Power Delivery Networks for Embedded Mobile SoCs: Architectural Advancements and Design Challenges

MUHAMMAD ABRAR AKRAM\(^1\), (Member, IEEE),
IN-CHUL HWANG\(^2\), (Senior Member, IEEE),
AND SOHMYUNG HA\(^1,3\), (Senior Member, IEEE)

\(^1\)Division of Engineering, New York University Abu Dhabi, Abu Dhabi, United Arab Emirates
\(^2\)Department of Electrical and Electronics Engineering, Kangwon National University, Chuncheon 200-701, South Korea
\(^3\)Tandon School of Engineering, New York University, New York, NY 10003, USA

Corresponding author: Sohmyung Ha (sohmyung@nyu.edu)

This work was supported by the National Research Foundation (NRF) through the Korean Government (MOE) under Grant NRF-2020R1A1A3073683.

\begin{abstract}
Conventional power delivery networks (PDNs) and power management techniques using off-chip power converters with bulky passive components cannot meet the ever-evolving power delivery requirements of high-performance modern system-on-chips (SoCs). In SoCs, heterogeneous components, including multi-core processors and mixed-signals peripheral circuits, require state-of-the-art PDNs to provide high-quality power-on-demand with minimum latency, simultaneously achieving the small-factor, high conversion efficiency, and minimum current consumption. To satisfy these power delivery requirements, various PDNs have been developed over the past decades, such as the conventional architectures using off-chip power converters, architectures using in-package power converters and fully-integrated power converters, and heterogeneous architectures (off-chip power converters and on-chip regulators). This paper reviews these architectural advancements of the PDNs and their advantages and limitations, which leads us to discuss a heterogeneous PDN structure consisting of a highly efficient off-chip switching-mode power converter and multiple highly precise small linear regulators integrated on chip at point-of-load locations. The heterogeneous PDN has been proved one of the most suitable architectures to achieve high-quality fine-grained on-chip power delivery and management in SoCs. This paper also discusses unified voltage and frequency regulators (UVFRs), which support dynamic-variation-aware dynamic voltage and frequency scaling (DVFS) for fine-grained power management in multi-core processors. Based on the UVFR, we propose a modified heterogeneous PDN using frequency-referenced digital low-dropout regulators (FR-DLDOs) for more efficient DVFS, eliminating the need for band-gap circuits to provide reference voltages. As an exemplary implementation of FR-DLDO for this PDN, we present an FR-DLDO with a transient-boost control, which accelerates the transient response. The transient-boost control is activated dynamically only when an abrupt change happens out of the steady state. The implemented FR-DLDO fabricated in a 40-nm CMOS process outperforms other FR-DLDOs in the figure-of-merit and peak power efficiency while driving 40 mA of load current.
\end{abstract}

\begin{IEEEkeywords}
Power delivery network (PDN), fine-grained power management, dynamic-voltage and frequency scaling (DVFS), dynamic-voltage scaling (DVS), low-dropout regulator (LDO), digital LDO, unified-voltage and frequency regulator (UVFR), frequency-referenced digital LDO (FR-DLDO), transient-boost control.
\end{IEEEkeywords}

\begin{glossary}
\begin{itemize}
\item \textbf{CMOS} Complimentary Metal-Oxide-Semiconductor
\item \textbf{DVS} Dynamic Voltage Scaling
\item \textbf{DVFS} Dynamic Voltage and Frequency Scaling
\item \textbf{FIVR} Fully Integrated Voltage Regulator
\item \textbf{FR-DLDO} Frequency-Referenced Digital LDO
\item \textbf{LDO} Low-Dropout Regulator
\item \textbf{MIM} Metal-Insulator-Metal
\end{itemize}
\end{glossary}
PDN  Power Delivery Network
PoL  Point-of-Load
PSiP  Power Supplies in a Package
PWM  Pulse-Width Modulation
QOESC  Quadcore Elastic Switched-Capacitor
RPDN  Reconfigurable PDN
SoC  System-on-Chip
TDP  Thermal Design Power
TRO  Tunable-Replical Oscillator
UVFR  Unified-Voltage and Frequency Regulator

I. INTRODUCTION

The continuous and aggressive down-scaling of CMOS technologies has led to on-chip integration of various micro/nano-electronic circuits and systems on a SoC platform [1], [2]. In SoC platforms, heterogeneous components, including multi-core processors and other peripheral circuits, which are mixed-signal (analog and digital) in nature, are all integrated together on a single semiconductor chip [3]–[5]. Such heterogeneous integration enables versatile applications and high performances while maintaining the minimum energy-performance trade-off, high power efficiency, minimum footprint area, PDNs should be efficiently designed.

One of the most traditional PDN configurations is shown in Fig. 1(a). It uses off-chip power converters only. The battery power is first down-converted by off-chip switching-mode power converters, and then the power is delivered through board-level interconnects and I/O pads to the chip [11]. This PDN type with off-chip switching-mode power converters has a high power conversion efficiency. Still, it typically consumes a large board area due to off-chip converters and passive components. Moreover, it is difficult to meet modern SoCs’ huge current demands, which have a significantly high density of on-chip circuits. High currents through board-level interconnections and I/O pads inevitably cause large voltage drops and large power consumption, severely degrading the overall efficiency and the quality of on-chip power [12].

To reduce the voltage drops and power losses associated with board-level interconnects, power converters can be placed inside the package together with the chip, as shown in Fig. 1(b). This configuration is called PSiP. Within the same package of the load device (SoC), different chips such as switching drivers, controllers, and passive components are placed together. They can effectively lower the parasitic impedance effects both on the board and the package. However, the typical in-package integration technologies are still limited in supporting the increased number of on-chip power domains with large load current requirements [13]. The PSiPs have been considered as a partially off-chip configuration and an intermediate power-supply technology in terms of cost, complexity, and performances [13], [14].

To meet the challenging and ever-increasing power delivery demands in recent SoCs, efforts have been made to advance in full on-chip integration of the power supplies and converters, as illustrated in Fig. 1(c). Several state-of-the-art power converters were fully integrated on chip by using on-die MIM capacitors, air-core inductors, or non-magnetic package-trace inductors [15]–[18]. These fully integrated power converters improve the delivered power’s quality and significantly reduce the transient times and the voltage droops in load transient occasions. For example, a fully integrated PDN [15] consisting of multiple FIVRs is implemented in Intel core SoC. It has a driving capacity of 700 A of load current while maintaining a peak power efficiency of 90%. In addition, it enables >50% improvement in the battery life for mobile products and ×2 – 3 increase in the peak available power. Although full integration of power supplies has multiple advantages, a single on-chip power converter for each voltage domain cannot supply sufficient regulated currents with required preciseness.

To maintain a high-quality regulated power supply all across each voltage domain, hundreds of ultra-small micro-power converters/regulators can be integrated on chip at the PoL locations within each individual voltage domain [10], [19], [20]. This PoL-distributed PDN is illustrated in Fig. 1(d). This fine-grained power delivery can significantly enhance power supplies’ quality and speed because PoL-embedded small regulators can achieve smooth, fast, and precise voltage regulation. The effectiveness and utilization of PoL-distributed PDN are evident from the recent commercial chip designs [6], [20]. Thanks to the PoL-integration of multiple micro-power regulators, high energy-performance trade-offs were achieved [6], [20]. Although full integration of switching-mode power converters has been demonstrated to some extent, however, the on-chip integration of...
large passive components still remains challenging. Besides, ultra-small regulators are generally less power efficient due to their inherent lossy nature (refer to Sec. II-B for more details). Thus, efforts are being made to mitigate these losses by system-level solutions through on-chip distribution networks [8].

To optimize the performances of PDNs, various equation-based verification and mathematical analyses have been presented. Because they are out of the main scope of this paper, some major works are briefly given here for readers’ reference [12], [21]–[26]. An analytical analysis based on a convex optimization method can be used for a PDN with on-chip/off-chip buck converters, providing an accurate and fast evaluation of important characteristic parameters, such as power efficiency, output stability, and DVS [21]. As another example, a geometric programming can be utilized to find the optimal design variables for different architectures such as on-chip power converters of power delivery systems [26].

To determine suitable locations of on-chip power regulators, various optimization algorithms that are widely used for facility location problems have been applied in PDNs [22]. [23] presents a comprehensive methodology based on Mason’s Gain Formula for modeling and analyzing distributed linear regulators and their interactions to optimize the PDN’s overall stability. To boost power/energy-efficiency gains, some other optimization techniques are also presented [12], [24], [25].

As briefly discussed so far, various PDN architectures and techniques have been proposed to date to achieve highly efficient fine-grained power delivery and management in SoCs with robustness and small form factor. These architectures and techniques are comprehensively discussed in the next sections. In the next section (Sec. II), the main power converter topologies are reviewed and compared. Sec. III presents a comprehensive comparison between typical PDNs using off-chip power converters and heterogeneous PDNs with an off-chip power converter and on-chip small
regulators. In Sec. IV, state-of-the-art heterogeneous PDNs are discussed. Sec. V presents a recently proposed new approach of power management using UVFRs along with state-of-the-art UVFR architectures for fine-grained dynamic voltage and frequency scaling. The architecture of a proposed heterogeneous PDN is presented in Sec. VI, and Sec. VII presents an implemented example of the regulator based on the principle of UVFR along with some measurement results. The paper is concluded in Sec. VIII.

II. TOPOLOGIES OF POWER CONVERTERS

DC-DC power converters can be largely categorized into two types, switching and linear converters, depending on the regulation control method. The switching-mode converters used to be preferred over the linear converters due to their higher power efficiencies (ideally near 100%). However, they normally require large off-chip passive components. On the contrary, the linear regulators typically consume much less area, so they are considered more suitable for full on-chip integration. Both the switching and linear converters are discussed more in detail in the following subsections.

A. SWITCHING-MODE CONVERTERS

The switching-mode power converters convert the input DC voltage ($V_{IN}$) to an output DC voltage ($V_{OUT}$) by using duty-cycled switching of the power MOSFETs, which are typically controlled by a PWM controller. There are two types of switching-mode power converters, i.e., buck and boost, depending on the voltage relation between $V_{IN}$ and $V_{OUT}$. The boost converter steps up $V_{IN}$ while the buck converter steps down the $V_{IN}$. The buck converter architecture shown in Fig. 2(a) is widely used to supply a wide range of output voltage with a high current delivery while offering a high power efficiency (e.g., $>90\%$). However, they intrinsically have large voltage and current ripples at the output because of their switching-based control and operation. To reduce these ripples, two techniques can be commonly utilized. First, the size of the LC filter can be increased but at the cost of more area. Second, the switching frequency can be increased, but at the cost of more power consumption. Moreover, most performance metrics of the switching-mode buck converters such as the power efficiency, output load current ($I_{LOAD}$), and transient response are also largely affected by the LC filter size. Hence, on-chip integration of the buck converters is greatly hindered by these bulky passive components. The required size of these passive components can be reduced for full on-chip integration by operating at an ultra-high clock frequency $F_{SW}$, but at the cost of increased power dissipation in the converter [11], [27]. As a result, the energy efficiency degrades significantly in on-chip buck converters.

Recently, some methods to integrate the passives, i.e., inductors and capacitors, on chip for buck converters have been investigated. Such methods include the techniques using air-core inductors [28]–[31], non-magnetic package trace inductors [15], stacked inductors [32] and on-die MIM capacitors [33]–[35]. These FIVRs can achieve faster transient responses than the regulators using off-chip passive components. However, the integrated passive components still suffer from significant area overheads and poor quality factors owing to their large equivalent series resistance (ESR, $\approx200m\Omega/nH$). Moreover, for full integration, the converters should operate at a very high frequency of $F_{SW}$ as $100s\,MHz$. Such a fast $F_{SW}$ incurs high switching losses in the power switching devices and their corresponding drivers, thereby degrading the overall power efficiency. Due to these area and power overheads, on-chip integration of switching power converters is still considered undesirable for SoCs, especially ones with multiple voltage domains like the one shown in Fig. 1(d).

B. LINEAR CONVERTERS (LOW DROPOUT REGULATORS)

The linear power converters, widely known as LDOs, convert $V_{IN}$ to a regulated $V_{OUT}$ against variations of the load current and the input voltage by comparing the feedback voltage with a reference voltage ($V_{REF}$) using an error-amplifier, as shown in Fig. 2(b). LDOs can provide a more precise and ripple-free supply voltage and current to a load circuit with much less area overhead than switching mode power converters (buck converters) [36]–[39]. They are considered as a better candidate for PoL, power supply. It is because LDOs are typically easier to integrate on chip since they require active devices only and no passive components to down-convert the voltage [36], [38], [40]–[43]. Moreover, compared to buck converters, LDOs can offer better process and voltage scalability, smaller silicon area overhead, higher rejection capability against the power supply noise, and faster
response times to load current changes. Due to these advantages, state-of-the-art LDOs have been integrated at PoL locations to implement low-cost, low-power distributed power management (Fig. 1(d)) [40]-[44].

However, the power efficiency of LDOs is intrinsically limited by the dropout voltage $V_{DO} = V_{IN} - V_{OUT}$. The power loss from the resistive division increases as the dropout voltage increases [3], [45]. The power efficiency of a LDO is directly related to the dropout voltage as follows:

$$\text{Power Efficiency} = \frac{I_{LOAD}}{I_{LOAD} + I_Q} \times \frac{V_{OUT}}{V_{IN}}$$

Thus, to increase power efficiency, the dropout voltage should be reduced.

Also, the performance of analog LDOs (Fig. 2(b)) is mainly dependent on the gain of the error amplifier. Since the circuits should operate at near-threshold voltage (NTV) levels in modern SoCs, it has become greatly challenging to attain a sufficient EA gain. Analog LDOs’ transient response, line/load regulation, and regulation range are adversely affected by such insufficient EA gains.

Contrarily, DLDOs are suitable to operate at NTV levels, so they have been extensively developed over the past few years [46]-[48]. In addition, the dropout voltage in DLDOs is typically quite small (40 mV - 50 mV) compared to analog LDOs ($\sim$200 mV), thus making DLDOs a better candidate in terms of the maximum achievable power efficiency. Therefore, distributed PDNs (Fig. 1(d)) using multiple small DLDOs are most suitable and largely accepted to highly efficient fine-grained power delivery and management.

### III. CONVENTIONAL PDNs

As discussed above, buck converters are more power-efficient, but they are inappropriate for on-chip integration due to the large physical sizes of the required passive elements. Alternatively, LDOs are more suitable for delivering high-quality power to load circuits with reasonable area overheads. But, they suffer from limited power efficiencies due to their dropout voltages. Hence, PDNs with either switching-mode or linear converters only suffer from either significant area or power overhead. These overheads for each PDN type are further illustrated in Fig. 3. In-package power converters (PSiPs) partially trade off the power and area overheads, but both their area and power overheads are not small enough yet. As shown in Fig. 3, a desirable PDN should have both small power loss and area.

To utilize the advantages of buck and LDO converters simultaneously, a heterogeneous PDN is commonly adopted [11], [12], [19], [20], [22], [24], [45], [49], [50]. As illustrated in Fig. 4, the heterogeneous PDN converts the input power by using off-chip buck converters first and then regulates these converted powers by using on-chip LDOs. This de-coupling of the power conversion from the power regulation lowers the power and area overheads. A more detailed discussion on typical PDNs using off-chip power converters and heterogeneous PDNs is presented below in this section.

### A. TYPICAL PDN USING OFF-CHIP CONVERTERS

A simplified diagram of a typical PDN with off-chip buck converters for powering a mobile SoC is shown in Fig. 5(a). Here the off-chip buck converters supply power directly to individual blocks and multiple processor units of the SoC through on-board power supply rails. Across these power rails, intermittent but large IR drops happen when the multi-core CPUs’ power density increases [21], [51]. These large and fluctuating IR drops worsen the response latency of the PDN to load current changes, resulting in further large voltage drops in the power supply [52]. Such a slow response of the conventional PDN in a scenario of voltage scaling is illustrated in Fig. 5(b).

In addition, to power multi-core CPUs, one off-chip buck converter is typically implemented using a shared power rail, as shown in Fig. 5(a). However, in the shared power
rail, the supply voltage typically remains fixed at a certain voltage level by a single off-chip buck converter. Thus, all the CPU cores that share the power rail have the same power supply level regardless of each core’s workload. This results in wastage of the energy resources, as shown in Fig. 5(c). Hence, the typical PDN using off-chip buck converters results in a reduction of energy efficiency [52], [53].

B. HETEROGENEOUS PDN

To overcome the limitations of the conventional PDNs, heterogeneous PDNs using on-chip LDOs have been extensively investigated [19], [20], [22], [24], [49], [50]. A simplified block diagram of the heterogeneous PDNs for a mobile SoC is shown in 6(a). Here an off-chip buck converter, which is highly power efficient, serves as a master power converter while multiple PoL-embedded LDOs serve as slave regulators dedicated to each CPU core or block. As shown, dedicated LDOs are integrated to supply a precise ripple-free voltage to each CPU core or block. Each of the dedicated LDOs produces and varies the required voltage for each CPU depending on the workload condition, thus saving a significant amount of energy [24] as shown in Fig. 6(b). In addition, because these LDOs are embedded at PoL, i.e., near the target circuit, the IR drop is significantly reduced. Hence, a fast and precise voltage scaling can be implemented more easily for each CPU core or block, unlike the typical PDN using off-chip converters [49], as shown in Fig. 6(b) and (c).

However, even with substantial research and development in on-chip power delivery and power management, there are still several challenges to address further to achieve energy-efficient fine-grained power delivery and management. These design considerations and state-of-the-art power delivery and management techniques are discussed in the next section.

IV. STATE-OF-THE-ART PDNs

A. DISTRIBUTED HETEROGENEOUS PDN

The heterogeneous PDN can optimize the performance and power consumption for each core of modern processors to some extent while supplying power to multiple voltage domains, as shown in Fig. 6. But, the IR drops induced across the on-chip parasitic resistances have become substantially large due to the increase in the power density of microprocessors over the past decade [54]. These IR drops can no longer be ignored during the design process in the ultra-deep-submicron CMOS process, where the operating voltages are already very low. Energy losses and fluctuations associated with the IR drops are expected to be worse in the future as workloads in the CPU cores for cognitive computing and artificial intelligence becomes more heterogeneous and specialized [54]. To guarantee the performance in such workload conditions, the minimum supply voltage for the core must be larger than a critical voltage. Therefore, it requires a guardband above the critical voltage, causing an additional large power penalty [25], [54].

In addition, the power efficiency of PDNs can also be degraded by the poor efficiency of LDOs owing to their large dropout voltages. Moreover, on-chip LDOs may also suffer from a slow loop latency, which results in a severe voltage drop during peak load transients, further requiring a larger V_{DD} guardband. To address these challenges of the heterogeneous PDNs, a heterogeneous PDN with multiple distributed digital LDOs is proposed [54]. Digital LDOs can maintain a low dropout voltage of 40 - 50 mV, offering a higher power efficiency as compared to their analog counterpart. A simplified block diagram of an exemplary heterogeneous distributed PDN is shown in Fig. 7. Instead of using a single large LDO, each core embeds 9 LDOs, all supplied by the common off-chip buck converter. As shown, these multiple LDOs are connected like a power grid to supply power to the core and operate in a cooperative fashion to significantly reduce IR drops compared to typical heterogeneous PDNs. Because the LDOs are placed in close proximity to the load, noises incurred by the parasitic elements across the power lines can be reduced as well.

Furthermore, by sharing the information with neighboring LDOs, the PDN can significantly enhance the transient response [54]. Fig. 8 illustrates the cooperative regulation scheme, which can enhance the transient response in two
aspects. First, when a large-load-current transient occurs anywhere in the power grid, the load current does not flow from the most adjacent LDO only but also from other neighboring LDOs to quickly mitigate the voltage drop. Second, in each local LDO, the comparison information about the local output voltage is shared with neighboring LDOs. Upon receiving this information, the neighboring LDOs can attentively tune the voltage in consideration of both the shared information and the local voltage. Using this distributed heterogeneous PDN, the overall voltage drop was reduced by more than 66%, and a sub-ns transient response time was achieved against more than 500 mA of load current [54].

B. FULLY INTEGRATED RECONFIGURABLE PDN (RPDN)

As an efficient PDN for complex multi-core processors, per-core FIVRs with on-chip inductors and multiple local power gates (PGs) were demonstrated [15], [20], [28], [55]. The FIVRs can offer high-current high-power-density fine-grained power management with a fast transient response because they can operate on a high-speed switching frequency \( F_{SW} \). For instance, in the PDN [15] for an Intel processor, the power conversion stage is implemented on the motherboard, while the power regulation comprises 31 FIVRs. These FIVRs are synchronous multiphase buck converters while operating at a fast \( F_{SW} = 140 \text{ MHz} \) and up to 16 phases. The FIVRs in [15] achieved a current density of 1.3 A/mm\(^2\) and a power efficiency of 88%, while driving four voltage domains of the core with a total load current of 50 A at \( F_{SW} = 140 \text{ MHz} \). To achieve such performances, however, the integration of the air-core inductors (16 inductors for 16 phases) requires a massive silicon area. Moreover, as the aggressive CMOS scaling continues, each core shrinks, and the number of cores per die increases. Hence, the required number of inductors is also increased. On the contrary, each inductor’s required footprint does not scale down, posing a severe challenge in integrating these inductors [56]. Therefore, the FIVRs are not suitable to achieve efficient power delivery and power management in multi-core SoCs because of their huge area overhead.

To overcome the area constraints of the PDN using FIVRs only, a fully integrated autonomous reconfigurable PDN (RPDN) was recently proposed [56]. As shown in Fig. 9, this RPDN reduces the number of FIVRs to just two. Each FIVR is dynamically shared among four co-located cores via per-core DLDOs [56]. Each per-core DLDO consists of a local LDO control (LLC) block, which controls two local power gates (PGs), as shown in Fig. 9. With these techniques, the 2-input/4-output RPDN shown in Fig. 9 delivers a better performance on demand while maximizing the overall energy efficiency along with better core count and scalability [56]. However, the fully integrated RPDN using on-chip inductors still occupies a significant valuable area on the package, conflicting with the stringent requirements of modern enterprise microprocessors [55]. Moreover, due to its limited achievable on-chip inductance, the peak output power of the FIVRs is also restricted.

Based on this fully integrated RPDN with 2-input/4-output FIVRs, a fully-integrated QOESC converter was proposed in [57]. The top-level structure of the PDN using the QOESCs for 4 CPU cores is shown in Fig. 10. The PDN consists of a switched-capacitor (SC) power converter and four LDOs. The SC converter comprises 32 equal partitions of capacitors and switches, operating as a power conversion stage. Each of the LDOs operates as a regulator for each core. This SC design uses an extended binary (EXB) scheme, which uses two flying capacitors to produce three ratios, each with
1/4 resolution. For each output ratio, 32 time-interleaved phases are generated with equal partitions to reduce the output voltage ripples at the SC output. While each core’s power is supplied by its own LDO, which is a phase-locked DLDO in this case. The QOESC routes the power on demand by sharing the total capacitance of the SC network across all the four cores and delivering power to each core in a time-interleaved manner. For example, if the current demand of a particular core increases, more resources (capacitors and switches of the SC network) are dynamically allocated to the core. If the power demand increases further, the corresponding SCs are dynamically configured to change the conversion ratio to provide a higher output voltage. The QOESC with this resource-sharing technique achieved a power efficiency of >87% and 2.5× core frequency range enhancement. However, QOESC’s load-current driving capability per core is just up to 5 mA, which is not sufficient for a typical high current demand of multi-core processors.

V. UVFRs FOR FINE-GRAINED DVFS

The fully integrated heterogeneous PDNs using distributed LDOs [28], [55]–[57] increase the energy and power efficiencies of the multi-core processors. However, as the number of cores in a microprocessor increases due to an increase in computing demands, the performance is often limited by the system’s TDP rather than the total number of the cores, which can be integrated into the processor chip [55], [58]. To achieve optimal use of computing resources, high power efficiency is of paramount importance. In power delivery and management, dynamic voltage and frequency scaling (DVFS) is a well-established method for dynamic thermal management. It is extensively utilized to increase the power efficiency in multi-core processors and SoC platforms [59], [60]. This design strategy is more needed for cases with different workloads. For example, some cores on the processor may not be required to operate at the maximum speed. Then, applying the same supply voltage to all cores would be wasteful.

In addition to TDP, variations of the device, circuit, and system parameters degrade the performances of processor cores in an SoC [61]. These variations are categorized into static and dynamic parameter variations. The static parameter variations occur during manufacturing processes. The parameters are different across dies, but the variations are static over time [61]. Contrarily, the dynamic parameter variations happen over time during processor operations as environmental and workload conditions change. The dynamic parameter variations include supply voltage \(V_{DD}\) drops, temperature-dependent variations, transistor aging, and processor power variations due to workload fluctuations [62]. A key feature of the dynamic parameter variations is the time scale, over which the parameter changes meaningfully. An example of slow-changing variations is the temperature fluctuation, while the \(V_{DD}\) drop is a representative example of fast-changing variations [62].

In commercial processors, static parameter variations can be easily taken care of. Because these variations occur during the manufacturing process, they can be easily detected. For that purpose, each processor is tested using either the maximum clock frequency \(F_{MAX}\) at a constant supply voltage \(V_{DD}\) or using the minimum \(V_{DD}\) at a constant clock frequency \(F_{CLK}\) across multiple DVFS conditions per die. These \(F_{MAX}\) or \(V_{min}\) tests enable to adapt \(F_{CLK}\) and \(V_{DD}\) per die to compensate the static variations. In contrast, the dynamic parameter variations require the processor to either operate at a \(F_{CLK}\) lower than the \(F_{MAX}\) for a target \(V_{DD}\) (i.e., \(F_{CLK} \text{guardband}\)) or at a \(V_{DD}\) higher than the \(V_{MIN}\) for a target \(F_{CLK}\) (i.e., \(V_{DD} \text{guardband}\)) [62]. Due to these guardbands, the processor cannot exploit the opportunities for higher performance by increasing \(F_{CLK}\) or for lower energy by reducing \(V_{DD}\). Hence, their performance and energy efficiency are compromised. Furthermore, the guardband is often required to be increased further due to the wide DVFS ranges of state-of-the-art processors.

These guardbands cannot be reduced sufficiently by the voltage-referenced voltage regulators, which are discussed so far in this paper. These regulators require on-die parameter monitoring circuits and an adaptive control circuit inside the
To address these challenges, variation-aware voltage and frequency regulators have been proposed over the past few years [62]–[68]. Due to their main operation principle, i.e., regulating the voltage and frequency simultaneously, they are broadly known as UVFRs. UVFRs can aggressively reduce the guardband by adapting the voltage and frequency together to the supply voltage \( V_{DD} \) and temperature variations. Because UVFRs continuously track not only \( V_{DD} \) but also the temperature variations, the voltage margin that should be added due to variations in other conventional systems can be taken away even at low \( V_{DD} \), improving the power efficiency of the SoC. Conventional and state-of-the-art topologies of UVFRs for fine-grained DVFS and variation-aware adaptive voltage and frequency scaling are discussed in the following subsections.

### A. TYPICAL VOLTAGE AND FREQUENCY REGULATORS FOR DVFS

An exemplary DVFS system based on two independent loops for the voltage and frequency (\( V_{REG} \) and \( F_{REG} \)) is shown in Fig. 12. The voltage regulation loop, which can be implemented by a buck converter, switched-capacitor converter, or LDO, generates \( V_{REG} \) the regulated voltage based on a reference voltage (\( V_{REF} \)), as shown in Fig. 12(a). In the frequency regulation loop, a phase-locked loop (PLL) is usually employed to regulate the clock frequency (\( F_{REG} \)) based on a reference clock (\( F_{REF} \)). In this system, the two loops are independent and don’t affect each other. Thus, in response to any voltage drop in \( V_{REG} \), which are commonly caused by load current transients, \( F_{REG} \) is not adapted accordingly as shown in Fig. 12(b). Then, this occurrence degrades the timing slack and may cause a failure of timing margin in the processor [62]–[64]. To avoid such issues, a voltage guardband should be introduced. But it is gravely undesirable because of its overhead in power consumption, design complexity, and area [62], [64].

### B. STATE-OF-THE-ART UVFRs

To minimize the guardband, it was proposed recently to combine the \( V_{REG} \) and \( F_{REG} \) regulation loops into a single unified loop. This loop can be based on an LDO [65], switched capacitor [64], or buck converter [66]–[68]. Buck- and LDO-based UVFRs are shown in Fig. 13(a) and (b). In both schemes, the basic operation principle of the unified \( V_{REG} \) and \( F_{REG} \) regulation is almost identical. As illustrated, the UVFR systems generate \( F_{REG} \) from a TRO. Based on \( F_{REG} \), \( V_{REG} \) is regulated and supplied to the digital logic load, which is typically a processor. At the same time, \( V_{REG} \) is also supplied to the TRO. This single control loop implements an adaptive clocking by using the \( V_{REG} \)-powered TRO. The TRO has a replica path for the load circuit’s critical path, mimicking both the logic delay and the voltage drop of the critical path [64]–[66]. As a result, when any slow-down of signal progression happens in the critical path due to either \( V_{REG} \) drops or other dynamic variations, the clock also slows down accordingly. Such clock stretching (slow-down of the clock) ensures avoiding any failure in the timing margin. Similarly, in the events of voltage overshoots, the clock speeds up accordingly, maintaining the timing margins in the processor, as shown in Fig. 11, to measure specific dynamic parameters (e.g., temperature, supply voltage, and aging) and adjust \( F_{CLK} \) and \( V_{DD} \) accordingly to compensate the dynamic parameter variations. This technique has several drawbacks. One of the most significant ones is the area and cost overheads caused by the additional sensors and circuits. Another one is that it requires a long recovery time to mitigate the fast-changing dynamic parameter variations like fast-switching load currents, which causes severe supply voltage \( V_{DD} \) drops.
that UVFRs require load circuits by 27%. Furthermore, it should also be noted the UVFR in [65] reduced the overall supply voltage of the PDN using frequency-referenced DLDOs. A simplified block diagram of the proposed heterogeneous PDN is shown in Fig. 14. As shown, a global clock is utilized as the reference (BGR) for each FR-DLDO. Hence, multiple BGRs are the output voltage and frequency (VREF and FREQUENCY-REFERENCED DLDO) for the designated load circuit and the local output voltage and frequency (FCi) depending on the workload demand of the load circuit. VREF and FREQUENCY-REFERENCED DLDO generate both VREF and FCi, which may happen due to different workload conditions, and adjusts (slow-down or speed-up) FREG accordingly. After meeting these workload

VI. PROPOSED HETEROGENEOUS PDN USING FREQUENCY-REFERENCED DLDOs

Considering the advantages of UVFRs, we propose a modified heterogeneous PDN that uses FR-DLDOs for fine-grained DVFS. A simplified block diagram of the proposed heterogeneous PDN is shown in Fig. 14. As shown, an off-chip buck converter is utilized as a master power supply to multiple frequency-referenced DLDOs, which are all integrated inside the SoC. The FR-DLDOs operate as slave power supplies dedicated to each load circuit block. They are fully integrated right at the PoL locations. In the proposed PDN, the FR-DLDOs do not require VREF from a BGR circuit. Instead, a global clock is utilized as the reference frequency (FREF) for each FR-DLDO. Hence, multiple BGRs required for typical LDOs and PDNs, are no longer needed in the proposed heterogeneous PDN, thereby significantly reducing power distribution rails, board area, and component counts. Moreover, each FR-DLDO sets the local supply voltage (VCI) for the designated load circuit and the local output frequency (FCi), depending on the workload demand of the load circuit. VCI and FCi are the output voltage and frequency from each FR-DLDO, and they are used as the local power supply and the operating clock for the designated local load circuit, respectively. The FR-DLDO generates and regulates the fast switching current demands of SoC load circuits. The UVFR in [65] has an enhanced loop response time and a faster load-transient response time (TR). But its current driving capability is also minimal at 6 mA. To overcome these limitations of the FR-DLDOs [65], [69], a fast-transient FR-DLDO with a large current-driving capacity and wide regulation range is presented in the next section [63]. The FR-DLDO achieves a faster transient response time due to its transient-boost control. In addition, its 10-bit binary-weighted PMOS switch array can drive a large load current up to 40 mA.

VII. IMPLEMENTATION EXAMPLE OF FREQUENCY-REFERENCED DLDO

A. OVERALL ARCHITECTURE AND OPERATION

Fig. 15 shows a block diagram of the implemented FR-DLDO (presented in [63]) for the proposed heterogeneous PDN. The FR-DLDO generates both VREF and FREF by using a single regulation loop with a TRO, as shown in Fig. 15. This single unified loop tracks the change in VREF, which may happen due to different workload conditions, and adjusts (slow-down or speed-up) FREF accordingly. After meeting these workload

FIGURE 14. Simplified block diagram of the proposed heterogeneous PDN using frequency-referenced DLDOs.

FIGURE 15. Block diagram of the implemented frequency-referenced DLDO.
conditions through multiple iterations of the feedback loop, the loop is put in the steady state when $F_{\text{REG}} \approx F_{\text{DIV}}$. In a typical FR-DLDO [69], a typical gated voltage-controlled oscillator (VCO) is utilized in a multiplying delay-locked loop (DLL). The VCO does not model the critical path of the target load circuit in terms of the workload requirement. Thus, this FR-DLDO’s output voltage scaling is independent of the frequency scaling, and it cannot optimize $V_{\text{REG}}$ and $F_{\text{REG}}$ according to the variations in the load circuit. Contrarily, in the proposed FR-DLDO, the TRO is powered by $V_{\text{REG}}$, generating the core clock $F_{\text{REG}}$. Thus, the clock intrinsically adapts to the voltage ($V_{\text{REG}}$) and temperature variations while variations in the load circuit’s critical path delays are compensated. As a result, the timing margins are maintained nearly constant in the implemented FR-DLDO [63]. However, the critical path modeling of the TRO is totally load-circuit-specific. If the load circuit changes, the TRO should be reconfigured. In addition, the FR-DLDO should be integrated with the close proximity of the load circuits on the same chip to ensure a good matching between the critical path delay of the load circuits and the TRO under PVT variations.

In the implemented FR-DLDO shown in Fig. 15, the phase difference between the two incoming clocks $F_{\text{REG}}$ and $F_{\text{DIV}}$ is detected by a bang-bang phase frequency detector (BBPFD). The phase-difference outputs ($UP/DN$) of the BBPFD are fed to a digital loop filter (DLF). The DLF operates as both a low-pass filter and a loop controller. The filter counts up or down according to the polarity of $UP/DN$ and accordingly turns on/off PMOS power transistors by controlling a 10-bit output signal (SW[9:0]). In the proposed DLF, the up/down counter does not reset to zero but keeps the previous counter value at the end of the reference cycle. By controlling SW[9:0], $V_{\text{REG}}$ is regulated to the level required for the TRO to generate $F_{\text{REG}}$ in order to lock the loop. Here the divider value ($\div N$) sets the frequency of $F_{\text{REG}}$ and the level of $V_{\text{REG}}$. The relationship between two frequencies is $F_{\text{REG}} = N \times F_{\text{REF}}$. When the target frequency is reached at $F_{\text{REG}} = N \times F_{\text{REF}}$, the phase and frequency of both $F_{\text{REG}}$ and $F_{\text{DIV}}$ (the divided frequency of $F_{\text{REG}}$) get locked. $V_{\text{REG}}$ is now settled to the level where the TRO generates the target $F_{\text{REG}}$. Under this $V_{\text{REG}}$, the load circuit can operate without a timing error at $F_{\text{REG}}$. Contrarily, when $F_{\text{DIV}} \neq F_{\text{REG}}/N$ is slower than $F_{\text{REF}}$, the loop increases $V_{\text{REG}}$, thus increasing the frequency of the TRO until $F_{\text{DIV}}$ becomes identical enough to be locked with $F_{\text{REF}}$. Hence, this unified $V_{\text{REG}}$ - $F_{\text{REG}}$ control loop allows the FR-DLDO to accommodate two goals, i.e., 1) adjusting $F_{\text{REG}}$ in response to any supply ($V_{\text{REG}}$) drop or dynamic variations in a short term, and 2) simultaneously regulating the clock ($F_{\text{REG}}$) to a fixed reference ($F_{\text{REF}}$) in a longer term. Simplified block diagram of the implemented FR-DLDO is shown in Fig. 16(a). The FR-DLDO generates a tightly coupled pair of $V_{\text{REF}}$ and $F_{\text{REG}}$. $V_{\text{REG}}$ serves as the power supply of the TRO, which generates $F_{\text{REG}}$. The overall behavior of the FR-DLDO can be modeled in small signal as shown in Fig. 16(b). $V_{\text{REG}}$ and $F_{\text{REG}}$ can be estimated from the small-signal model as follows:

$$V_{\text{REG}} = \frac{K_{PD} \cdot K_{PMOS}}{s^2 + p_1 s + K_{PMOS} \cdot K_{TRO}/N}$$  \hspace{1cm} (2)$$

$$F_{\text{REG}} = \frac{K_{PD} \cdot K_{PMOS} \cdot K_{TRO}}{s^2 + p_1 s + K_{PMOS} \cdot K_{TRO}/N}$$  \hspace{1cm} (3)$$

where $K_{PD}$ is the gain of BBPD, $K_{PMOS}$ is the gain of PMOS output stage in V/рад, $K_{TRO}$ is the gain of the tunable-replica oscillator in Hz/V, and $p_1$ is the output pole frequency in Hz.

As Fig. 16(b) and the equations show, the FR-DLDO is a second-order system consisting of 1) a pole at DC due to the accumulator operation and 2) the load-dependent output pole due to $C_{OUT}$ and the load. We can get the steady-state open-loop transfer function $O(s)$ in s domain as follows:

$$O(s) = H_0 \cdot \frac{e^{-s/\tau_{ref}}}{s} \cdot \frac{1 + s/z_1}{1 + s/p_1}$$  \hspace{1cm} (4)$$

$p_1$ and $z_1$ at the output node are given as follows:

$$p_1 = \frac{1}{C_L \cdot (R_{ESR} + 1/g_{ds} \parallel R_L)}$$  \hspace{1cm} (5)$$

$$z_1 = \frac{1}{C_L \cdot R_{ESR}}$$  \hspace{1cm} (6)$$

where $H_0$ is the loop gain which mainly depends on the gains of $K_{PD}$, $K_{PMOS}$, and $K_{TRO}$, $R_{ESR}$ represents the equivalent series resistance of $C_L$, and $g_{ds}$ is the sum of the output conductance of PMOS transistors at steady state. The FR-DLDO uses a minimal value of $C_L$ as 10 pF, which makes $p_1$ and
TABLE 1. Performance comparison of the proposed FR-DLDO with state-of-the-art frequency-referenced power converters.

| Topology | Our Work [63] | [65] | [69] | [68] |
|----------|---------------|------|------|------|
| Process [nm] | FR-DLDO | FR-DLDO | FR-DLDO | FR-Buck |
| V_{DD} [V] | 0.6 – 1.2 | 0.6 – 1.0 | 0.38 | 1.5 – 1.8 |
| F_{REF} [MHz] | 37.4 (nominal) | 0.1 – 500 | 5 | 10 |
| V_{REG} [V] | 0.56 – 1.16 | 0.38 – 0.81 | 0.12 – 0.32 | 0.6 – 1.0 |
| Max. I_{LOAD} [mA] | 40 | 6 | 1 | 400 |
| C_{OUT} [nF] | 0.01 | 0.2 | 100 | 1000 |
| ΔV_{REG} [mV] @ ΔI_{LOAD} [mA] | 133 @ 37 | 163 @ 3 | 70 @ 0.2 | 38 @ 123 |
| T_{R} [ns] | 400 | 200 | 100,000 | 3500 |
| I_{Q} [μA] | 147.5 – 634 | 36 – 330 | 6.85 | N. R |
| Load Regulation [mV/mA] | 0.014 | < 1.8 | 0.11 | N. R |
| Current Efficiency [%] | 99.5 @ 0.56 V | 99.4 @ 0.38 V | 99.3 | N. R |
| Peak Power Efficiency [%] | 93 @ 0.56 V | 73 @ 0.81 V | 84 @ 0.32 V | 91.9 |
| FOM [ps] | 0.62 | 130 | 1198750 | N. R |
| Area [mm²] | 0.02 | 0.02 | 0.026 | 0.236 |

FOM = \frac{C_{OUT} \times \Delta V_{REG} \times I_{Q}}{\Delta I_{LOAD}}

Peak Power Efficiency = \frac{I_{LOAD} \times I_{Q} \times V_{REG}}{V_{DD} \times 100}

1) TRANSIENT-BOOST CONTROL

To enhance the load transient response and reduce the voltage undershoot (ΔV_{REG}) for load current step changes, a transient-boost control was proposed [63]. Its circuit diagram is shown in Fig. 17(a). It continuously measures the time difference between the two incoming clocks, F_{REF} and F_{DIV}, by observing the BBPFD outputs (UP and DN). When F_{DIV} ≈ F_{REF}, the lock signal (Lock) becomes high, and the

z\_1 to be placed at very high frequencies. Then, O(s) can be approximated as a first-order system, the loop is stable.

FR-DLDO operates in the steady state. Any deviation in V_{REG} during the steady state changes the polarity of Lock from high to low, activating the transient-boost mode.

Such deviations typically happen when V_{REG} experiences a load current step, inducing a voltage undershoot. Then, the transient-boost control circuit generates a boosted clock F_{DLF} (=16 × F_{REF}) by using 16 delay cells, 1-shot pulse, and digital logic, quickly recovering V_{REG}. Once V_{REG} recovers from the undershoot, the loop control switches back to the normal loop operation with Lock = 1. The load transient response with the transient-boost control is illustrated in Fig. 17(b). As shown, the boosted clock significantly reduces the peak of voltage undershoot ΔV_{REG} and the transient response time (T_{R}) by shortening the feedback loop response time (t_{RES}). It is because the peak of ΔV_{REG} mainly depends on the t_{RES} with a given load current step ΔI_{LOAD} and the output capacitor (C_{L}) [45]. ΔV_{REG} can be roughly calculated as follows:

ΔV_{REG} ≈ \frac{ΔI_{LOAD} \times t_{RES}}{C_{L}} (7)
The proposed transient-boost control significantly shortens $t_{RES}$, thus improving the load transient response in terms of $\Delta V_{REG}$ and $T_R$.

2) RESULTS AND MEASUREMENTS

The FR-DLDO was designed and fabricated in a 40-nm CMOS process, occupying an active area of 0.02 mm$^2$, as shown in Fig. 18. At a nominal $V_{DD} = 1.2$ V and $F_{REF} = 37.4$ MHz, the FR-DLDO achieved a wide voltage and frequency regulation ranges of 0.56 V - 1.16 V and 112.2 MHz to 1.91 GHz, respectively. These regulation ranges were acquired by varying the divider value $N$ between 3 and 51. In addition to $N$, the reference frequency of the proposed FR-DLDO can be varied from 10 MHz to 100 MHz. The FR-DLDO consumes 147.5 $\mu$A of the quiescent current ($I_Q$) when supplying $V_{REG}$ and $F_{REG}$ of 0.56 V and 112.2 MHz, respectively.

The measured load transient responses of the proposed FR-DLDO with and without using the proposed transient-boost control at $V_{REG} = 1.1$ V, $F_{REF} = 37.4$ MHz, and the output capacitor $C_{OUT} = 10$ pF are shown in Fig. 19. As shown in Fig. 19(a), for a load current step of 37 mA changing from 3 mA to 40 mA with an edge time ($T_{EDGE}$) of 25 ns, the FR-DLDO recovered $V_{REG}$ with $\Delta V_{REG}$ of 300 mV and settles within 5 $\mu$s of the transient response time ($T_R$) without using the transient-boost control. On the contrary, with the assistance of the transient-boost control, $\Delta V_{REG}$ was reduced to 133 mV, and $V_{REG}$ was fully recovered and settled within 400 ns of $T_R$, as shown in Fig. 19(b). The proposed transient-boost control reduced $\Delta V_{REG}$ and $T_R$ by 55.6%, and 92%, respectively. In addition, the FR-DLDO achieved a load regulation of 0.014 mV/mA while driving a load current of 40 mA. The performance summary of the FR-DLDO and other state-of-the-art FR-DLDOs is shown in Table 1. The proposed FR-DLDO outperforms other state-of-the-art FR-DLDOs [65], [69] and frequency-referenced buck converter [67] in the regulation range, peak power efficiency, and figure-of-merit (FOM).

VIII. CONCLUSION

PDNs and power management techniques have gained prime importance in SoC designs due to the integration of diverse heterogeneous circuits and systems on a single chip, requiring multiple voltage domains and high-quality fast power delivery. To meet these challenging power delivery and management demands, PDNs and power management systems are required to perform intelligent, energy-efficient, fine-grained, and dynamically controlled on-chip power management. The shortcomings of the typical PDNs using off-chip power converters have been well overcome over the years by using state-of-the-art techniques of PSiP and on-chip integration of power converters. On-chip integration of just a few power converters is no longer enough to address highly-efficient fine-grained power management challenges. However, on-chip integration of a large number of power converters drastically increases the occupied silicon area. Therefore, state-of-the-art PDNs with heterogeneous structures have been utilized as a good alternative to fully-on-chip integrated converters. In the heterogeneous PDNs, highly power-efficient switching-mode power converters (buck) are utilized as off-chip converters, and multiple small-sized LDOs are integrated on chip at PoL locations inside the SoC. This technique delivers highly precise on-demand voltages to multiple voltage domains of the SoC. The state-of-the-art PDNs, including the distributed heterogeneous PDN, fully-integrated reconfigurable PDN, and quad-output elastic switched-capacitor PDN using per-core digital LDOs, are discussed thoroughly in the paper. These PDNs enable fine-grained power management with high peak-power efficiencies and high current densities for multi-core processors. In addition to these PDNs, frequency-operated UVFRs have gained a lot of attention over the past few years due to their capabilities of variation-aware fine-grained DVFS in
multi-core processors. The UVFRs can achieve on-the-fly DVFS over PVT variations without interrupting the timing margins of load circuits. The UVFRs help achieve high energy efficiencies in the load circuits thanks to their simultaneous voltage and frequency scaling, which is tailored to the load circuit’s workload demand. Based on the UVFR operation, we proposed a distributed heterogeneous PDN with FR-DLDOs to achieve fine-grained DVFS in an SoC. In the proposed PDN, a buck converter is utilized as a master power supply to multiple FR-DLDOs, which are integrated inside the SoC as slave power supplies dedicated to each load circuit. An implementation of FR-DLDO for the proposed heterogeneous PDN is presented. A transient-boost control is used in the FR-DLDO to mitigate the transient response time and voltage drops dynamically. It is activated only when the regulated output voltage faces any deviation during the steady state. The 0.02-mm² FR-DLDO prototype fabricated in 40 nm CMOS process outperforms other FR-DLDOs by achieving a minimum figure-of-merit, maximum peak power and current efficiencies, and wide regulation range.

The distributed heterogeneous and fully-integrated PDNs are expected to be adopted more widely because distributed digital LDOs in these PDNs can supply a uniform power to each core with a small IR drop and help the thermal management, as proposed in a recent design [70]. Furthermore, with the advancement of synthesizable digital LDOs for distributed PDNs, PDN designs will become easier and process scalable. Some synthesized digital LDOs were recently demonstrated with promising performance metrics [71], [72]. In addition to the distributed PDNs, the synthesizable design of UVFR is recently adopted due to its scalability [67]. Moreover, buck-converter-based UVFRs have been also emerging because they are capable of producing multiple voltage outputs with a single voltage only, i.e., single input, multiple output (SIMO). A SIMO buck converter with an adaptive-clocking scheme for UVFR operation was proposed, achieving significant performance gains in terms of the voltage margin reduction and power efficiency enhancement [73]. Over the decades, PDNs for large power delivery have been improved significantly in terms of power efficiency, cost, area, and many other factors. In addition to this, there has been another recent trend for PDNs, i.e., a continuously growing demand for ultra-low-power and low-cost power delivery and management systems for wearable and IoT devices.

REFERENCES

[1] N. Zlatanov, “CPU vs. SoC—The battle for the future computing,” in Proc. Int. Syst.-Chip Conf., 2015.
[2] M. Hammes, C. Kranz, D. Seippel, J. Kissing, and A. Leyk, “Evolution on SoC integration: GSM baseband-radio in 0.13 μm CMOS extended by fully integrated power management unit,” IEEE J. Solid-State Circuits, vol. 43, no. 1, pp. 236–245, Jan. 2008.
[3] M. A. Akram, “Fine-grained power delivery and management with all-digital capacitor-less low-dropout regulators,” Ph.D. dissertation, Dept. Elect. Electron. Eng., Kangwon Natl. Univ., Chuncheon-si, South Korea, 2019, doi: 10.1341/002.2.10671.15528.
[4] X. Zhou, P.-L. Wong, P. Xu, F. C. Lee, and A. Q. Huang, “Investigation of candidate VRM topologies for future microprocessors,” IEEE Trans. Power Electron., vol. 15, no. 6, pp. 1172–1182, Nov. 2000.
[5] R. Mirfakhridinov, “An analytical comparison of alternative control techniques for powering next-generation microprocessors,” in Proc. Texas Instrum. Seminar, 2001.
[6] T. Singh, A. Schaefer, S. Rangarajan, D. John, C. Henrion, R. Schreiber, M. Rodriguez, S. Kosonocky, S. Naftziger, and A. Novak, “Zen: An energy-efficient high-performance x86 core,” IEEE J. Solid-State Circuits, vol. 53, no. 1, pp. 102–114, Jan. 2018.
[7] C. Tokunaga, J. F. Ryan, C. Augustine, J. P. Kulkarni, Y.-C. Shih, S. T. Kim, R. Jain, K. Bowman, A. Raychowdhury, M. M. Khellah, J. W. Tschanz, and V. De, “A graphics execution core in 22 nm CMOS featuring adaptive clocking, selective boosting and state-retentive sleep,” in IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers, Feb. 2014, pp. 108–109.
[8] S. B. Nasir, “Fine-grain on-chip power management using digital and digitally-assisted linear voltage regulators,” Ph.D. dissertation, School Elect. Comput. Eng., Georgia Inst. Technol., Atlanta, GA, USA, 2017, p. 149. [Online]. Available: http://hdl.handle.net/1853/60708.
[9] D. E. Lackey, P. S. Zachow, T. R. Bednar, D. W. Stout, S. W. Gould, and J. M. Cohn, “Managing power and performance for system-on-chip designs using voltage islands,” in Proc. IEEE/ACM Int. Conf. Comput. Aided Design (ICCAD), 2002, pp. 195–202.
[10] W. Godycki, C. Tong, I. Bukreyev, A. Apsel, and C. Batten, “Enabling realistic fine-grain voltage scaling with reconfigurable power distribution networks,” in Proc. 47th Annu. IEEE/ACM Int. Symp. Microarchitecture, Dec. 2014, pp. 381–393.
[11] I. P. Vaisaband, R. Jakushokas, M. Popovich, A. V. Mezhiba, S. Kose, and E. G. Friedman, On-Chip Power Delivery and Management, Cham, Switzerland: Springer, 2016, p. 742.
[12] I. P. Vaisaband, “Power delivery and management in nanoscale ICs,” Ph.D. dissertation, Dept. Elect. Comput. Eng., Univ. Rochester, Rochester, NY, USA, 2015, p. 251. [Online]. Available: http://hdl.handle.net/1802/30066.
[13] K. Onizuka, K. Inagaki, H. Kawaguchi, M. Takamiyi, and T. Sakurai, “Stacked-chip implementation of on-chip buck converter for distributed power supply system in SiPs,” IEEE J. Solid-State Circuits, vol. 42, no. 11, pp. 2404–2410, Nov. 2007.
[14] R. Foley, P. Waldron, J. Slowey, A. Alderman, B. Narveson, and S. C. O’Mathuna, “Technology roadmaping for power supply in package (PSP) and power supply on chip (PwrSoCs),” in Proc. 25th Annu. IEEE Appl. Power Electron. Conf. Expo. (APEC), Feb. 2010, pp. 525–532.
[15] E. A. Burton, G. Schrom, F. Paillet, J. Douglas, W. J. Lambert, R. Radhakrishnan, and M. J. Hill, “FIVR—Fully integrated voltage regulators on 4th generation intel core SoCs,” in Proc. IEEE Appl. Power Electron. Conf. Expo. (APEC), Mar. 2014, pp. 432–439.
[16] N. Sturcken, E. O’Sullivan, N. Wang, P. Herget, B. Webb, L. Romankiw, M. Petracca, R. Davi, C. Fontana, G. Decad, I. Kymissis, A. Peterchey, L. Carloni, W. Gallagher, and K. Shepard, “A 2.5D integrated voltage regulator using coupled-magnetic-core inducers on silicon interposer delivering 10.8A/mm²,” in Proc. IEEE Int. Solid-State Circuits Conf. (ISSCC), Feb. 2012, pp. 400–402.
[17] H. Jia, J. Lu, X. Wang, K. Padmanaban, and J. Z. Shen, “Integration of a monolithic buck converter power IC and bondwire inducers with ferrite epoxy glob cores,” IEEE Trans. Power Electron., vol. 26, no. 6, pp. 1627–1630, Jun. 2011.
[18] Y. Ahn, H. Nam, and J. Roh, “A 50-MHz fully integrated low-swing buck converter using packaging inducers,” IEEE Trans. Power Electron., vol. 27, no. 10, pp. 4347–4356, Oct. 2012.
[19] Z. Toprak-Deniz, M. Sferling, M. Bulzacchelli, G. Still, R. Kruse, S. Kim, D. Boerstorfer, T. Gloekler, R. Robertazzi, K. Stawiasz, T. Diemont, G. English, D. Hui, P. Muersch, and J. Pach, “4.2 distributed system of digitally controlled microregulators enabling per-core DVFS for the POWER8 microprocessor,” in Proc. IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers, Feb. 2014, pp. 98–99.
[20] E. J. Fluhr, J. Friedrich, D. Drees, V. Zyuban, G. Still, R. Gonzalez, A. Hall, D. Hogenmiller, F. Malgioglio, R. Nett, J. Paredes, J. Pille, D. Plass, R. Puri, P. Restle, D. Sloan, K. Stawiasz, Z. T. Deniz, D. Wendel, and R. Pach, “5.1 POWER8: 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.
[21] X. Wang, J. Xu, Z. Wang, K. J. Chen, X. Wu, Z. Wang, P. Yang, and L. H. K. Duong, “An analytical study of power delivery systems for many-core processors using on-chip and off-chip voltage regulators,” IEEE Trans. Comput.-Aided Design Integr. Circuits Syst., vol. 34, no. 9, pp. 1401–1414, Sep. 2015.
S. Joo and S. Kim, “Output-capacitor-free LDO design methodologies for high EMI immunity,” IEEE Trans. Electromagn. Compat., vol. 60, no. 2, pp. 497–506, Apr. 2018.

C. Yang, K. Ye, and M. Tan, “A 0.5-V capless LDO with 30-dB PSRR at 10-kHz using a lightweight local generated supply,” IEEE Trans. Circuits Syst. II, Exp. Briefs, vol. 67, no. 10, pp. 1785–1789, Oct. 2020.

Y.-H. Lin, K.-L. Zheng, and K.-H. Chen, “Smooth pole tracking technique by power MOSFET array in low-dropout regulators,” IEEE Trans. Power Electron., vol. 23, no. 5, pp. 2421–2427, Sep. 2008.

M. A. Akram, I.-C. Hwang, and S. Ha, “Architectural advancement of digital low-dropout regulators,” IEEE Access, vol. 8, pp. 137838–137855, 2020.

L. K. Abram, W. Hong, and I.-C. Hwang, “Fast transient fully standard-cell-based all digital low-dropout regulator with 99.97% current efficiency,” IEEE Trans. Power Electron., vol. 33, no. 9, pp. 8011–8019, Sep. 2018.

M. A. Akram, W. Hong, and I.-C. Hwang, “Capacitorless self-locked all-digital low-dropout regulator,” IEEE J. Solid-State Circuits, vol. 54, no. 1, pp. 266–276, Jan. 2019.

M. A. Akram, S.-S. Kim, S. Ha, and I.-C. Hwang, “Output-capacitorless tri-loop digital low dropout regulator achieving 99.91% current efficiency and 2.87 fs FOM,” IEEE Trans. Power Electron., vol. 36, no. 2, pp. 2044–2058, Feb. 2021.

S. K"ose and E. G. Friedman, “Distributed power network co-design with on-chip power supplies and decoupling capacitors,” in Proc. Int. Workshop Syst. Level Interconnect Predict., 2011, pp. 1–5.

J. Vaidhyanathan and E. G. Friedman, “Dynamic power management with power network-on-chip,” in Proc. IEEE 12th Int. New Syst. Conf. Form. (NEWCAS), Jun. 2014, pp. 225–228.

H. Li, J. Xu, Z. Wang, P. Yang, R. K. V. Maeda, and Z. Tian, “Adaptive power delivery system management for many-core processors with on/off-chip voltage regulators,” in Proc. Design, Autom. Test Eur. Conf. Exhib. (DATE), Mar. 2017, pp. 1265–1268.

W. Kim, M. S. Gupta, G.-Y. Wei, and D. Brooks, “System level analysis of fast, per-CORE DVFS using on-chip switching regulators,” in Proc. IEEE 14th Int. Symp. High Perform. Comput. Archit., Feb. 2008, pp. 123–134.

L. T. Clark, E. J. Hoffman, J. Miller, M. Biyani, L. Liao, S. Straszus, M. Morrow, K. E. Velarde, and M. A. Yarch, “An embedded 32-b microprocessor core for low-power and high-performance applications,” IEEE J. Solid-State Circuits, vol. 36, no. 11, pp. 1599–1608, Nov. 2001.

Y. Lu, F. Yang, F. Chen, and P. K. T. Mok, “A distributed power delivery grid based on analog-assisted digital LDOS with cooperative regulation and IR-drop reduction,” IEEE Trans. Circuits Syst. I, Reg. Papers, vol. 67, no. 8, pp. 2859–2871, Aug. 2020.

M. E. Perez, M. A. Sperling, F. J. Balzachelli, Z. Toprak-Deniz, and T. E. Demers, “Distributed networks of LDO micro-regulators providing submicronsecond DVFS and IR drop compensation for a 24-core microprocessor in 14-nm SOI CMOS,” IEEE J. Solid-State Circuits, vol. 55, no. 5, pp. 731–743, Mar. 2020.

Z. K. Ahmed, H. K. Krishnamurthy, S. Weng, X. Liu, C. Schaefer, N. Desai, K. Ravichandran, J. W. Tschanz, and V. De, “An autonomous reconfigurable power delivery network (RPDN) for many-core SoCs featuring dynamic current steering,” in Proc. IEEE Symp. VLSI Circuits, Jun. 2020, pp. 1–2.

S. Gangopadhyay, J. W. Tschanz, and A. Raychowdhury, “A quad-output elastic switched capacitor converter and per-core LDO with 87% power efficiency and 2.5% core-frequency range improvement,” in Proc. IEEE Int. Symp. Circuits Syst. (ISCAS), Oct. 2020, pp. 1–5.

C. Isci, A. Buyuktosunoglu, C.-Y. Cher, P. Bose, and M. Martonosi, “An analysis of efficient multi-core global power management policies: Maximizing performance for a given power budget,” in Proc. 39th Annu. IEEE/ACM Int. Symp. Microarchitectu (MICRO), Dec. 2006, pp. 347–358.

J. Ganci and M. H. Chowdhury, “A hybrid scheme for on-chip voltage regulation in system-on-a-chip (SOC),” IEEE Trans. Very Large Scale Integ. (VLSI) Syst., vol. 19, no. 11, pp. 1949–1959, Nov. 2011.

M. K. Tavana, M. H. Hajsazemi, D. Pathak, J. Savvidis, and H. Homayoun, “Elasticcore: Enabling dynamic heterogeneity with joint core and voltage/frequency scaling,” in Proc. 52nd ACM/EDAC/IEEE Design Autom. Conf. (DAC), Jun. 2015, pp. 1–6.

K. A. Bowman, S. G. Duvall, and J. D. Meindl, “Impact of die-to-die and within-die parameter fluctuations on the maximum clock frequency distribution for gigascale integration,” IEEE J. Solid-State Circuits, vol. 37, no. 2, pp. 183–190, Feb. 2002.
S. Gangopadhyay, S. B. Nasir, A. Subramanian, V. Sathe, and K. A. Bowman, “Adaptive and resilient circuits: A tutorial on improving processor performance, energy efficiency, and yield via dynamic variation,” *IEEE Solid State Circuits Mag.*, vol. 10, no. 3, pp. 16–25, Summer 2018.

M. A. Akram and I.-C. Hwang, “Fully digital fast transient phase-locked digital LDO-embedded MDLL for DVFS applications,” *Anal. Int. Circuits Signal Process.*, vol. 93, no. 1, pp. 123–136, Oct. 2017.

F. U. Rahman, S. Kim, N. John, R. Kumar, X. Li, R. Pamula, K. A. Bowman, and V. S. Sathe, “A unified clock and switched-capacitor-based power delivery architecture for variation tolerance in low-voltage SoC domains,” *IEEE J. Solid-State Circuits*, vol. 54, no. 4, pp. 1173–1184, Apr. 2019.

S. Gangopadhyay, S. B. Nasir, A. Subramanian, V. Sathe, and A. Raychowdhury, “UVFR: A unified voltage and frequency regulator with 500 MHz/0.84 V to 100 KHz/0.27 V operating range, 99.4% current efficiency and 27% supply guardband reduction,” in *Proc. Conf. 42nd Eur. Solid-State Circuits (ESSCIRC)*, Sep. 2016, pp. 321–324.

X. Sun, F. U. Rahman, V. R. Pamula, S. Kim, X. Li, N. John, and V. S. Sathe, “An all-digital fused PLL-buck architecture for 82% average 3σ-margin reduction in a 0.6-to-1.0-v cortex-M4 processor,” *IEEE J. Solid-State Circuits*, vol. 54, no. 11, pp. 3215–3225, Nov. 2019.

X. Sun, A. Boora, R. Pamula, C.-H. Huang, D. Peña-Colaiocco, and V. S. Sathe, “UniCaP-2: Phase-locked adaptive clocking with rapid clock cycle recovery in 65 nm CMOS,” in *Proc. IEEE Symp. VLSI Circuits*, Jun. 2020, pp. 1–2.

F. Atallah, K. Bowman, H. Nguyen, J. Jeong, D. Yingling, Y. Sun, B. Appel, A. Polomik, M. Harinath, J. Morelli, T. Moore, N. Reeves, A. Cassier, and A. Raychowdhury, “A 7 nm all-digital unified voltage and frequency regulator based on a high-bandwidth 2-phase buck converter with package inductors,” in *IEEE Int. Solid-State Circuits Conf. (ISSCC)* Dig. Tech. Papers, Feb. 2019, pp. 316–318.

Y. Kim and P. Li, “A 0.38 v near/sub-Vt digitally controlled low-dropout regulator with enhanced power supply noise rejection in 90 nm CMOS process,” *IET Circuits, Devices Syst.*, vol. 7, no. 1, pp. 31–41, Jan. 2013.

D.-H. Jung, T.-H. Kong, J.-H. Yang, S. Kim, K. Jim, P. Park, M. Choi, and J. Shin, “A distributed digital LDO with time-multiplexing calibration loop achieving 40A/mm² current density and 1 mA-to-6.4 A ultra-wide load range in 5 nm FinFET CMOS,” in *IEEE Int. Solid-State Circuits Conf. (ISSCC)* Dig. Tech. Papers, Feb. 2021, pp. 414–415.

S. Bang, W. Lim, C. Augustine, A. Malavasi, M. Khellah, J. Tschanz, and V. De, “A fully synthesizable distributed and scalable all-digital LDO in 10 nm CMOS,” in *IEEE Int. Solid-State Circuits Conf. (ISSCC)* Dig. Tech. Papers, Feb. 2020, pp. 380–382.

J. Oh, J.-E. Park, Y.-H. Hwang, and D.-K. Jeong, “A 480 mA output-capacitor-free synthesizable digital LDO using CMP-triggered oscillator and droop detector with 99.99 current efficiency, 1.3µs response time, and 9.8A/mm² current density,” in *IEEE Int. Solid-State Circuits Conf. (ISSCC)* Dig. Tech. Papers, Feb. 2020, pp. 382–384.

C.-H. Huang, X. Sun, Y. Chen, R. Pamula, A. Mandal, and V. Sathe, “A single-inductor 4-output SoC with dynamic droop allocation and adaptive clocking for enhanced performance and energy efficiency in 65 nm CMOS,” in *IEEE Int. Solid-State Circuits Conf. (ISSCC)* Dig. Tech. Papers, Feb. 2021, pp. 416–417.

**MUHAMMAD ABRAR AKRAM** (Member, IEEE) received the B.S. degree in electrical engineering from the University of Punjab, Lahore, Pakistan, in 2013, and the combined M.S. and Ph.D. degrees in electrical and medical convergent engineering from Kangwon National University, Chuncheon, South Korea, in 2019.

He is currently a Postdoctoral Associate with New York University Abu Dhabi, United Arab Emirates. His research interests include analog/mixed-signal CMOS integrated circuits for power management in mobile devices and mm-scaled bio-implants.

Dr. Akram was a recipient of the Korean Semiconductor Industry Association (KSIDA) President Award in 2016, the KSIDA Company Special Award in 2018, and the Best Paper Award at International Conference on Electronics, Information, and Communication (ICEIC) in 2021.

**IN-CHUL HWANG** (Senior Member, IEEE) received the B.S., M.S., and Ph.D. degrees from Korea University, Seoul, South Korea, in 1993, 1995, and 2000, respectively.

He was a Postdoctoral Associate with the Coordinated Science Laboratory, University of Illinois at Urbana–Champaign, Urbana, IL, USA, from 2000 to 2001. In 2001, he joined the Samsung Electronics, System LSI Division, Ilsan-gu, South Korea, where he led a Design Team for developing CMOS Global System for Mobile Communication (GSM)/Enhanced Data for GSM Evolution/Wideband Code Division Multiple Access RF transceiver. Since 2007, he has been with the Department of Electrical and Electronics Engineering, Kangwon National University, Chuncheon, South Korea, where he is currently a Professor. He was a Visiting Scholar with the Georgia Institute of Technology, Atlanta, GA, USA, in 2012. His current research interests include digital phase-locked loops for radio-frequency integrated circuits and power management circuits for dynamic-voltage and frequency scaling applications. He was a recipient of the LG Design Contest Grand Prize in 1999 and the DAC Student Design Contest Award in 2002.

**SOHYOUNG HA** (Senior Member, IEEE) received the B.S. (summa cum laude) and M.S. degrees in electrical engineering from the Korea Advanced Institute of Science and Technology (KAIST), Daejeon, South Korea, in 2004 and 2006, respectively, and the M.S. and Ph.D. degrees in bioengineering from the Department of Bioengineering, University of California, San Diego, La Jolla, CA, USA, in 2015 and 2016, respectively.

From 2006 to 2010, he worked with Samsung Electronics as a Mixed-Signal Circuit Designer of commercial multimedia devices. After this extended career in industry, he returned to academia as a Fulbright Scholar. Since 2016, he has been with New York University Abu Dhabi, United Arab Emirates, and New York University. He is currently an Assistant Professor with New York University Abu Dhabi, and a Global Network Assistant Professor of electrical and computer engineering with New York University, New York, NY, USA.

Dr. Ha received a Fulbright Fellowship in 2010 and the Engelson Best Ph.D. Thesis Award for Biomedical Engineering from the Department of Bioengineering, University of California, San Diego, CA, USA, in 2016. He has served as an Associate Editor for *Smart Health* (Elsevier), IEEE *Transactions on Biomedical Circuits and Systems*, and Frontiers in *Electronics*. He is also a member of the Analog Signal Processing Technical Committee and the Biomedical and Life Science Circuits and Systems Technical Committee of IEEE Circuits and Systems Society.