard and U. R. Karpuzcu are with the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN 55455 USA (e-mail: khata006@umn.edu; ukarpuzc@umn.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TVLSI.2018.2871381

required power to different components of the processor chip. The load current and temperature can vary significantly among these components for different processor applications [8]. All of these variations partially contribute to different aging mechanisms of on-chip voltage regulators, which should be considered to avoid overdesign for a targeted lifetime. Additionally, in certain processor components that can show higher degrees of tolerance to errors, the regulators can be intentionally underdesigned to save valuable chip area and potentially power-conversion efficiency. In other words, a heterogeneous distributed power delivery network can be designed consisting of different voltage regulators with accurate voltage regulators that house additional circuitry to mitigate the aging-induced supply voltage variations and approximate voltage regulators that are intentionally underdesigned to mitigate just enough aging-induced variations. The quality of the supply voltage directly affects the data path delay and signal quality, and fluctuations in the supply voltage result in delay uncertainty and clock jitter. Accordingly, the supply noise tolerance of certain processor components is investigated as an area quality control knob where the quality of the supply voltage can be compromised to save valuable chip area.

Several studies have been performed regarding the reliability issues in nanometer CMOS designs [9]–[11]. There is, however, quite limited amount of work on the reliability of on-chip voltage regulators. Device aging on the immunity level of electromagnetic interference for low-dropout (LDO) regulators is characterized in [12]. A method of distributing the aging stress by rotating the phase to shed at light load is proposed in [13] to enhance the light-load efficiency for multiphase buck converters. An algorithm to uniformly distribute the current provided by the power transistor array of a digital LDO (DLDO) is proposed in [14] to reduce hot spots and ensure reliable silicon operation. The reliability of metal wires connected to on-chip voltage regulators is investigated in [15]. Nonetheless, the quantitative analysis of aging effects on on-chip voltage regulators considering load current characteristics and temperature variations as well as efficient reliability enhancement techniques under arbitrary load conditions have not yet been investigated.

As compared to other voltage regulator types, the emerging DLDO has gained impetus due to the design simplicity, easiness for integration, high power density, and fast response [16]–[18]. DLDOs have demonstrated major advantages in modern processors including the recent IBM POWER8 processor [19]. More importantly, as compared to the analog LDOs, DLDO can provide certain advantages for low-power and low-voltage IoT applications due to its capability for low supply voltage operations [20]. However, as pMOS is used as the power transistor for DLDOs, NBTI-induced degradations largely affect important performance metrics such as the maximum output current capability $I_{\max}$, load response time T_R, and magnitude of the droop ΔV as defined in [21]. Meanwhile, the combined NBTI- and PBTI-induced control loop degradations can potentially increase the mode of inherent limit cycle oscillations (LCOs) within DLDOs and adversely affect the steady-state output voltage ripple performance. It is, therefore, imperative to investigate aging mitigation techniques for DLDOs to achieve reliable operation of critical components. Alternatively, when a circuit component can tolerate higher degrees of errors, the DLDOs can be designed with minimal area OH, achieving heterogeneous power delivery. A voltage regulator is proposed in this paper that can be designed at the design time based on the supply noise resiliency requirement of the circuitry it powers. Since the number of voltage regulators can be as high as several hundred in modern processors [19], the area and number of voltage regulators can be easily scaled, thereby to satisfy the diverse needs of systems that house components with varying degrees of noise tolerance.

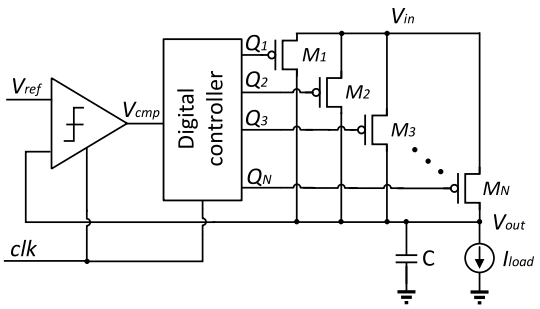


Fig. 1. Schematic of a conventional DLDO.

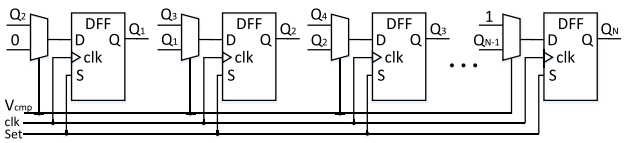
The rest of this paper is organized as follows. Background information regarding conventional DLDO regulator and BTI is introduced in Section II. BTI-induced DLDO performance degradation including $I_{\max}$, T_R, ΔV, and mode of LCOs is demonstrated theoretically in Section III. The proposed aging-aware (AA) DLDO is described in Section IV. Benefits evaluation of the proposed AA DLDO through simulation of an IBM POWER8 like processor is provided in Section V. Tradeoff between the area OH of voltage regulators and program output quality is detailed in Section VI. Concluding remarks are offered in Section VII.

II. Background

A. Conventional DLDO Regulator

The schematic of a conventional DLDO [16] is shown in Fig. 1. DLDO is composed of N parallel pMOS transistors M_i $(i = 1, \ldots, N)$ connected between the input voltage V_{in} and output voltage V_{out}, and feedback control loop implemented with a clocked comparator and digital controller. The value of V_{out} and reference voltage V_{ref} are compared through the comparator at the rising edge of the clock signal clk. More (less) number of M_i is turned ON through the digital controller output signals Q_i $(i = 1, \ldots, N)$ if $V_{\text{out}} < V_{\text{ref}}$, $V_{\text{cmp}} = H$ ($V_{\text{out}} > V_{\text{ref}}$, $V_{\text{cmp}} = L$). A bidirectional shift register (bDSR), as shown in Fig. 2(a), is conventionally implemented for the digital controller to turn ON (OFF) power transistors M_1 to M_m (M_{m+1} to M_N) with the value of m decided by the load current I_{load}. At a certain step $k+1$, M_{m+1} (M_m) is turned ON (OFF) if $V_{\text{cmp}} = H$ ($V_{\text{cmp}} = L$) and bDSR shifts right (left) as demonstrated in Fig. 2(b).

DLDO needs to be able to supply the maximum possible load current $I_{\max}$. It is, however, demonstrated that within



(a)

Q_1	Q_2	Q_3	Q_4	Q_5	Q_6	· · ·	· · ·	Q_{N-1}	Q_N
(1) Initialize: all M_i turned off									
1	1	1	1	1	1	· · ·	· · ·	1	1
(2) Step k									
0	0	0	0	1	1	· · ·	· · ·	1	1
(3-a) Step k+1 if V_{cmp}=H: Shift right ➡									
0	0	0	0	0	1	· · ·	· · ·	1	1
(3-b) Step k+1 if V_{cmp}=L: Shift left ⬅									
0	0	0	1	1	1	· · ·	· · ·	1	1

(b)

Fig. 2. Digital controller for conventional DLDO. (a) bDSR. (b) Operation of bDSR.

most practical applications, including but not limited to processors [22], [23] and smartphones [13], less than the average power is consumed most of the time. The application environment of DLDO together with the conventional activation scheme of M_i leads to the heavy use of M_1 to M_m and less or even no use of M_{m+1} to M_N. This scheme can, therefore, introduce serious degradation to M_1 to M_m due to NBTI. The BTI mechanism and the subsequent DLDO performance deteriorations due to degradation of M_is are discussed in Sections II-B and III-A–III-C, respectively. Meanwhile, the error tolerance capability of different functional blocks can be different, which necessitates area–quality tradeoff for aging mitigation-induced area OH as will be discussed in Section VI.

Furthermore, DLDOs experience LCOs in steady state due to inherent quantization errors [16], [24], [25]. The number of power transistor M that is periodically turned ON or OFF in steady state is the mode of LCO. A larger LCO mode M under a certain load current I_{load} and clock frequency f_{clk} condition may lead to larger steady-state output voltage ripple as demonstrated in [24] and [25]. Larger delay between the clocked comparator and shift register is shown to be detrimental to LCOs [26]. The BTI-induced control loop degradation can potentially further exacerbate the LCO mode as will be explored in Section III-D.

B. Bias Temperature Instability

NBTI (PBTI) can introduce significant V_{th} degradations to pMOS (nMOS) transistors due to negatively (positively) applied gate-to-source voltage V_{gs}. The increase in $|V_{\text{th}}|$ due to BTI is considered to be related to the generation of interface traps at the Si/SiO_2 interface when there is a gate voltage [27]. $|V_{\text{th}}|$ increases when electrical stress is applied and partially recovers when stress is removed. This process is commonly explained using a reaction–diffusion model [27]. The V_{th} degradation can be estimated during each stress and recovery phase using a cycle-to-cycle model and can also be evaluated using a long-term reliability model [6], [10], [28]. As the long-term reliability evaluation is the focus of this paper, the analytical model for long-term worst case threshold voltage degradation ΔV_{th} estimation in [6] is adopted in this paper as

$$\Delta V_{\text{th}} = \chi K_{\text{lt}} \sqrt{C_{\text{ox}}(|V_{\text{gs}}| - |V_{\text{th}}|)} e^{\frac{-E_a}{kT}} (\alpha t)^{\frac{1}{6}} \tag{1}$$

where C_{ox}, k, T, α, and t are, respectively, the oxide capacitance, Boltzmann's constant, temperature, the fraction of time (activity factor) when the device is under stress, and the operation time. K_{lt} and E_a are the fitting parameters to match the model with the experimental data [6]. $\chi = 1$ ($\chi = 0.5$) for NBTI (PBTI) induced ΔV_{th}. Note that BTI recovery phase is already included in the model.

III. Aging-Induced DLDO Performance Degradation

$I_{\max}$, T_R, ΔV, and mode of LCO are among the most important design parameters for DLDOs. The effect of aging-induced degradations on these important performance metrics is examined in this section.

A. Maximum Current Supply Capability

Without NBTI-induced degradations, $I_{\max} = N I_{\text{pMOS}}$, where I_{pMOS} is the maximum output current of a single pMOS stage. For DLDO, $|V_{\text{gs}}|$ in (1) is equal to V_{in} when M_i is active. The pMOS transistor M_i operates in the linear region when turned ON and the ON-resistance R_{ON} of a single pMOS stage can be approximated as [6]

$$R_{\text{ON}} \approx [(W/L)\mu_p C_{\text{ox}}(V_{\text{in}} - |V_{\text{th}}|)]^{-1} \tag{2}$$

where W, L, μ_p, and C_{ox} are, respectively, the width, length, mobility, and oxide capacitance of M_i. I_{pMOS} can thus be expressed as

$$I_{\text{pMOS}} = \frac{V_{\text{sd}}}{R_{\text{ON}}} = (V_{\text{in}} - V_{\text{out}})(W/L)\mu_p C_{\text{ox}}(V_{\text{in}} - |V_{\text{th}}|) \tag{3}$$

where V_{sd} is the source–drain voltage of M_i. NBTI-induced degradation factor DF_i for M_i can be defined as

$$DF_i = \frac{I^{\text{deg}}_{\text{pMOS}_i}}{I_{\text{pMOS}}} = \frac{V_{\text{in}} - |V_{\text{th}}| - \Delta V_{\text{th}_i}}{V_{\text{in}} - |V_{\text{th}}|} \tag{4}$$

where ΔV_{th_i} and $I^{\text{deg}}_{\text{pMOS}_i}$ are, respectively, the NBTI-induced V_{th} degradation and the degraded I_{pMOS} for M_i. Degraded $I_{\max}$ can be expressed as

$$I^{\text{deg}}_{\max} = I_{\text{pMOS}} \sum_{i=1}^{N} DF_i. \tag{5}$$

As an example, the percentage I_{pMOS} degradation $1 - DF_i$ for a smaller value of i, considering M_i is active most of the time, is shown in Fig. 3 as a function of time under different temperatures. Equations (1) and (4) are leveraged for evaluation, where transistor model parameters are adopted from a 32-nm metal gate, high-k strained-Si CMOS technology within the predictive technology model (PTM) model library [28]. A supply voltage $V_{\text{in}} = 1.1$ V is used for estimation. PTM is adopted for the aging-induced deterioration analysis and subsequent DLDO simulations as it is widely used for BTI

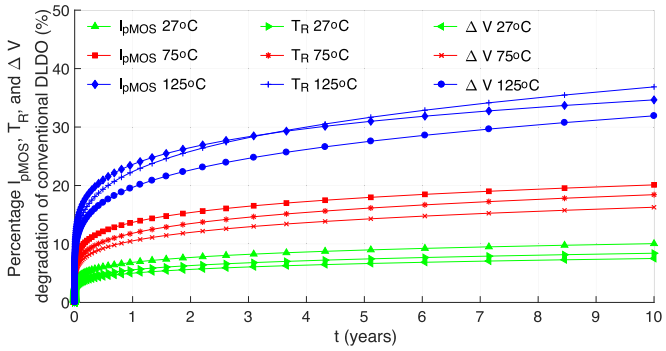


Fig. 3. Percentage I_{pMOS}, T_R, and ΔV degradation of conventional bDSR-based DLDO.

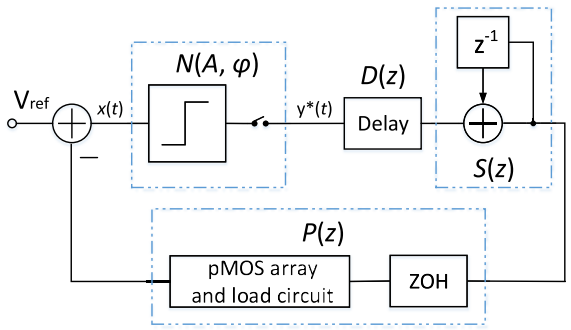


Fig. 4. Nonlinear sampled feedback model of a conventional DLDO.

study due to the availability of fitting parameter values in the ΔV_{th} degradation model [5], [6], [9]–[11]. As shown in Fig. 3, NBTI can induce significant I_{pMOS} degradations, especially at high temperatures. Also, most degradation occurs in the first 2 years. Beyond 2 years, the degradation typically plateaus to within 10%. Degraded I_{pMOS} can further lead to reduced $I_{\max}$ and lower output voltage regulation capability under high load current. Moreover, as discussed in Sections III-B and III-C, degraded I_{pMOS} also exacerbates T_R and ΔV, necessitating reliability enhancement techniques.

B. Load Response Time

Load response time T_R measures how fast the feedback loop responds to a step load. T_R can be estimated as [29]

$$T_R = RC \ln\left(1 + \frac{\Delta i_{\text{load}}}{I_{\text{pMOS}} f_{\text{clk}} RC}\right) \tag{6}$$

where R, C, and Δi_{load} are, respectively, the average DLDO output resistance before and after Δi_{load}, load capacitance, and amplitude of the load change. Considering NBTI effect, degraded T_R can be expressed as

$$T_R^{\text{deg}} = RC \ln\left(1 + \frac{\Delta i_{\text{load}}}{DF I_{\text{pMOS}} f_{\text{clk}} RC}\right). \tag{7}$$

As $0 < DF < 1$ and $T_R < T_R^{\text{deg}}$, NBTI-induced degradation slows down DLDO response. Based on the DLDO design specifications that will be discussed in Section V-A2, worst case T_R degradation is illustrated in Fig. 3.

C. Magnitude of the Droop

Magnitude of the droop ΔV reflects the V_{out} noise profile under transient response and can be estimated as [29]

$$\Delta V = R\Delta i_{\text{load}} - I_{\text{pMOS}} f_{\text{clk}} R^2 C \ln\left(1 + \frac{\Delta i_{\text{load}}}{I_{\text{pMOS}} f_{\text{clk}} RC}\right). \tag{8}$$

Considering NBTI effect, degraded ΔV can be expressed as

$$\Delta V_{\text{deg}} = R\Delta i_{\text{load}} - DFI_{\text{pMOS}} f_{\text{clk}} R^2 C \ln\left(1 + \frac{\Delta i_{\text{load}}}{DFI_{\text{pMOS}} f_{\text{clk}} RC}\right). \tag{9}$$

Let $\Delta i_{\text{load}}/I_{\text{pMOS}} f_{\text{clk}} RC = A$, $A > 0$. Under $0 < DF < 1$, the following holds:

$$1 + A > \left(1 + \frac{A}{DF}\right)^{DF} \tag{10}$$

thus

$$I_{\text{pMOS}} f_{\text{clk}} R^2 C \ln\left(1 + \frac{\Delta i_{\text{load}}}{I_{\text{pMOS}} f_{\text{clk}} RC}\right) > DFI_{\text{pMOS}} f_{\text{clk}} R^2 C \ln\left(1 + \frac{\Delta i_{\text{load}}}{DFI_{\text{pMOS}} f_{\text{clk}} RC}\right) \tag{11}$$

and $\Delta V < \Delta V_{\text{deg}}$, which means NBTI can degrade the transient voltage noise profile. Worst case ΔV degradation based on the DLDO design specifications in Section V-A2 is illustrated in Fig. 3.

D. Limit Cycle Oscillation

In conventional DLDOs, when the shift register turns ON/OFF the pass transistor, the output voltage of the DLDO cannot change instantaneously due to the output pole of the DLDO. The delay between the operation of the shift register and fluctuation of the output voltage, together with the quantization effects of the comparator and the delay between the sampling instant and the time of pMOS array actuation lead to the occurrence of LCO. Such behavior can be examined by a nonlinear sampled feedback model developed in [24] to determine the possible modes and amplitudes of LCOs.

The model consists of $N(A, \varphi)$, $P(z)$, $S(z)$, and $D(z)$ as shown in Fig. 4, which represent, respectively, the describing function of the clocked comparator, transfer function of the zero-order hold together with the pMOS array and load circuit, transfer function of the shift register, and delay element between the comparator and shift register. A and φ stand for the LCO amplitude and the phase shift of $x(t)$, respectively.

$N(A, \varphi)$, $P(z)$, $S(z)$, and $D(z)$ can be expressed, respectively, as [25], [26]

$$N(A, \varphi) = \frac{2D}{MTA} \sum_{m=0}^{M-1} \sin\left(\frac{\pi}{2M} + \frac{m\pi}{M}\right) \angle\left(\frac{\pi}{2M} - \varphi\right) \tag{12}$$

$$P(z) = K_{\text{OUT}} \frac{1 - e^{-F_l T}}{F_l (z - e^{-F_l T})} \tag{13}$$

$$S(z) = \frac{z}{z - 1} \tag{14}$$

$$D(z) = z^{-1} \tag{15}$$

where $K_{\rm OUT} = K_{\rm dc} I_{\rm pMOS}$, $T = 1/f_{\rm clk}$, $F_l = 1/(R_L || R_{\rm pMOS})C$, and $\varphi \in (0, \pi/M)$. D, F_l, $K_{\rm OUT}$, $K_{\rm dc}$, R_L, and $R_{\rm pMOS}$ are, respectively, the amplitude of comparator output, load pole, gain of $P(z)$, direct current (dc) proportional constant, load resistance, and resistance of power transistor array.

The mode and amplitude of LCO can be determined by the following Nyquist criterion:

$$N(A,\varphi)P(e^{j\omega T})S(e^{j\omega T})D(e^{j\omega T}) = 1\angle(-\pi) \tag{16}$$

where $\omega = \pi/TM$ is the angular LCO frequency. The phase shift $\varphi_{\rm LCO}$ for a steady LCO can thus be expressed as [25]

$$\varphi_{\rm LCO} = \frac{\pi}{2} - \frac{\pi}{2M} - \tan^{-1}\left(\frac{\pi}{MTF_l}\right). \tag{17}$$

$\varphi_{\rm LCO}$ needs to be within $(0, \pi/M)$ for mode M to exist.

Transistor aging can lead to increased path delay [30]. Considering BTI-induced propagation delay degradation of the clocked comparator and shift register, the delay element in Fig. 4 becomes

$$D'(z) = z^{-1} z^{-\frac{t_c^d}{T}} z^{-\frac{\left(t_s^d - t_c^d\right)}{T}} = z^{-1-\frac{t_s^d}{T}} \tag{18}$$

where t_c^d and t_s^d are, respectively, the degraded propagation delay of the clocked comparator and shift register. Note that t_c^d is canceled out in $D'(z)$, and thus, the propagation delay of clocked comparator has negligible effects on the mode of LCO. $\varphi_{\rm LCO}$ then becomes

$$\varphi'_{\rm LCO} = \frac{\pi}{2} - \frac{\pi}{2M} - \tan^{-1}\left(\frac{\pi}{MTF_l}\right) - \frac{\pi t_s^d}{MT}. \tag{19}$$

The negative effect of the propagation delay of the shift register on LCO can be explained as follows. If an LCO mode M_a exists and the propagation delay of the shift register is not considered, the phase shift $\varphi_{\rm LCO}$ is within $(0, \pi/M_a)$. That is, $0 < \pi/2 - \pi/2M_a - \tan^{-1}(\pi/M_a TF_l) < \pi/M_a$. For a larger LCO mode $M_a + 1$ to exist, the following condition needs to be satisfied:

$$0 < \frac{\pi}{2} - \frac{\pi}{2(M_a+1)} - \tan^{-1}\left(\frac{\pi}{(M_a+1)TF_l}\right) < \pi/(M_a+1). \tag{20}$$

Typically

$$\frac{\pi}{2} - \frac{\pi}{2(M_a+1)} - \tan^{-1}\left(\frac{\pi}{(M_a+1)TF_l}\right) > \frac{\pi}{2} - \frac{\pi}{2M_a} - \tan^{-1}\left(\frac{\pi}{M_a TF_l}\right) \tag{21}$$

and if $\pi/2 - \pi/2M_a - \tan^{-1}(\pi/M_a TF_l)$ is very close to π/M_a, it is likely that

$$\varphi_{\rm LCO}|_{M=M_a+1} = \frac{\pi}{2} - \frac{\pi}{2(M_a+1)} - \tan^{-1}\left(\frac{\pi}{(M_a+1)TF_l}\right) > \pi/M_a > \pi/(M_a+1) \tag{22}$$

such that LCO mode $M_a + 1$ cannot exist as (20) is violated.

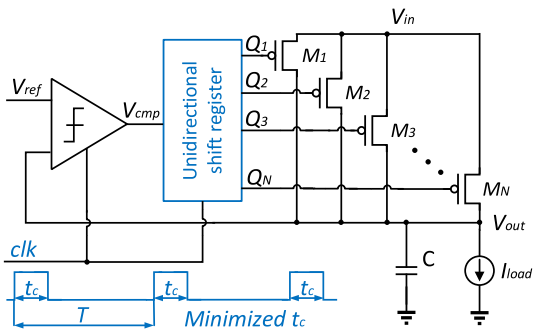


Fig. 5. Schematic of the proposed AA DLDO.

However, if the propagation delay of the shift register is included, for LCO mode $M_a + 1$, $\varphi_{\rm LCO}$ becomes

$$\varphi'_{\rm LCO}|_{M=M_a+1} = \frac{\pi}{2} - \frac{\pi}{2(M_a+1)} - \tan^{-1}\left(\frac{\pi}{(M_a+1)TF_l}\right) - \frac{\pi t_s^d}{(M_a+1)T}. \tag{23}$$

The contribution of $\pi t_s^d/(M_a+1)T$ term may push $\varphi'_{\rm LCO}|_{M=M_a+1}$ to be within the range of $(0, \pi/(M_a+1))$, making a larger LCO mode M_a+1 possible. This demonstrates the potential negative effect of the propagation delay of the shift register on LCO.

Note that aging-induced propagation delay degradation is not a sufficient condition to incite a larger LCO mode. It is, however, as will be discussed in Sections IV and V, due to a small aging-induced shift register delay degradation, the lower boundary of the timing constraint for normal DLDO operation can be significantly smaller than half of the clock cycle such that beneficial effects of the reduced clock pulsewidth scheme can be achieved.

IV. Aging-Aware DLDO Voltage Regulator

Considering the side effects of power transistor array and control loop degradations, an AA DLDO voltage regulator is proposed as shown in Fig. 5. The proposed AA DLDO adopts a unidirectional shift register (uDSR) and reduced clock pulsewidth triggering to mitigate, respectively, $I_{\rm pMOS}$, T_R, and ΔV degradation and LCOs. The proposed uDSR and reduced clock pulsewidth triggering are explained in Sections IV-A and IV-B, respectively. Power and area OH of the proposed techniques as well as compatibility analysis are provided in Section IV-C.

A. Unidirectional Shift Register

To mitigate NBTI-induced $I_{\rm pMOS}$, T_R, and ΔV degradations, distributing the electrical stress among all available power transistors as evenly as possible under arbitrary load current conditions is essential. Reliability is not considered in conventional bDSR-based DLDO designs, and therefore, too much stress is exerted on a small portion of M_is. An uDSR is thus proposed to evenly distribute the electrical stress among all of M_is to realize an AA DLDO voltage regulator and enhance reliability.

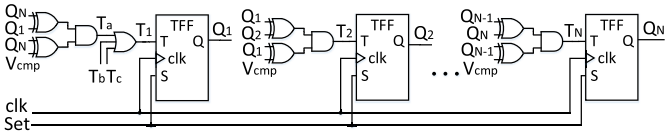


Fig. 6. Proposed uDSR for AA DLDO.

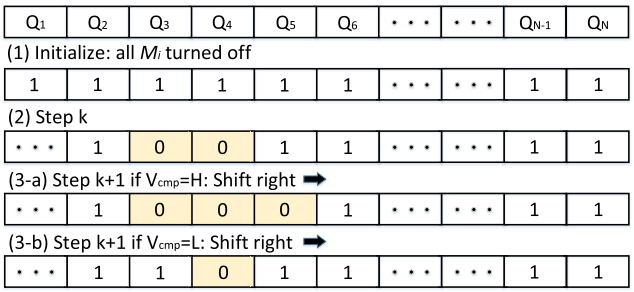


Fig. 7. Operation of the proposed uDSR.

The schematic and operation of the proposed uDSR are shown, respectively, in Figs. 6 and 7. The elementary D flip-flop (DFF) and multiplexer within bDSR, as shown in Fig. 2(a), are replaced with T flip-flop (TFF) and simple logic gates within the proposed uDSR, respectively. The rest of the DLDO including power transistor array and clocked comparator remains unchanged. The idea is to balance the utilization of each available M_i under all load current conditions. To achieve this objective, control signals Q_{i-1} and Q_i for two adjacent power transistors M_{i-1} and M_i, respectively, are XORed to determine if M_{i-1} and M_i are at the boundary of active and inactive power transistor portions. Normally, there are two such boundaries if at least one power transistor is active, as shown in Fig. 7. Q_{i-1} and output of the comparator V_{cmp} are thus XORed to decide which power transistor at the boundaries need to be turned ON/OFF at the rising edge of the clock signal. Inactive (active) power transistor at the right (left) boundary is turned ON (OFF) if V_{cmp} is logic high (low). A uDSR is realized through this activation/deactivation scheme, as demonstrated in Fig. 7. Q_{i-1} for the first stage is Q_N from the last stage, and thus a loop is formed. Considering the initialization step when all M_is are OFF and the full load current condition when all M_is are ON, additional control signals are inserted as T_b and T_c in the first stage, to avoid inaction under these two situations, where $T_b = Q_1 \cdot Q_2 \ldots Q_N \cdot V_{\text{cmp}}$ and $T_c = \overline{Q_1 + Q_2 + \cdots + Q_N + V_{\text{cmp}}}$. The logic functions for T_b and T_c can be implemented with n-input AND/NOR gates [31].

1) Steady-State Operation: Under steady-state conditions, LCO [16] occurs to supply the required current. The number of active power transistors changes dynamically at the rising edge of each clock cycle. Due to LCO, the changing number of active power transistors leads to the flip of control logics and power transistors for both conventional and the proposed DLDO. The number of active/inactive power transistors is the same during each clock cycle for both bDSR and uDSR control if all other simulation settings except the digital controller are the same. The only functional difference between the two controllers is which portion of the power transistor array is active during each clock cycle as illustrated in the following.

For example, Fig. 8 illustrates the different operations of bDSR and uDSR controlled DLDO at steady state with LCO mode $M = 2$ for simplicity. The LCO mode M indicates the number of switching power transistors for conventional bDSR-based DLDO at steady state. Fig. 8(a) demonstrates the operation of a conventional bDSR. Assuming at step k (rising edge of the kth clock cycle) power transistors M_1 and M_2 are active, due to mode 2 LCO and bDSR control (right shift with increasing number of active power transistor and left shift with decreasing number of active power transistor), power transistors M_3 and M_4 become active at, respectively, step $k+1$ and step $k+2$ (rising edge of the $(k+1)$th and $(k+2)$th clock cycle). Power transistors M_4 and M_3 become inactive at, respectively, step $k+3$ and step $k+4$. The subsequent steps will repeat steps $k+1$ to $k+4$. Fig. 8(b) demonstrates the operation of the proposed uDSR. Assuming at step k power transistors M_3 and M_4 are active, due to mode 2 LCO and uDSR control [power transistor is always activated on the right side of active power transistor region and deactivated on the left side of active power transistor region, i.e., the colored region in Fig. 8(b)], power transistors M_5 and M_6 become active at, respectively, step $k+1$ and step $k+2$. Power transistors M_3 and M_4 become inactive at, respectively, step $k+3$ and step $k+4$. The subsequent steps will follow the same activation/deactivation pattern. The location of the colored region dynamically shifts right (unidirectional shift). For a long-term reliability concern, each M_i is active for six clock cycles before it becomes inactive. When power transistor M_N becomes active, the next activated power transistor will be M_1 such that a loop is formed and electrical stress can be more evenly distributed among all of the power transistors as compared to bDSR operation.

The simulated power transistor array gate voltages with both bDSR and uDSR implementation under $I_{\text{load}} = 300$ mA is illustrated in Fig. 9. The detailed DLDO design specifications are explained in Section V-A. As shown in Fig. 9, for bDSR control, power transistor M_as experience electrical stress all the time while M_bs are always OFF. For uDSR control, three randomly picked adjacent power transistor gate signals Q_{59}, Q_{60}, and Q_{61} together with two additional further separated gate signals Q_{20} and Q_{120} are demonstrated. The falling edge of Q_{60} (Q_{61}) demonstrates delay as compared to Q_{59} (Q_{60}). However, the percentage of time when power transistor M_i ($1 \le i \le N$) is active is the same for all M_is, and thus, the electrical stress can be more evenly distributed.

2) Transient Load Operation: Under transient load condition, operations of the bDSR and uDSR follow a similar activation/deactivation pattern demonstrated in Fig. 8. If $V_{\text{out}} < V_{\text{ref}}$ ($V_{\text{out}} > V_{\text{ref}}$) due to increased (decreased) load current, for bDSR, inactive (active) power transistors at the right boundary of the colored region in Fig. 8(a) are gradually turned ON (OFF) to supply the required output current and regulate V_{out}. The colored region always locates at the left side of the power transistor array. While for the proposed uDSR, inactive (active) power transistors at the right (left) boundary in Fig. 8(b) are

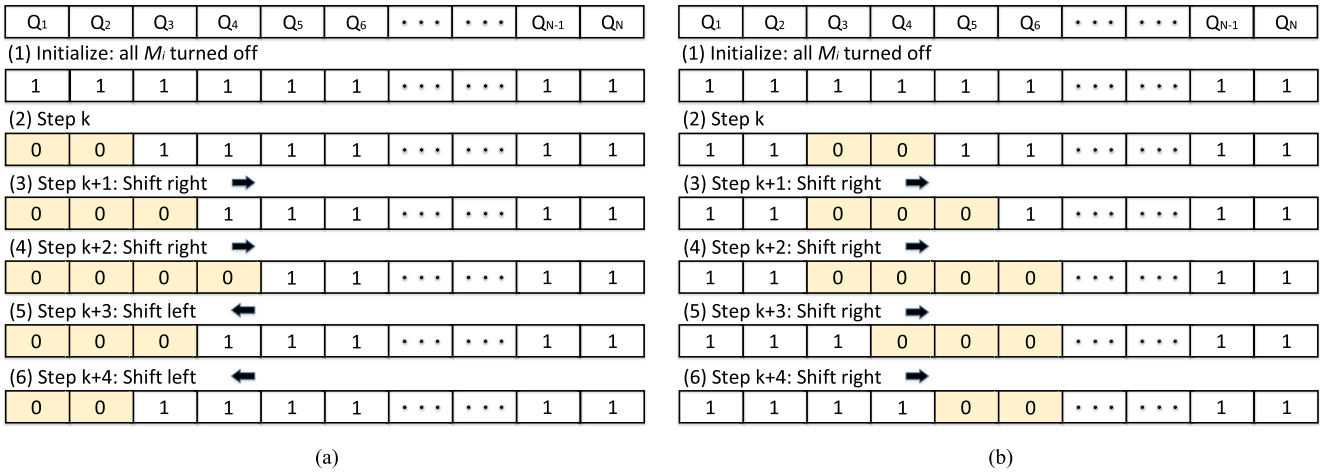


Fig. 8. Digital controller operation for DLDO at steady state with mode 2 LCO. (a) Operation of a conventional bDSR. (b) Operation of the proposed uDSR.

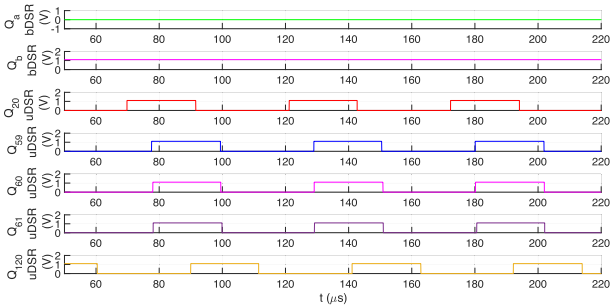


Fig. 9. Representative simulated steady-state gate signals of power transistors with bDSR and uDSR control, where Q_a $(1 \leq a < I_{\text{load}}N/I_{\max} - M)$ and Q_b $(I_{\text{load}}N/I_{\max} + M < b \leq N)$ are, respectively, gate signal of active power transistor M_a and inactive power transistor M_b with bDSR control. Q_is $(1 \leq i \leq N)$ all have similar waveform with uDSR control.

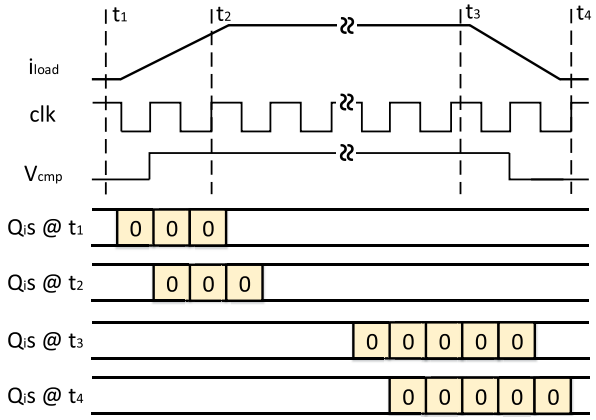


Fig. 10. Conceptual transisent waveforms and active power transistor locations of the proposed uDSR-based DLDO.

gradually turned ON (OFF) and the colored region dynamically moves right at all times, leading to a more balanced distribution of electrical stress. The operation of uDSR under transient load condition can be elaborated with the conceptual transient waveforms and active power transistor locations at different time instances in Fig. 10. A step load current with a few clock cycles of rise and fall time is utilized for illustration. Assume at t_1 before the load increase, there are three active power transistors on the left side of the power transistor array, the deactivation of power transistor at the left boundary at the next clock rising edge, and the activation of power transistor at the right boundary at the following clock rising edge lead to the updated active power transistor locations at t_2. The number of active power transistors continues to increase after t_2 and due to the steady-state operation of the uDSR following Fig. 8(b), active power transistors with an increased number move right to reach the new locations at t_3. After experiencing one more activation and deactivation of power transistor due to load decrease, the updated locations at t_4 (the second clock rising edge after t_3) are demonstrated at the bottom in Fig. 10.

Thus, regardless of the load current conditions, electrical stress can always be more evenly distributed among all of the available power transistors. Furthermore, as compared to conventional bDSR-based DLDO, the number of activated/deactivated power transistors per clock cycle remains the same, and thus, bDSR and uDSR have the same transfer function $S(z)$ in (14). Leveraging uDSR to evenly distribute electrical stress within the power transistor array does not negatively affect control loop performance.

B. Reduced Clock Pulsewidth

Dual-clock edge triggering has been employed in [26] and [32] to reduce the control signal delay, where the clocked comparator and shift register are triggered at the rising and falling edge of the clock signal, respectively. Considering the potential side effect of the control loop delay element $D'(z)$ on LCO as discussed in Section III-D, a reduced clock pulsewidth t_c, as shown in Fig. 5, is proposed to minimize the delay element. With dual-clock edge-triggering implementation of the control loop, the following condition needs to be satisfied regarding t_c for proper operation of the uDSR-based DLDO:

$$t_c > t_c^d + t_l^d + t_t^{\text{st}} \tag{24}$$

where t_l^d and t_t^{st} are, respectively, the total propagation delay of the logic gates connected to the first stage TFF within

the uDSR and the setup time of the TFF. Aging-induced degradation of t_c^d, t_l^d, and t_t^{st} needs to be considered with the targeted lifetime to decide the value of t_c. The one-shot pulse generator in [33] can be leveraged for reduced pulsewidth clock generation.

Within the proposed AA DLDO, φ_{LCO} becomes

$$\varphi''_{\text{LCO}} = \frac{\pi}{2} + \frac{\pi}{2M} - \tan^{-1}\left(\frac{\pi}{MTF_l}\right) - \frac{\pi\left(t_s^d + t_c\right)}{MT}. \quad (25)$$

The effectiveness of the proposed reduced clock pulsewidth DLDO regarding LCO mode reduction will be shown in Section V-B.

C. Discussion

1) Overhead: Considering the similar area of DFF and TFF, the proposed uDSR only induces ∼3.8% area OH per control stage compared to bDSR. The total area OH including the one-shot pulse generator is ∼2.6% of a single active DLDO area designed with μA current supply capability [16]. As little extra transistors are added per control stage and the bDSR only consumes a few μW power [16], the uDSR-induced power OH is also negligible. With larger I_{pMOS} for higher load current rating, both the area and power OH can be significantly less. Note that the area OH discussed here is different from the area OH that will be discussed in Section VI to compensate aging-induced degradation.

2) Compatibility With Quiescent Current Saving Technique: Freeze mode operation and clock gating technique [32], [34], [35] are adopted in the literature to save quiescent current at steady state. For freeze mode operation, the DLDO control circuit can be disabled once the number of active power transistors converges to save the quiescent current. In this case, the operation of the proposed unidirectional shift would also be stopped. However, after many load current changes and different steady-state operations for long-term reliability concern, the active power transistor region still moves rightward and electrical stress can also be more evenly distributed among all of the power transistors as compared to the conventional bidirectional shift method. Furthermore, the sliding clock gating technique in [32] can also be utilized to save the steady-state quiescent current. The power transistor array and the control flip-flops are divided into multiple sections with equal number within each section in [32]. During steady-state operation, if the left boundary of the active power transistor region falls within one section and the right boundary falls within another section, other sections not covering the two boundaries can be temporarily clock gated to save quiescent current. The active power transistor region still dynamically moves rightward to evenly distribute the electrical stress and the clock-gated sections also dynamically change. For this case, as not all flip-flops are clock gated, the steady-state quiescent current can be higher than that in the freeze mode operation discussed earlier. Thus, the proposed unidirectional shift scheme is still beneficial even when a steady-state quiescent current saving technique is employed. However, a tradeoff exists between the steady-state quiescent current saving and reliability enhancement enabled by the proposed scheme.

3) Compatibility With Binary-Weighted DLDO: The proposed technique has certain limitations and it may not be directly applied to recent DLDO designs with binary-weighted power transistors [36], [37]. It can, however, serve as a basis for reliability enhancement of more advanced DLDO designs. Furthermore, it may still be beneficial for DLDO designs with binary-weighted power transistors but implemented in an equivalent or similar form as in [32] and [35].

DLDO designs utilizing binary-weighted power transistors can be seen as a special case of barrel-based DLDOs with 2^N equal-weighted pMOS fingers and programmable loop gain capability such as in [32]. In [32], a shift of one, two, or three pMOS transistors in a single clock cycle is realized. As shown with the operation of our proposed AA DLDO in Fig. 7, the power transistor on the right (left) boundary of the colored region (active power transistors) is activated (deactivated) when $V_{\text{out}} < V_{\text{ref}}$ ($V_{\text{out}} > V_{\text{ref}}$) at the rising edge of the clock signal. If the barrel shifter in [32] is modified such that whenever one, two, or three pMOS transistors need to be activated (deactivated) in a single clock cycle, they are activated (deactivated) on the right (left) boundary of the active power transistor region. Electrical stress can still be more evenly distributed among all the power transistors if the DLDO is active for a sufficient amount of time (long-term reliability concern). In this sense, the proposed idea of unidirectional shift can be beneficial for DLDO designs with binary-weighted power transistors if they are implemented in the equivalent form of [32] with binary-weighted step sizes.

The analog-assisted triloop DLDO design in [35] is realized with the same intention as DLDO designs with binary-weighted power transistors for faster convergence. Within the digital portion of the triloop DLDO design in [35], three different power transistor sizes of $1\times$, $L\times$, and $(L \times M)\times$, with different number of power transistors L, M, and H, respectively, are adopted. Within each group of L, M, and H power transistors, a conventional bDSR is utilized as a controller. This bDSR within each group can be replaced with our proposed uDSR for reliability enhancement as power transistor sizes are the same within each group.

V. Evaluation

To evaluate the benefits of the proposed AA DLDO architecture in terms of reliability enhancement and to provide design insights for a targeted lifetime, an IBM POWER8 like microprocessor [19] simulation platform is constructed.

A. Simulation Framework

1) IBM POWER8 Like Microprocessor: IBM POWER8 microprocessor [19] is among one of the state-of-the-art server-class processors and, thus, a representative for evaluation of the proposed AA DLDO scheme. The corresponding technology and architecture parameters listed in Table I are adopted from [8]. The IBM POWER8 like microprocessor core as shown in Fig. 11, includes a load store unit (LSU), an execution unit (EXU), an instruction fetch unit (IFU), an instruction scheduling unit (ISU), an L1 data cache inside LSU, an L1 instruction cache inside IFU, and

TABLE I

TECHNOLOGY AND ARCHITECTURE PARAMETERS

Technology node: 22nm, Frequency: 4.0GHz TDP: 150W, Area: $441mm^2$, Vdd: 1.03V
cores: 8, issue width: 8 64 architectured FRF, 32 architectured IRF
L1-I cache: 32KB, 8-way, 64B, LRU, 1-cycle hit L1-D cache: 64KB, 8-way, 64B, LRU, 1-cycle hit L2 cache: 512KB, 8-way, 128B, LRU, 11-cycle hit L3 cache: 64MB, 8-way, 128B, LRU, 30-cycle hit

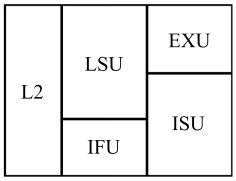


Fig. 11. Schematic demonstrating the floor plan of one core within an IBM POWER8 like microprocessor chip.

TABLE II

LOAD CHARACTERISTICS OF DIFFERENT FUNCTIONAL BLOCKS WITHIN ONE CORE OF AN IBM POWER8 LIKE MICROPROCESSOR CHIP UNDER ALL EXPERIMENTED BENCHMARKS

	IFU	LSU	ISU	EXU	L2
Min I_{load} (A)	0.091	0.172	0.125	0.251	0.178
Max I_{load} (A)	3.245	12.092	1.356	5.056	2.195
Avg I_{load} (A)	1.138	0.908	0.201	1.294	1.719

a private L2. All benchmarks are from SPALSH2x [38] and cover a wide range of representative application domains. Analysis is restricted to the region of interest of the benchmarks and eight threads are involved in the simulations. The load characteristics of different functional blocks, as shown in Fig. 11, under all experimented benchmarks are summarized in Table II.

2) DLDO Design Specifications: Distributed microregulators are implemented in IBM POWER8 microprocessor [39]. In this simulation example, a switch array of 256 pMOS transistors, which is typical in DLDO designs [16], is implemented in each microregulator. Two different DLDO designs with bDSR and uDSR controls are implemented using 32-nm PTM CMOS technology where $V_{\text{in}} = 1.1$ V and $V_{\text{out}} = 1$ V as in [39]. $I_{\text{pMOS}} = 2$ mA and $I_{\text{max}} = 512$ mA are used in the simulations, leading to 7, 24, 3, 10, and 5 microregulators (DLDOs) in, respectively, IFU, LSU, ISU, EXU, and L2 blocks to be able to supply the maximum load current across all benchmarks in each block. Load current of each block is assumed to be supplied by microregulators within that block, which is reasonable due to the principle of spatial locality [40] regarding current distribution. Each microregulator within a certain block is assumed to provide equal current due to the availability of current balancing scheme implemented within IBM POWER8 microprocessor [21]. $f_{\text{clk}} = 10$ MHz and $C = 15$ nF are used for each DLDO to achieve smaller than 10% Vdd transient voltage noise [8] most of the time. The total output capacitance of 735 nF is comparable to 750 nF used in [39]. As resonant clock meshes [41] are already deployed within IBM POWER8 processor, the complexity and OH of generating and distributing the clock signal for the DLDOs can be frequency dividers consisting of simple flip-flops and localized routing wires.

TABLE III

CONVENTIONAL DLDO PERFORMANCE DEGRADATION FOR DIFFERENT FUNCTIONAL BLOCKS UNDER ALL EXPERIMENTED BENCHMARKS FOR A 5-YEAR TIME FRAME

	IFU	LSU	ISU	EXU	L2
% I_{pMOS} degradation	16.2	21.4	15.3	16.6	15.1
% T_R degradation	9.4	12.9	8.9	9.7	8.7
% ΔV degradation	6.4	8.7	6.1	6.6	6

3) Evaluation of Aging-Induced Performance Degradation: Equations (1), (3), (6), and (8) are leveraged for the evaluation of aging-induced performance degradation. A typical temperature profile [8], [42] of 90 °C, 69 °C, 67 °C, 63 °C, and 62 °C for, respectively, LSU, EXU, IFU, ISU, and L2 is adopted for evaluations. The activity factors for both DLDO designs under different benchmarks and functional blocks are estimated through simulations in Cadence Virtuoso. The worst case I_{pMOS} degradations are used for evaluations of both designs, which is reasonable due to load characteristics of typical applications [23] and the consequent heavy use of a portion of M_is in conventional DLDOs.

B. Simulation Results

1) Performance Degradation Within Conventional DLDO: Conventional DLDO performance degradation regarding I_{pMOS}, T_R, and ΔV for different functional blocks for a 5-year time frame is summarized in Table III. These degradations apply to all the experimented benchmarks as the worst case I_{pMOS} degradation is considered. As shown in Table III, NBTI can induce serious I_{pMOS}, T_R, and ΔV degradations for all functional blocks. I_{pMOS} degradation can lead to the deterioration of DLDO V_{out} regulation capability and possible V_{out} drop under large load current conditions. Larger than 10% V_{out} drop can lead to voltage emergencies and potential execution errors for microprocessors. Similarly, T_R and ΔV degradations can, respectively, increase the duration and frequency of voltage emergencies, which can slow down microprocessor executions as further actions may need to be taken to remedy the errors. Moreover, for a longer targeted lifetime of more than 5 years, the degradations are expected to be more disastrous as I_{pMOS} degradations are even worse, as seen from Fig. 3, which may not be tolerable for critical applications where the replacement of the devices can be costly or even impossible.

2) I_{pMOS}, T_R, and ΔV Mitigation With Proposed Aging-Aware DLDO: Simulation results for all benchmarks are summarized in Figs. 12–14 regarding, respectively, I_{pMOS}, T_R, and ΔV degradation mitigation of the proposed uDSR-based DLDO as compared to the conventional DLDO design for a 5-year time frame. Up to 39.6%, 43.2%, and 42% performance improvement is achieved for, respectively, I_{pMOS}, T_R, and ΔV. The highest performance improvement is obtained for LSU with the highest operation temperature. Even at the lowest operation temperature within L2, degradation mitigations of

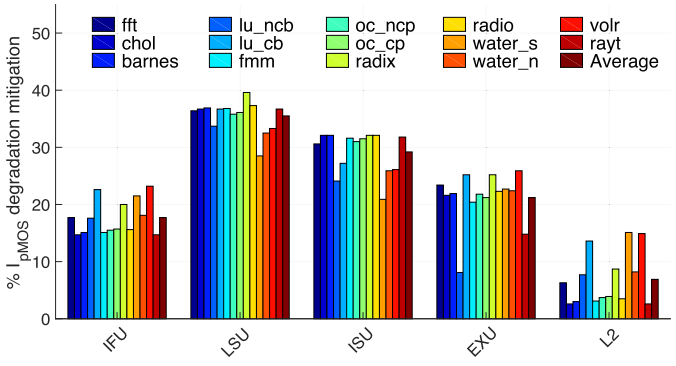


Fig. 12. Percentage I_{pMOS} degradation mitigation of the proposed AA DLDO as compared to the conventional DLDO design for different functional blocks under all experimented benchmarks.

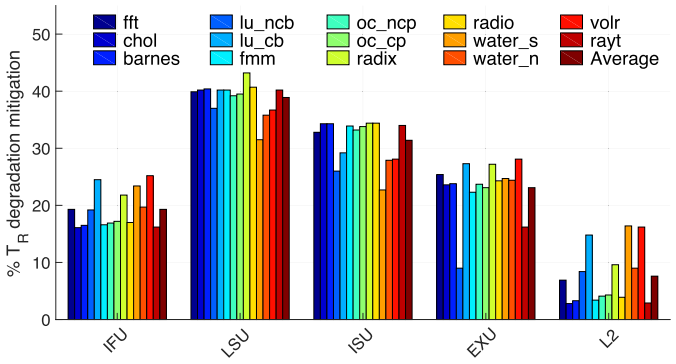


Fig. 13. Percentage T_R degradation mitigation of the proposed AA DLDO as compared to the conventional DLDO design for different functional blocks under all experimented benchmarks.

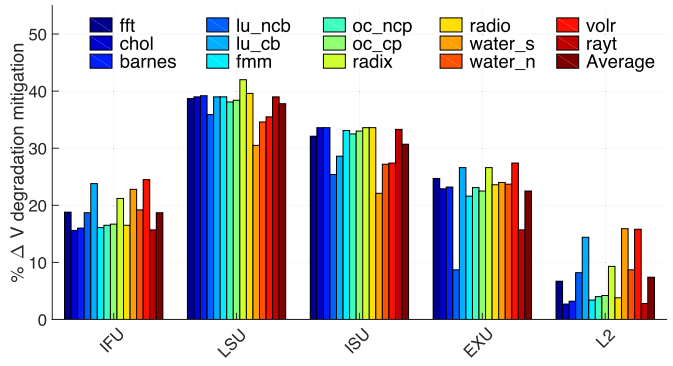


Fig. 14. Percentage ΔV degradation mitigation of the proposed AA DLDO as compared to the conventional DLDO design for different functional blocks under all experimented benchmarks.

up to 15.1%, 16.4%, and 15.9% are achieved for, respectively, I_{pMOS}, T_R, and ΔV.

3) LCO Mitigation With Proposed Aging-Aware DLDO: To verify the benefits of the proposed reduced clock pulsewidth DLDO regarding LCO mitigation, the theoretical maximum LCO mode for dual-edge-triggered and reduced clock pulsewidth DLDOs with uDSR implementation is, respectively, examined by considering BTI-induced threshold voltage degradation of the control loop. An average IBM POWER8 microprocessor temperature profile of 70 °C is

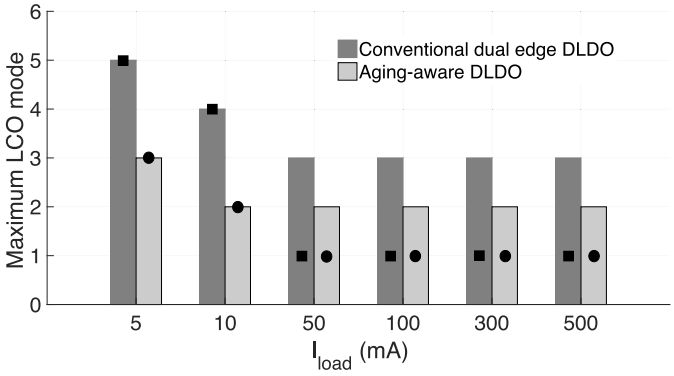


Fig. 15. Maximum LCO mode with simulation results superimposed for conventional and AA DLDO under different load current conditions after a 5-year aging period.

TABLE IV
TFF SETUP TIME, LOGIC DELAY, AND COMPARATOR DELAY BEFORE AND AFTER A 5-YEAR AGING PERIOD

	TFF setup time	Logic delay	Comparator delay
Fresh (ps)	170	209.6	171.5
Aged 5 yrs (ps)	180	227.4	225

utilized for V_{th} degradation evaluation. NBTI and PBTI are considered as the major V_{th} degradation factor for pMOS and nMOS transistors in the control loop, respectively. Under different load current conditions, the activity factor of each transistor within the control loop is obtained through simulations in Cadence Virtuoso. Equation (1) is then leveraged to calculate the V_{th} degradation for each transistor within a 5-year time frame. The calculated V_{th} degradation is embedded in each transistor by adopting the subcircuit model for BTI effect in [43] within Cadence Virtuoso simulations. The fresh and aged TFF setup time t_t^{st}, logic delay t_l^d, and comparator delay t_c^d are summarized in Table IV. The aged t_t^{st}, t_l^d, and t_c^d are approximately load current independent. $t_c = 1$ ns is adopted to satisfy timing constraint in (24). The maximum LCO mode for dual-edge-triggered and reduced clock pulsewidth DLDOs under different load current conditions after a 5-year aging period is illustrated in Fig. 15.

As seen from Fig. 15, with reduced clock pulsewidth considering aging imposed limitations, the maximum LCO mode can be greatly reduced, especially under light-load conditions. The simulated steady-state output voltages for both conventional dual-edge (CDE) triggered DLDO and the proposed AA DLDO under 10-mA load current are demonstrated in Fig. 16. LCO mode reduction from 4 to 2 and 3 times output voltage ripple amplitude reduction are achieved. As the minimum and average I_{load} can be way smaller than the maximum I_{load} shown in Table II, especially for LSU, light- and medium-load conditions are experienced most of the time such that outstanding benefits can be achieved with the proposed AA DLDO considering the negligible power and area OH induced.

Furthermore, in many applications, the clock frequency can be much higher than 10 MHz such as 1 GHz in [44]. However, the 1-GHz sampling clock sacrifices the quiescent

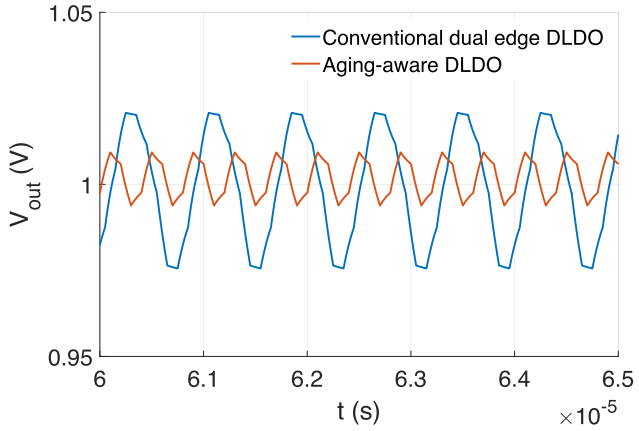


Fig. 16. Simulated output voltage ripple and LCO mode reduction with the proposed AA DLDO under 10-mA load current.

TABLE V
MAXIMUM LCO MODE UNDER DIFFERENT SAMPLING CLOCK FREQUENCY AND LOAD CURRENT CONDITION FOR CDE AND AA DLDO

CDE/AA LCO mode	Sampling clock frequency f_{clk} (MHz)				
I_{load} (mA)	10	50	100	300	500
10	4/2	8/6	11/9	20/18	27/27
100	3/2	3/2	4/3	6/6	8/8
500	3/2	3/2	3/2	3/3	4/4

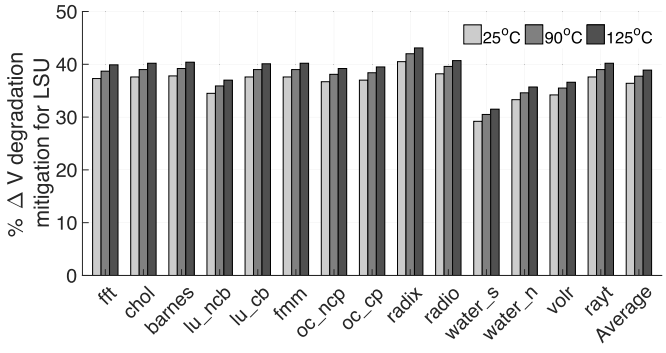


Fig. 17. Percentage ΔV degradation mitigation of the proposed AA DLDO as compared to the conventional DLDO design for LSU under all experimented benchmarks and different temperature profile.

current. Recent works such as [32] and [45] utilize a high clock frequency for fast transient and a much lower frequency for steady-state operation. For a better verification of LCO improvement utilizing the proposed reduced clock pulsewidth scheme, the maximum LCO mode under different sampling clock frequencies and load current conditions for both CDE and AA DLDO is shown in Table V. Seen from the table, the proposed reduced clock pulsewidth scheme demonstrates the maximum LCO mode reduction under a wide f_{clk} range, especially under light-load current condition. For a clock frequency of 1 GHz, there would be no room to further reduce the pulsewidth due to the timing constraint. However, as discussed earlier, clock frequency utilized at steady-state operation is typically much lower.

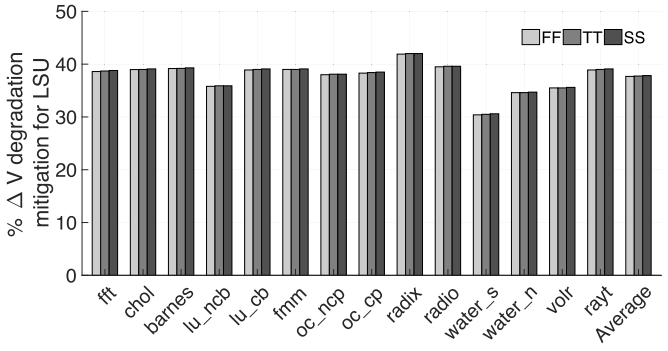


Fig. 18. Percentage ΔV degradation mitigation of the proposed AA DLDO as compared to the conventional DLDO design for LSU under all experimented benchmarks and different process corners (FF = fast-fast, TT = typical-typical, and SS = slow-slow).

4) Discussion: Regarding the temperature variation effects within each functional unit, representative results for ΔV degradation mitigation are shown in Fig. 17. Fig. 17 demonstrates the percentage ΔV degradation mitigation of the proposed AA DLDO as compared to the conventional DLDO for LSU unit under all experimented benchmarks and different temperature profiles. As shown in Fig. 17, as the temperature increases, the percentage ΔV degradation mitigation only slightly increases since the performance degradation for conventional and proposed DLDOs both increase. As compared to temperature variations, load current profile plays a more important role regarding the relative advantages of the proposed AA DLDO, as shown in Fig. 14.

The proposed approach is still beneficial considering process variations. Representative results are shown in Fig. 18 with percentage ΔV degradation mitigation of the proposed AA DLDO as compared to the conventional DLDO for LSU unit under all considered benchmarks and different process corners. For a fair comparison, the conventional and proposed DLDO are assumed to have the same process corner. As shown with the results, process variations have negligible effects on the effectiveness of the proposed approach. Comparison with related work is shown in Table VI.

VI. TRADEOFF BETWEEN AREA OVERHEAD AND PROGRAM OUTPUT QUALITY

Considering aging effects, regulators are typically designed and optimized for the expected service life of the processor. Deploying regulators optimized for a shorter service life cannot guarantee error-free operation, however, if such regulators are confined to feed error-tolerant loads, the service life can be traded for lower hardware complexity, which almost always directly translates into area savings. Note that the area represents a scarce on-chip resource for distributed voltage regulators as many of these regulators are squeezed between various circuit blocks. Such area savings can enable a higher number of on-chip voltage regulators, and hence enhance the scalability of on-chip voltage regulation. To illustrate this point, the percentage area OH within each functional unit to achieve the same fresh ΔV performance utilizing the conventional DLDO is examined in Fig. 19. The relative area

TABLE VI
COMPARISON WITH PREVIOUS AA ON-CHIP VOLTAGE REGULATORS

	[13]	[14]	[46]	This work
Year	2014	2015	2017	2018
Regulator type	Buck	DLDO	DLDO	DLDO
Technique	Rotating phase-shedding	Row rotation scheme	Code roaming algorithm	Unidirectional shift controller
Wide load range	No	Yes	Yes	Yes
Additional controller	Yes	Yes	Yes	No
Circuit overhead	DFFs and logic gates	Multiple decoders	Decoder	Modification of original controller
Improvement	Light load efficiency	Hotspot, metal track rupture, transistor breakdown	Hotspot, electromigration	I_{pMOS}, T_R, ΔV, LCO

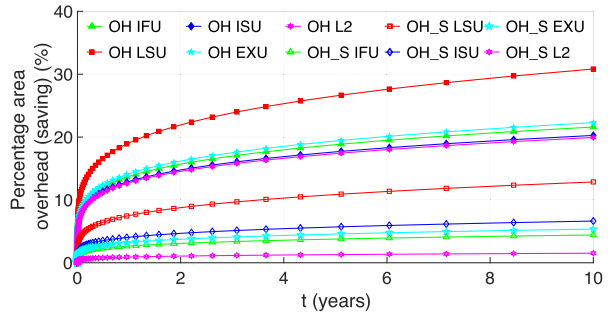


Fig. 19. Percentage area OH utilizing conventional DLDO and percentage area OH saving (OH_S) utilizing AA DLDO for ΔV degradation mitigation within each functional unit.

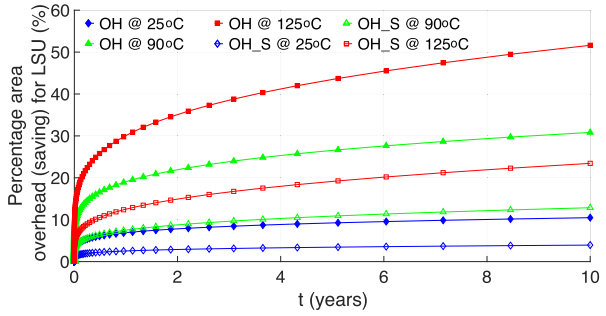


Fig. 20. Percentage area OH utilizing conventional DLDO and percentage area OH saving utilizing AA DLDO for ΔV degradation mitigation within LSU under different temperature profile.

between pMOS array and shift register and output capacitance is based on the data in [45] for estimation. Adding extra output capacitance to mitigate ΔV degradation is considered in the estimation. The percentage area OH is relative to the original DLDO area including output capacitance designed in an aging-unaware fashion. The percentage area OH saving within each functional unit for ΔV degradation mitigation utilizing the proposed AA DLDO is also demonstrated in Fig. 19. As shown in Fig. 19, a large area OH can be introduced to mitigate aging-induced transient voltage noise degradation for conventional DLDOs. Similar to the trend demonstrated in Fig. 3, the area penalty required to compensate for the aging-related deterioration of ΔV is significant, especially in the first 2 years. The percentage area OH also plateaus to within 10% after 2 years. These trends need to be considered to realize optimal design based on different application environment and lifetime targets. Furthermore, leveraging the proposed AA DLDO, due to mitigation of aging-induced ΔV degradation, significant area OH savings compared to the conventional DLDO case can be achieved as shown in Fig. 19. Our proof-of-concept analysis reveals approximately 1% total DLDO area, which corresponds to ∼36% active DLDO area, savings for per year service life reduction.

The temperature variation effects on percentage area OH (saving) within LSU are demonstrated in Fig. 20. As seen from the figure, as the temperature increases, the percentage area OH needed for the conventional DLDO to mitigate ΔV degradation increases significantly. The percentage area OH saving achieved by the AA DLDO also greatly increases. Although the relative benefits of AA DLDO do not improve significantly as the temperature increases shown in Fig. 17,

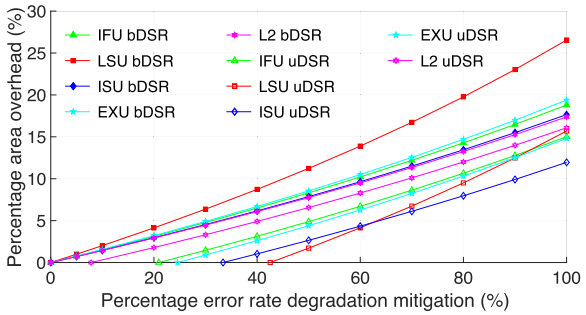


Fig. 21. Percentage area OH within each functional unit for percentage error rate degradation mitigation utilizing bDSR and uDSR-based DLDO.

the area OH saving is considerable due to the relatively large ratio between the area of output capacitance and that of active DLDO.

For a proof-of-concept analysis, considering a 5-year aging period, the percentage area OH within each functional unit for percentage error rate degradation mitigation utilizing bDSR and uDSR-based DLDOs is demonstrated in Fig. 21 based on the relationship between the error rate and supply voltage demonstrated in [47] and [48]. The percentage error rate degradation mitigation is with respect to the degraded error rate utilizing bDSR-based DLDO and a 100% error rate degradation mitigation means the same error rate within each functional unit is achieved as the fresh one after a 5-year aging period. As seen from Fig. 21, contrary to the bDSR curves, the uDSR curves do not start from the origin, which means with negligible area OH, uDSR-based DLDO achieves a certain amount of error rate degradation mitigation compared to bDSR-based DLDO. Also, for the same amount of error rate

degradation mitigation, the area OH needed for uDSR-based DLDO is lower than that of bDSR-based DLDO.

VII. Conclusion

As an emerging and essential part of the modern processor power delivery network, DLDO regulators experience serious aging-induced performance degradations including I_{pMOS}, T_R, and ΔV. In particular, DLDO degradation can increase noise in the supply voltage and further deteriorate the program output quality. Area OH needed to fully compensate these degradations can be significant, especially when a conventional DLDO design is utilized. Algorithmic noise tolerance of different processor components is leveraged as an area–quality control knob to alleviate the area OH requirement through scalable on-chip voltage regulation at design time. Furthermore, DLDO designed in an AA fashion is proposed to mitigate aging-induced performance degradations with negligible power and area OH. With reduced DLDO performance degradation, a significantly better area and quality tradeoff can be achieved due to AA DLDO-induced area OH savings. Therefore, more efficient scalable on-chip voltage regulation can be realized with the proposed AA DLDO. Up to 43.2% transient and 3× steady-state DLDO performance improvement as well as more than 10% area OH saving can be achieved utilizing the proposed AA paradigm.

References

[1] V. De, "Fine-grain power management in manycore processor and System-on-Chip (SoC) designs," in *Proc. ICCAD*, Nov. 2015, pp. 159–164.

[2] I. P. Vaisband, R. Jakushokas, M. Popovich, A. V. Mezhiba, S. Köse, and E. G. Friedman, *On-Chip Power Delivery and Management*, 4th ed. Cham, Switzerland: Springer, 2016.

[3] Y.-K. Chen *et al.*, "Convergence of recognition, mining, and synthesis workloads and its implications," *Proc. IEEE*, vol. 38, no. 5, pp. 790–807, May 2008.

[4] H. Cho, L. Leem, and S. Mitra, "ERSA: Error resilient system architecture for probabilistic applications," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 31, no. 4, pp. 546–558, Apr. 2012.

[5] M. M. Mahmoud, N. Soin, and H. A. H. Fahmy, "Design framework to overcome aging degradation of the 16 nm VLSI technology circuits," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 33, no. 5, pp. 691–703, May 2014.

[6] D. Rossi, V. Tenentes, S. Yang, S. Khursheed, and B. M. Al-Hashimi, "Reliable power gating with NBTI aging benefits," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 24, no. 8, pp. 2735–2744, Aug. 2016.

[7] T.-B. Chan, J. Sartori, P. Gupta, and R. Kumar, "On the efficacy of NBTI mitigation techniques," in *Proc. DATE*, Mar. 2011, pp. 1–6.

[8] S. K. Khatamifard, L. Wang, W. Yu, S. Köse, and U. R. Karpuzcu, "ThermoGater: Thermally-aware on-chip voltage regulation," in *Proc. ISCA*, Jun. 2017, pp. 120–132.

[9] K.-C. Wu, I.-C. Lin, Y.-T. Wang, and S.-S. Yang, "BTI-aware sleep transistor sizing algorithm for reliable power gating designs," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 33, no. 10, pp. 1591–1595, Oct. 2014.

[10] I. Agbo *et al.*, "Integral impact of BTI, PVT variation, and workload on SRAM sense amplifier," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 25, no. 4, pp. 1444–1454, Apr. 2017.

[11] J. Fang and S. S. Sapatnekar, "The impact of BTI variations on timing in digital logic circuits," *IEEE Trans. Device Mater. Rel.*, vol. 13, no. 1, pp. 277–286, Mar. 2013.

[12] J. Wu, A. Boyer, J. Li, S. B. Dhia, and R. Shen, "Characterization of changes in LDO susceptibility after electrical stress," *IEEE Trans. Electromagn. Compat.*, vol. 55, no. 5, pp. 883–890, Oct. 2013.

[13] Y. Ahn, I. Jeon, and J. Roh, "A multiphase buck converter with a rotating phase-shedding scheme for efficient light-load control," *IEEE J. Solid-State Circuits*, vol. 49, no. 11, pp. 2673–2683, Nov. 2014.

[14] P. Patra, R. Muthukaruppan, and S. Mangal, "A reliable digitally synthesizable linear drop-out regulator design for 14 nm SoC," in *Proc. IEEE INIS*, Dec. 2015, pp. 73–76.

[15] L. Wang, S. K. Khatamifard, O. A. Uzun, U. R. Karpuzcu, and S. Köse, "Efficiency, stability, and reliability implications of unbalanced current sharing among distributed on-chip voltage regulators," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 25, no. 11, pp. 3019–3032, Nov. 2017.

[16] Y. Okuma *et al.*, "0.5-V input digital LDO with 98.7% current efficiency and 2.7-μ A quiescent current in 65 nm CMOS," in *Proc. IEEE CICC*, 2010, pp. 1–4.

[17] L. Wang, S. K. Khatamifard, U. R. Karpuzcu, and S. Köse, "Mitigation of NBTI induced performance degradation in on-chip digital LDOs," in *Proc. DATE*, Mar. 2018, pp. 803–808.

[18] S. Köse, "Regulator-gating: Adaptive management of on-chip voltage regulators," in *Proc. GLSVLSI*, 2014, pp. 105–110.

[19] E. J. Fluhr *et al.*, "5.1 POWER8: A 12-core server-class processor in 22 nm SOI with 7.6 Tb/s off-chip bandwidth," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2014, pp. 96–97.

[20] M. Alioto, *Enabling the Internet of Things: From Integrated Circuits to Integrated Systems*. Cham, Switzerland: Springer, 2017.

[21] J. F. Bulzacchelli *et al.*, "Dual-loop system of distributed microregulators with high DC accuracy, load response time below 500 ps, and 85-mV dropout voltage," *IEEE J. Solid-State Circuits*, vol. 47, no. 4, pp. 863–874, Apr. 2012.

[22] A. Marongiu, L. Benini, A. Acquaviva, and A. Bartolini, "Analysis of power management strategies for a large-scale SoC platform in 65 nm technology," in *Proc. 11th EUROMICRO Conf. Digit. Syst. Design Archit., Methods Tools*, Sep. 2008, pp. 259–266.

[23] D. Pathak, H. Homayoun, and I. Savidis, "Smart grid on chip: Work load-balanced on-chip power delivery," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 25, no. 9, pp. 2538–2551, Sep. 2017.

[24] S. B. Nasir and A. Raychowdhury, "On limit cycle oscillations in discrete-time digital linear regulators," in *Proc. IEEE APEC*, Mar. 2015, pp. 371–376.

[25] M. Huang, Y. Lu, S.-W. Sin, S. P. U, R. P. Martins, and W.-H. Ki, "Limit cycle oscillation reduction for digital low dropout regulators," *IEEE Trans. Circuits Syst., II, Exp. Briefs*, vol. 63, no. 9, pp. 903–907, Sep. 2016.

[26] S. B. Nasir and A. Raychowdhury. (Jan. 2015). "A model study of an all-digital, discrete-time and embedded linear regulator." [Online]. Available: https://arxiv.org/abs/1501.00579

[27] M. A. Alam and S. Mahapatra, "A comprehensive model of PMOS NBTI degradation," *Microelectron. Rel.*, vol. 45, no. 1, pp. 71–81, 2005.

[28] Y. Cao, *Predictive Technology Model for Robust Nanoelectronic Design*. New York, NY, USA: Springer, 2011.

[29] S. Leitner, P. West, C. Lu, and H. Wang, "Digital LDO modeling for early design space exploration," in *Proc. IEEE SOCC*, Sep. 2016, pp. 7–12.

[30] H. Yi, T. Yoneda, M. Inoue, Y. Sato, S. Kajihara, and H. Fujiwara, "A failure prediction strategy for transistor aging," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 20, no. 11, pp. 1951–1959, Nov. 2012.

[31] M. Alioto and G. Palumbo, "NAND/NOR adiabatic gates: Power consumption evaluation and comparison versus the fan-in," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 49, no. 9, pp. 1253–1262, Sep. 2002.

[32] S. B. Nasir, S. Gangopadhyay, and A. Raychowdhury, "All-digital low-dropout regulator with adaptive control and reduced dynamic stability for digital load circuits," *IEEE Trans. Power Electron.*, vol. 31, no. 12, pp. 8293–8302, Dec. 2016.

[33] V. R. H. Lorentz *et al.*, "Lossless average inductor current sensor for CMOS integrated DC–DC converters operating at high frequencies," *Analog Integr. Circuits Signal Process.*, vol. 62, no. 3, pp. 333–344, 2009.

[34] Y. H. Lee *et al.*, "A low quiescent current asynchronous digital-LDO with PLL-modulated fast-DVS power management in 40 nm SoC for MIPS performance improvement," *IEEE J. Solid-State Circuits*, vol. 48, no. 4, pp. 1018–1030, Apr. 2013.

[35] M. Huang, Y. Lu, S.-P. U, and R. P. Martins, "An analog-assisted tri-loop digital low-dropout regulator," *IEEE J. Solid-State Circuits*, vol. 53, no. 1, pp. 20–34, Jan. 2018.

[36] D. Kim and M. Seok, "A fully integrated digital low-dropout regulator based on event-driven explicit time-coding architecture," *IEEE J. Solid-State Circuits*, vol. 52, no. 11, pp. 3071–3080, Nov. 2017.
[37] L. G. Salem, J. Warchall, and P. P. Mercier, "A successive approximation recursive digital low-dropout voltage regulator with PD compensation and sub-LSB duty control," *IEEE J. Solid-State Circuits*, vol. 53, no. 1, pp. 35–49, Jan. 2018.
[38] S. C. Woo, M. Ohara, E. Torrie, J. P. Singh, and A. Gupta, "The SPLASH-2 programs: Characterization and methodological considerations," in *Proc. ISCA*, Jun. 1995, pp. 24–36.
[39] Z. Toprak-Deniz *et al.*, "5.2 Distributed system of digitally controlled microregulators enabling per-core DVFS for the POWER8 microprocessor," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2014, pp. 98–99.
[40] S. Köse and E. G. Friedman, "Efficient algorithms for fast IR drop analysis exploiting locality," *Integration*, vol. 45, pp. 149–161, Mar. 2012.
[41] E. J. Fluhr *et al.*, "The 12-core POWER8 processor with 7.6 Tb/s IO bandwidth, integrated voltage regulation, and resonant clocking," *IEEE J. Solid-State Circuits*, vol. 50, no. 1, pp. 10–23, Jan. 2015.
[42] S. Köse, "Thermal implications of on-chip voltage regulation: Upcoming challenges and possible solutions," in *Proc. DAC*, Jun. 2014, pp. 1–6.
[43] W. Wang, V. Reddy, A. T. Krishnan, R. Vattikonda, S. Krishnan, and Y. Cao, "Compact modeling and simulation of circuit reliability for 65-nm CMOS technology," *IEEE Trans. Device Mater. Rel.*, vol. 7, no. 4, pp. 509–517, Dec. 2007.
[44] M. Onouchi *et al.*, "A 1.39-V input fast-transient-response digital LDO composed of low-voltage MOS transistors in 40-nm CMOS process," in *Proc. IEEE Asian Solid-State Circuit Conf.*, Nov. 2011, pp. 37–40.
[45] M. Huang, Y. Lu, S. Sin, S. U, and R. P. Martins, "A fully integrated digital LDO with coarse–fine-tuning and burst-mode operation," *IEEE Trans. Circuits Syst., II, Exp. Briefs*, vol. 63, no. 7, pp. 683–687, Jul. 2016.
[46] R. Muthukaruppan *et al.*, "A digitally controlled linear regulator for per-core wide-range DVFS of atom cores in 14 nm tri-gate CMOS featuring non-linear control, adaptive gain and code roaming," in *Proc. IEEE ESSCIRC*, Sep. 2017, pp. 275–278.
[47] D. Ernst *et al.*, "Razor: A low-power pipeline based on circuit-level timing speculation," in *Proc. MICRO*, Dec. 2003, pp. 7–18.
[48] 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.*, vol. 21, no. 1, pp. 3–13, Feb. 2008.

Longfei Wang received the B.S. degree in electronic information engineering from Southwest Jiaotong University, Chengdu, China, in 2010 and the M.S. degree in electrical engineering from Texas Tech University, Lubbock, TX, USA, in 2013. He is currently working toward the Ph.D. degree in electrical engineering at the University of South Florida, Tampa, FL, USA.

His current research interests include on-chip voltage regulation and power management.

S. Karen Khatamifard received the B.S. degree in electrical engineering from the Sharif University of Technology, Tehran, Iran, in 2013. He is currently working toward the Ph.D. degree at the University of Minnesota, Minneapolis, MN, USA.

In 2013, he joined the ALTAI Group, University of Minnesota. His current research interests include computer architecture.

Ulya R. Karpuzcu received the M.S. and Ph.D. degrees in electrical and computer engineering from the University of Illinois at Urbana–Champaign, Champaign, IL, USA.

She is currently an Associate Professor of Electrical and Computer Engineering at the University of Minnesota, Minneapolis, MN, USA. Her current research interests include the impact of technology on computing systems, energy-efficient computer architecture, hardware reliability, and approximate computing.

Dr. Karpuzcu is a Fulbright Fellow. She was a recipient of the NSF CAREER Award.

Selçuk Köse (S'10–M'12) received the B.S. degree in electrical and electronics engineering from Bilkent University, Ankara, Turkey, in 2006, and the M.S. and Ph.D. degrees in electrical engineering from the University of Rochester, Rochester, NY, USA, in 2008 and 2012, respectively.

He was at TUBITAK, Ankara, Intel Corporation, Santa Clara, CA, USA, and Freescale Semiconductor, Tempe, AZ, USA. He is currently an Associate Professor at the Department of Electrical Engineering, University of South Florida, Tampa, FL, USA. His current research interests include integrated voltage regulation, 3-D integration, hardware security, and green computing.

Dr. Köse has served on the Technical Program and Organization Committees of various conferences. He was a recipient of the NSF CAREER Award, the Cisco Research Award, the USF College of Engineering Outstanding Junior Researcher Award, and the USF Outstanding Faculty Award. He is an Associate Editor of the *Journal of Circuits, Systems, and Computers* and the *Microelectronics Journal*.