Four Monolithically Integrated Switched-Capacitor DC–DC Converters With Dynamic Capacitance Sharing in 65-nm CMOS

Ivan Bukreyev⬡, Christopher Torng, Waclaw (Wacek) Godycki, Christopher Batten, *Member, IEEE*, and Alyssa Apsel

Abstract—We present a network of *dynamic capacitance sharing* (DCS) switched-capacitor converters that increase the range of efficient voltage regulation for multiple independent loads while reducing area overhead. Since maximum power dissipation is fixed for a single chip due to thermal constraints, the proposed converters consider the overall power budget of multiple voltage-scalable loads to dynamically share energy storage area, allowing the dynamic allocation of energy storage clusters on-demand. Our DCS converters utilize a feedback control scheme including both capacitance and frequency modulation, which leads to the order of 10–100 ns voltage settling times. A test chip with 16 clusters and four regulator control loops is fabricated in 65-nm bulk CMOS process. For a 2.3 V input, our DCS converters achieve 0.742 V at 38.1 mA to 1.367 V at 298 mA output with peak efficiency of 70.9% at 550-mW/mm^2 power density. Regulator area for the four-load network is reduced by up to 70% when operating under a power constraint compared with the stand-alone per-load regulators capable of supporting an equivalent range of operating voltages.

Index Terms—CMOS switched capacitor voltage regulator, dynamic capacitance sharing, power/thermal constraints.

I. INTRODUCTION

CMOS switched-capacitor (SC) voltage regulators (VRs) with high power densities are a promising means of integrating DC-DC conversion on chip. Such on-chip VRs can support multiple voltage levels with faster transient response times than their off-chip counterparts, while maintaining high efficiency [14], [21].

There are many applications in which VRs are required to support rapidly changing output voltage levels across multiple loads, but where die area and the power budget are constrained. One example is power delivery to the often duty-cycled computation and communication blocks in Internet of Things (IoT) nodes [6]. Another common application is power delivery to

multicore processors that support dynamic voltage and frequency scaling (DVFS) under thermal and power constraints. DVFS is a well-known technique in which clock frequency and supply voltage of digital loads are dynamically adjusted to conserve energy and/or improve performance. Multicore processors can achieve energy-efficiency and performance benefits by quickly transitioning cores between low-voltage and high-voltage operating modes [2], [7], [10], [12], [27]. On-chip VRs in such contexts must be designed with the challenge of (1) supporting a wide range of supply levels with (2) fast transition times, while occupying (3) small area and still providing (4) high conversion efficiency. In this paper, we explore this problem in the context of multiple on-chip DVFS-enabled loads, while noting that our proposed solutions can also be applied to the design of VRs targeting IoT and other similar heterogeneous systems of circuits.

Previous research shows that SCVRs with a large range of efficient voltage regulation and fast sub-microsecond voltage transition times can improve performance and reduce power consumption of multiple DVFS-enabled loads [7], [10], [12]. However, individual per-load VRs require enough energy-storage area to efficiently supply the highest-power operating mode, even if the load typically operates at lower-power modes. This over-design results in large, cost-prohibitive area overheads across VRs and exemplifies a key challenge for integrated voltage regulation. Previous work has also observed that from embedded platforms to servers, power and thermal constraints prevent all loads from using the highest-power operating mode simultaneously [7], [8], [16], [26]. For example, ARM big.LITTLE power allocation schemes restrict the highest-power combinations of voltages and frequencies across cores in order to remain under the thermal design power (TDP) [16], [26]. Even if TDP is not an issue, the average power could be bounded across functional blocks of battery-powered embedded systems or energy-harvesting wake-up radios. This key observation, in conjunction with analysis in the theory section of this paper, suggests that instead of providing enough total energy-storage area for all individual VRs to support the highest-power operating mode at once, a significantly smaller total energy-storage area (enough to run a few loads at the highest-power operating mode) can be shared across VRs and used as needed.

This paper presents a novel circuit architecture and technique called *dynamic capacitance sharing* (DCS) in which

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

2

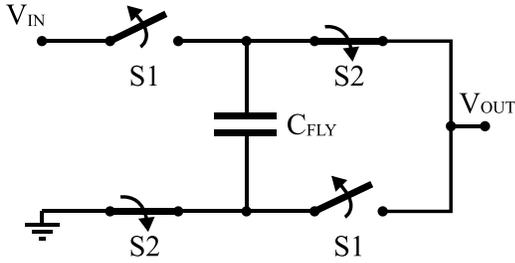IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS–I: REGULAR PAPERS

Fig. 1. A simple two-phase, series-parallel SCVR with a 2:1 conversion mode. The flying capacitor C_{fly} switches from series configuration (S1) to parallel configuration (S2) at the rate determined by the switching frequency.
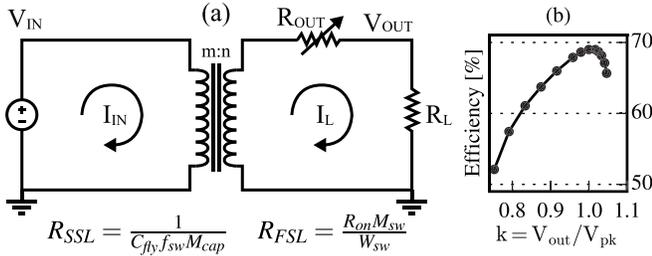

Fig. 2. (a) An SCVR modeled as an ideal transformer with a variable output resistance. (b) Efficiency scaling based on measured data from Fig. 13 for 5 clusters.

a pool of energy-storage *clusters* and independent regulator control loops can be combined to create multiple variable-area VRs on demand. The proposed circuit topology exploits the fact that stand-alone VRs suffer from deficit capacitance at the high-power end and excess capacitance at the low-power end when providing efficient regulation under a fixed area constraint. In addition, we exploit the fact that the cumulative average power draw across all loads is limited by chip-level power-thermal design constraints [7], [8], [16], [26]. Furthermore, by dynamically reallocating capacitance between regulators, the proposed topology utilizes both fine and coarse SCVR control simultaneously, thereby enabling fast transient response to voltage/current steps. By utilizing energy-storage sharing, DCS VRs can reduce the total regulator area by up to 70% for the same output voltage range, while simultaneously improving transient response and maintaining high efficiency.

II. CAPACITANCE SLACK ANALYSIS

There are many important parameters in the design of a single VR that can support a large range of output voltage (V_{out}) and load current (I_L) values, including input voltage (V_{IN}), regulator topology, conversion modes, and control scheme. Fig. 1 shows a simple two-phase, series-parallel SCVR with a 2:1 conversion mode. The basic principle behind SCVR operation is as follows. During the charging phase S1, the input is connected in series to the output through the flying capacitor (C_{fly}). During the discharge phase S2, the input is disconnected and the previously stored charge on C_{fly} is delivered to the output. Such a converter can be modeled as an ideal transformer with a variable output resistance, as shown in Fig. 2a. Adjusting the conversion ratio $m : n$ is

the most straightforward method for increasing V_{out} range. However, to achieve any intermediate output voltage, output resistance (R_{out}) modulation is still required. Regardless of the final SCVR implementation, a larger C_{fly} is required to efficiently support higher V_{out}. This requirement is exacerbated for multiple VRs supplying multiple independent loads, since each individual VR will require a large area that is typically under-utilized due to chip-level power-thermal constraints.

This section provides a theoretical motivation for dynamic capacitance sharing (DCS) by analyzing how varying the flying capacitance C_{fly} impacts SCVR performance across V_{out} for typical DVFS-enabled loads (e.g., IoT nodes, processors). For such loads, the output voltage and clock frequency scale together as pairs (linear to first order) in order to speed up and/or slow down individual loads executing work. Since power consumption is dominated by dynamic power, the output voltage, current, and power are approximately related as follows:

$$f_{clk} \propto V_{out}, \quad I_L \propto V_{out}^2, \quad P_{out} \propto V_{out}^3, \quad R_L = \frac{V_{out}}{I_L} \propto \frac{1}{V_{out}}. \tag{1}$$

In the following first-order analysis, we provide theoretical motivation for DCS by: (1) deriving an equation for C_{fly} as a function of efficiency and load; (2) analyzing required C_{fly} utilization vs. output voltage and efficiency requirements; (3) extending the above relationship of capacitance utilization to multiple DVFS-enabled loads subject to a chip-level power constraint.

A. Flying Capacitance vs. Efficiency and Load

To begin, we express the converter loss P_{loss} in terms of technology and SCVR design parameters as described in [14],

$$P_{loss} = \sqrt{\left(\frac{I_L^2}{C_{fly}f_{sw}M_{cap}}\right)^2 + \left(\frac{I_L^2 R_{on}M_{sw}}{W_{sw}}\right)^2}$$
$$+ C_{fly}f_{sw}V_{out}^2 k_{bott}M_{bott} + W_{sw}f_{sw}V_{sw}^2 C_{gate}, \tag{2}$$

where M_{cap}, M_{bott}, and M_{sw} are conversion topology constants. The technology parameters are R_{on} (switch resistance density in $\Omega \cdot m$), C_{gate} (gate capacitance density in F/m), and k_{bott} (bottom-plate capacitance ratio such that $C_{bott} = k_{bott}C_{fly}$). I_L, V_{out}, and V_{sw} are load current, load voltage, and switch gate voltage, respectively.

We make an important assumption that $W_{sw} = \alpha C_{fly}$, or that the switch width scales linearly with C_{fly}. The implications of this assumption will be discussed in Section IV-E. Although series resistance of the switches can impact charge flow through capacitors, the two loss terms can be approximated as a sum, which enables a simple and meaningful expression without significantly altering the results as explained in [14]. Equation (2) then becomes:

$$P_{loss} = \frac{I_L^2}{C_{fly}}\left(\frac{1}{f_{sw}M_{cap}} + R_{on}M_{sw}/\alpha\right)$$
$$+ C_{fly}f_{sw}\left(V_{out}^2 k_{bott}M_{bott} + V_{sw}^2 \alpha C_{gate}\right). \tag{3}$$

Next, we assume switches are driven by the output voltage such that $V_{sw} = V_{out}$, which is reasonable for a

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

BUKREYEV *et al.*: FOUR MONOLITHICALLY INTEGRATED SC DC–DC CONVERTERS

3

step-down converter. To find P_{loss} at optimum switching frequency, we remove the independent variable f_{sw} from (3) by finding the non-negative critical point of the partial derivative with respect to f_{sw}:

$$P_{loss} = 2I_L V_{out} \sqrt{(k_{bott}M_{bott} + \alpha C_{gate})/M_{cap}} + \frac{I_L^2 R_{on} M_{sw}}{\alpha C_{fly}}. \tag{4}$$

By relating P_{loss} to the output power, (4) can be solved for C_{fly} as a function of efficiency (η), load, and technology parameters:

$$C_{fly} = \frac{I_L R_{on} M_{sw}}{V_{out}\left(\frac{1}{\eta} - \beta\right)\alpha}, \tag{5}$$

$$\beta := 1 + 2\sqrt{\frac{1}{M_{cap}}(k_{bott}M_{bott} + \alpha C_{gate})}. \tag{6}$$

Since V_{out} and I_L for DVFS-enabled loads are related by (1), (6) relates C_{fly} to efficiency and load, and in turn sets peak flying capacitance (C_{pk}) for desired peak efficiency (η_{pk}) at the peak power (P_{pk}). Note that the constant α is picked at this time for optimum efficiency at the highest power mode.

B. Flying Capacitance Utilization

A classical SCVR will use C_{pk} for all output voltages. However, for $V_{out} < V_{pk}$ it is possible to use a strictly smaller C_{fly} to satisfy the load demand, albeit at the potential cost of lower efficiency. To capture load variability, we introduce a voltage scaling factor (k) such that

$$V_{out} = kV_{pk} \quad \text{for } k \leq 1. \tag{7}$$

When delivering power across a wide voltage range, a single-topology fixed-capacitance SCVR will achieve its peak efficiency only for one voltage value ($k = 1$). For any other V_{out}, the efficiency will decrease due to the required increase in R_{out} in order to maintain regulation, as illustrated in Fig. 2b.

To find the required C_{fly} for a wide range of output voltages at a fixed V_{IN}, it is important to estimate how the efficiency scales with k. We consider two boundary cases, while noting that efficiency scaling in a practical implementation will likely lie in between the two. First, if the topology conversion ratio can vary continuously[1] with k, the efficiency can remain equal to η_{pk} for the entire range of V_{out}. For such a flexible converter the required capacitance would be C_{req1}. Second, if the conversion ratio is kept fixed, the SCVR efficiency will roughly decrease linearly with k, similar to what is shown in Fig. 2b. For such a converter, the required capacitance would be C_{req2}. We now use (6) in conjunction with the scaling described by (1) to find C_{req1} and C_{req2}. The results are:

$$C_{req1} = kC_{pk}, \tag{8}$$

$$C_{req2} = \frac{k^2 R_{on} M_{sw}/\alpha}{R_{pk}\beta(1-k) + R_{on} M_{sw}/\alpha C_{pk}}. \tag{9}$$

Fig. 3. (a) Normalized C_{fly} utilization for continuous (C_{req1}) and fixed (C_{req2}) conversion topology. As k decreases, less C_{fly} is required. $k = 1$ corresponds to 1.27 V@309 mW from Fig. 13. (b) Corresponding slack vs. normalized maximum power. For lower $P_{tot,max}$, more C_{fly} is under-utilized.

Equations (8) and (9) are plotted in Fig. 3a for our process. Note that efficiencies are η_{pk} and $k\eta_{pk}$ for C_{req1} and C_{req2}, respectively. As k decreases, the required C_{fly} is significantly reduced due to the cubic power scaling of DVFS-enabled loads. For example, if efficiency is allowed to drop 10% with reduced voltage (i.e., $k = 0.9$), C_{fly} requirement may be reduced by up to 80%.

C. Extension to Multiple Loads Under a Power Budget

Based on the findings in other works [1], [7], [16], [18], [26], it is reasonable to have a chip-level maximum output power constraint $P_{tot,max}$[2] for multiple independent DVFS-enabled loads on the same die. Specifically, simultaneous operation at P_{pk} is never allowed for all N SCVR-load pairs such that $P_{tot,max} < NP_{pk}$. Then, the sum of N independent powers is bounded by:

$$\sum_{i=1}^{N} P_{out,i} \leq P_{tot,max}. \tag{10}$$

Given scaling from (1), we approximate the output power of each load as $P_{out,i} \approx k_i^3 P_{pk}$. The power constraint in (10) is then:

$$\sum_{i=1}^{N} k_i^3 P_{pk} \leq P_{tot,max}. \tag{11}$$

Based on the insights from Section II-B, a system under a power constraint could use strictly less total flying capacitance than NC_{pk}. If C_{fly} could be shared continuously among the N loads, we can solve for the maximum or worst-case total capacitance C_{opt} that could be used under this constraint using the method of Lagrange multipliers:

$$\begin{cases} C_{opt} = max\left(\{\sum_{i=1}^{N} C_{req}(k_i) : k_i \leq 1\}\right) \\ \text{constrained by } \sum_{i=1}^{N} k_i^3 = \dfrac{P_{tot,max}}{P_{pk}}. \end{cases} \tag{12}$$

---

[1]While topology constants in (6) also vary continuously with k, this effect is ignored in the analysis. Fixing constants in this derivation allows us to generate an upper bound on theoretical results.
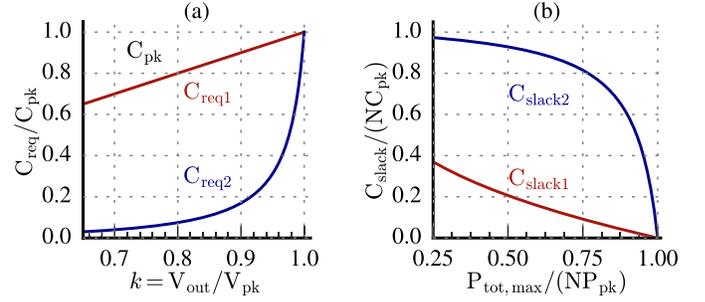
[2]This analysis does not include loss due to the finite efficiency of the converters under the power budget. $P_{tot,max}$ available to the loads is overestimated for non-ideal converters. More complete analysis is left to future work.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

4

IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS–I: REGULAR PAPERS

The general solution to the system is given by:

$$k_i = \sqrt[3]{\frac{P_{tot,max}}{NP_{pk}}}, \tag{13}$$

where all k_i are equal. We define capacitance slack as the difference between the capacitance that would be used by N VRs with fixed (C_{pk}) vs. N VRs with k-dependent capacitance (C_{req1}, C_{req2}) derived in Section II-B. The result is:

$$C_{slack1} = NC_{pk}\left(1 - \sqrt[3]{\frac{P_{tot,max}}{NP_{pk}}}\right), \tag{14}$$

$$C_{slack2} = NC_{pk} - N\frac{\sqrt[3]{\frac{P_{tot,max}}{NP_{pk}}}^2 R_{on}M_{sw}/\alpha}{R_{pk}\beta\left(1 - \sqrt[3]{\frac{P_{tot,max}}{NP_{pk}}}\right) + R_{on}M_{sw}/\alpha C_{pk}}. \tag{15}$$

Fig. 3b shows that capacitance slack increases as the normalized total power decreases, which presents the opportunity to save capacitance by reallocating it dynamically between converters, thus saving on-chip area.

III. DYNAMIC CAPACITANCE SHARING

The analysis presented in Section II suggests that lower-than-peak C_{fly} can be used to meet the load demand for $V_{out} < V_{pk}$. The analysis also suggests that for a given power constraint, under-utilized (slack) capacitance can be re-purposed to other converters in order to save chip area. To leverage this slack, we propose dynamic capacitance sharing (DCS) VRs that dynamically reconfigure C_{fly} via energy-storage *clusters*. DCS VRs are created by combining clusters with several fine-grain control loops and a global DVFS controller, are suitable for multiple independent SCVRs on the same die, and can enable more efficient use of chip area and expand the efficient range of voltage regulation. The remainder of this section discusses how coarse-grain DVFS control and fine-grain control loops form a dual feedback control scheme and analyzes their combined impact on the transient response.

A. Energy Storage Sharing and DVFS Controller

To control cluster allocation, we propose leveraging past work on lightweight DVFS controllers based on activity hints in low-power embedded processors [7], [15]. In this prior work, a multi-threaded application is instrumented with hint instructions that toggle activity bits indicating the status of the core executing the hint. The activity bits are read by a global DVFS controller, which then decides how to set the supply voltages and allocate available C_{fly}. The DVFS controller uses a simple lookup table designed offline that maps activity information into the appropriate voltage levels and C_{fly} to supply each load at high efficiency. At runtime, the controller dynamically reallocates units of capacitance across active loads and sets the appropriate V_{REF} levels. Although prior work used a simple lookup table, more sophisticated online adaptive DVFS controllers are also possible. To narrow the scope of this work, we assume control cluster allocation is performed via programmable registers that are
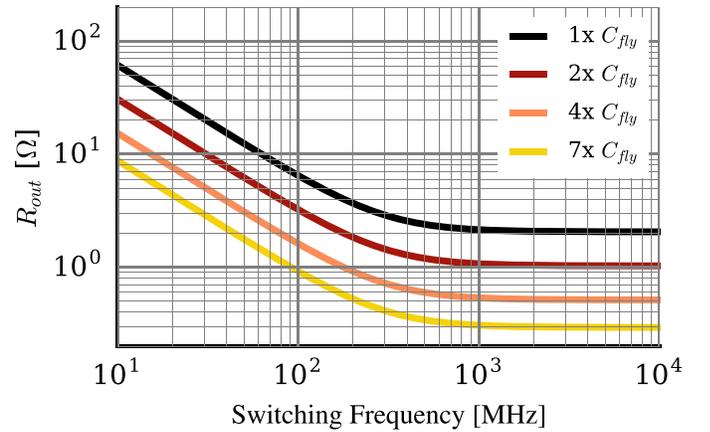

Fig. 4. Output resistance at SSL and FSL limits for 1–$7 \times C_{fly}$, computed based on the model in [20]. Note that $W_{sw} = \alpha C_{fly}$ where α is constant.

configured off-chip. This reduces design complexity while still allowing transient response measurements.

Fig. 4 shows how varying C_{fly} at fixed α (see Section II-A) affects an SCVR's output resistance at both the slow switching limit (SSL) and the fast switching limit (FSL). At the SSL, all cases can achieve a desired output resistance, however, lower C_{fly} is preferred since it permits higher f_{sw}, which in turn reduces feedback controller complexity. At the FSL, larger C_{fly} is required for achieving low output resistance that is otherwise not possible. By dynamically sharing C_{fly} across N loads, DCS VRs expand the efficient range of voltage regulation. Fine-grain control loops with low-power loads can give away C_{fly} to enable lower V_{out} for a given f_{sw}, while control loops with high-power loads can acquire additional C_{fly} to support higher V_{pk}.

B. Fine-Grain Feedback Control

A fast transient response is one of the most compelling reasons to integrate VRs on chip. SCVRs with fast transient response can quickly adapt to load-power requirements and save more energy during transition overhead. There are several popular control techniques for SCVRs: f_{sw} modulation via voltage-controlled oscillator (VCO) tuning [5], [13], [23], bound hysteretic control [3], [4], [9], [27], capacitance modulation and capacitance dithering [2], [17], and others [25]. This work utilizes a VCO-based frequency modulation feedback control scheme (Fig. 8b) due to the ease of implementation and theoretical analysis. A VCO-based approach allows for subharmonic oscillation-free operation for our distributed-capacitance architecture. A bound hysteretic or any other frequency controller could be used instead, but each has its own set of design trade-offs (e.g, the comparator clock frequency range and delay criterion).

C. Transient Response Analysis

By combining capacitance and frequency modulation into a dual control scheme, our DCS converters achieve improved transient response time. To analyze the benefits of this dual

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

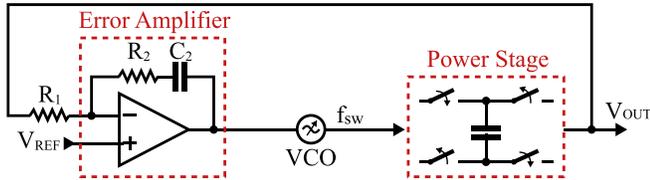BUKREYEV *et al.*: FOUR MONOLITHICALLY INTEGRATED SC DC–DC CONVERTERS

5

Fig. 5. An SCVR under VCO-based frequency modulation control.

control scheme, consider an SCVR capable of 2:1 and 3:2 conversion modes with a total flying capacitance of C_{pk}. Fig. 6a shows the simulated f_{sw} requirement for a fixed-capacitance SCVR under cubic power scaling. For a 2:1 conversion mode, such VR quickly hits SSL at the low-power end. At the same time, this VR requires a high relative f_{sw} adjustment range for the medium-power modes. Fig. 6b shows f_{sw} requirement for a SCVR that has C_{pk} divided into seven clusters that can be added or removed from the VR's active energy storage on the fly. Here f_{sw} variation is significantly reduced compared to a fixed-area scheme due to the coarse-grain capacitance modulation. As V_{out} increases, a more optimal C_{fly} is used before the VR is allowed to approach either the slow or fast switching limits.

Next we analyze how varying both C_{fly} and f_{sw} affects small-signal behavior of the closed-loop feedback controller (Fig. 5). Using state-space analysis, the steady-state switching frequency to output transfer function of the power stage derived in [22] is:

$$P(s) = \frac{G_o}{1 + \tau_0 s}, \quad (16)$$

where G_o is the power stage static gain of the form $\Delta V_{out}/\Delta f_{sw}$ and τ_0 is the power-stage time constant that is equal to $R_{out}C_{load}$ [23]. To compensate for the single pole of the power stage, a proportional-integral error amplifier is used with frequency response of the form:

$$C(s) = \frac{R_2}{R_1} \cdot \frac{R_2 C_2 s + 1}{R_2 C_2 s} = K_p \cdot \frac{T_i s + 1}{T_i s}. \quad (17)$$

The closed-loop system is formed by the power stage, error amplifier, and a linear VCO with a voltage-to-frequency gain of K_{VCO}. By inspection of the block diagram in Fig. 5, the open-loop transfer function is given by:

$$G(s) = P(s) \cdot C(s) \cdot K_{VCO} = \frac{G_o K_p K_{VCO}}{1 + \tau_0 s} \cdot \frac{1 + T_i s}{T_i s} \quad (18)$$

and the closed-loop transfer function is:

$$T(s) = \frac{G(s)}{1 + G(s)} = \frac{1 + T_i s}{1 + \frac{2\zeta}{\omega_0} s + \frac{s^2}{\omega_0^2}}, \quad (19)$$

where $\frac{2\zeta}{\omega_0} = T_i(1 + \frac{1}{K})$, $\omega_0^2 = \frac{K}{\tau_0 T_i}$, and $K = G_0 K_p K_{VCO}$. Given a power stage with τ_0 and G_0 and some steady-state operating point, the feedback parameters K_p, K_{VCO}, and T_i are set by following the methodology described in [22] to stabilize the system. We fix all feedback parameters and study how small-signal behavior of the system changes when varying C_{fly}.

For an underdamped or critically-damped system with the form of (19), a step response settling time to within 2% of the steady state is given by:

$$t_s = \frac{3.9}{\zeta \omega_0} = \frac{7.8\tau_0}{1 + K}. \quad (20)$$

For a fixed K_p and K_{VCO}, the best t_s can be achieved for a large power stage gain G_0 and small time constant τ_0. Increasing C_{fly} reduces τ_0 due to the reduced R_{out}, but simultaneously increases the gain G_0 of the power stage. Fig. 7 shows simulated V_{out} vs. f_{sw} for cubic power scaling within and near the linear SSL region governed by R_{SSL} (Fig. 2). For fast transient performance, an infinite G_0 for the entire range of output voltage is desirable. Achieving such an ideal response with a single C_{fly} over the entire voltage range is not possible as the regulator would need to adapt f_{sw} instantaneously across a very wide range of frequencies to accommodate both the SSL and FSL of the regulator. However, it is possible to achieve a high *average* G_0 (see dotted line in Fig. 7) by applying coarse-grain capacitance modulation. Toggling the capacitance at discrete intervals enables jumping across curves (see arrows in Fig. 7), resulting in a more desirable average response with a large effective G_0 throughout the operating range. Note that in a pure capacitance modulation scheme, f_{sw} would remain fixed while capacitance is dithered at a fine level. Such a technique can also achieve very fast transition times [2] but makes capacitance sharing challenging.

D. Summary

DCS converters have two significant advantages over a group of stand-alone SCVRs that provide the same efficient range of V_{out}. First, the overall regulator area is reduced because C_{fly} is shared. While doing so does not allow all loads to simultaneously operate at the highest-power mode, recall that chip-level power and/or thermal constraints preclude this scenario in practice [7], [8], [16], [26]. Second, transient response of each control loop is improved because the f_{sw} adjustment range is reduced when the optimized amount of C_{fly} is used to deliver power to the load. To our knowledge, the work presented in this paper describes the first regulator network that incorporates capacitance sharing with the goal of supporting DVFS-enabled loads with reduced area at the same range of V_{out} across multiple loads. Prior work has focused on capacitance modulation and/or dithering for voltage regulation [2], [3], [17]; however, excess capacitance was either converted to a different topology, dithered, or disconnected completely. Salem et al. [19] switch load instead of C_{fly} to achieve high-efficiency regulation, while removing the need for C_{fly} altogether. Jiang et al. [11] present dual output SCVR with unified and partitioned C_{fly} where the boundary of the power cells is dynamically shifted between two converters. Their work and our previous findings in [7] take advantage of capacitance over-provisioning in networks of SCVRs. In contrast, this work further extends energy-storage sharing to a network of four DVFS-enabled load-converter pairs and develops first-order theoretical models for capacitance slack. For DCS converters shared flying capacitance can

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

6

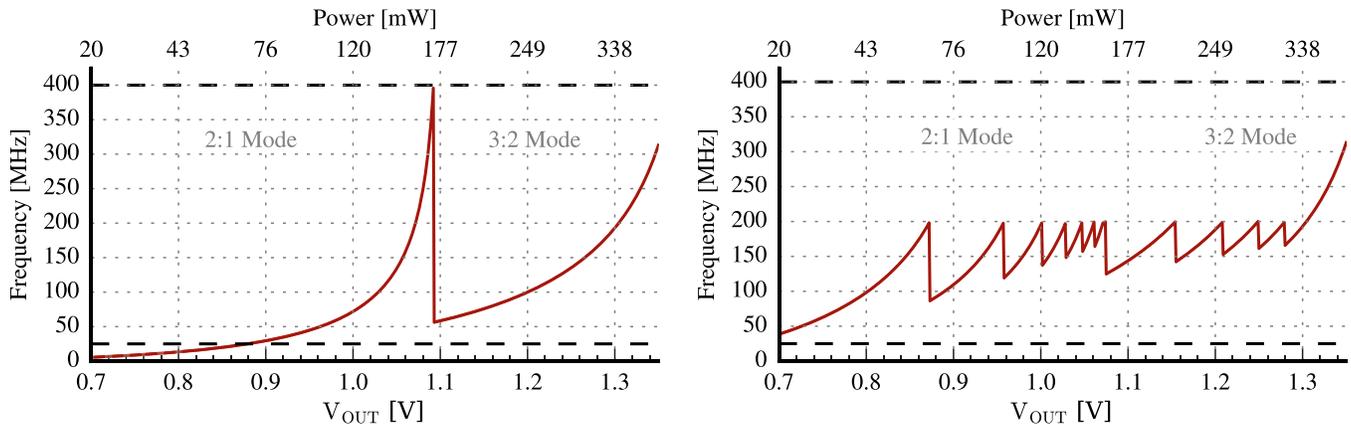IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS–I: REGULAR PAPERS

Fig. 6. Frequency vs. DVFS voltage scaling for fixed vs. variable-area SCVR capable of 2:1 and 3:2 conversion mode. Black lines represent switching frequency limits for our test chip. (a) Fixed-area SCVR with energy-storage equivalent of 7 clusters. (b) Variable-area SCVR that can select between 1 to 7 clusters. For both cases, maximum efficiency point is selected based on available conversion modes and clusters.
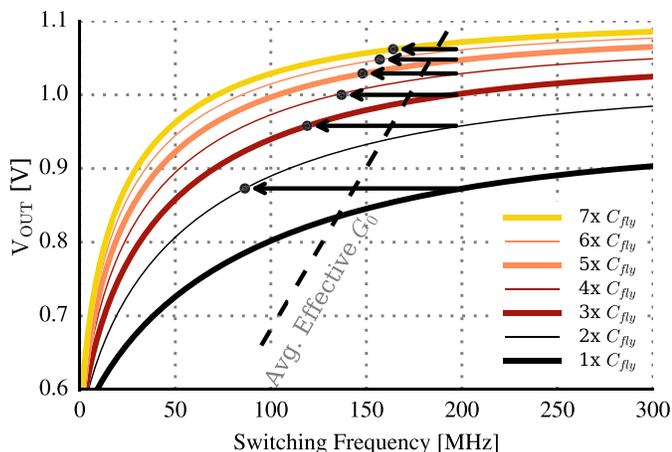

Fig. 7. Simulated V_{out} vs. switching frequency at cubic power scaling for varying C_{fly} values. Arrows indicate transitions to a different cluster count. Dotted line represents an average effective G_0 achieved with DCS.

be (1) arbitrarily assigned between all connected loads, (2) is used to enhance performance of both high-power and low-power independent loads, and (3) improves transient response times.

IV. CIRCUITS FOR DYNAMIC CAPACITANCE SHARING

There are three key design challenges relating to DCS VRs for multiple DVFS-enabled loads. First, each load has separate voltage and current requirements for optimal operation. Second, the highest power mode (V_{pk}) of each load sets the energy-storage requirement for the corresponding VR, constraining area. Third, a reconfigurable switch fabric that connects VRs to loads must be robust and have low resistance to enable fast, efficient, and reliable operation. To tackle these challenges and achieve a large range of V_{out} for multiple DVFS loads with a given fixed area, we have designed four DCS converters with 16 energy-storage clusters that can be dynamically reconfigured across the four loads. Fig. 8a shows the top-level block diagram with the primary circuit blocks used to construct DCS converters. Per-load VRs are formed

by connecting at least one cluster to a load with an associated control loop. The remainder of this section discusses the circuit design aspects of each block.

A. Fine-Grain Control Loop

Fig. 8b shows how each fine-grain per-load control loop consists of a dynamic latched comparator, a charge pump, and a VCO. The comparator compares V_{out} to V_{REF} and generates a digital-pulse waveform that is integrated by the charge pump. The resulting voltage output sets the VCO frequency in the tuning range of 100–1600 MHz. The undivided VCO output clocks the comparator. With the VCO gain and filter capacitance fixed, the overall control loop gain is set by adjusting the charge pump current via a programmable bias. For our DCS converters, switching frequency modulation performs only fine-grain load regulation. As discussed in Section III-A, cluster allocation performs coarse-grain regulation.

B. Cluster: Clock Generation Unit

The output clocks of all four control loops are routed to every cluster. Fig. 8c shows how each cluster then divides all VCO clocks by four to generate four sets of eight clock phases to drive the C_{FLY} cells. The time-domain cartoon in Fig. 8d shows example signal flow from a control loop to a C_{FLY} cell. Clock generation is done at the cluster level to allow rapid reconfiguration between the loads. Since clusters are reconfigured dynamically, the clocks regulating a particular cluster may shift from one control loop to another during runtime. To avoid metastable switching states during a reconfiguration event, each cluster continuously generates four sets of clock phases corresponding to each control loop and simply selects which ones are outputted (Fig. 8c, left). While incorporating more clock phases could improve converter efficiency and reduce output ripple [14], our design uses eight phases in order to minimize complexity and area overhead of the clock generation circuitry.

C. Cluster: Interleaved Flying Capacitance

Each cluster contains eight cells of switched flying capacitance ($\approx 384\,\text{pF}$) in series-parallel configuration, based on

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

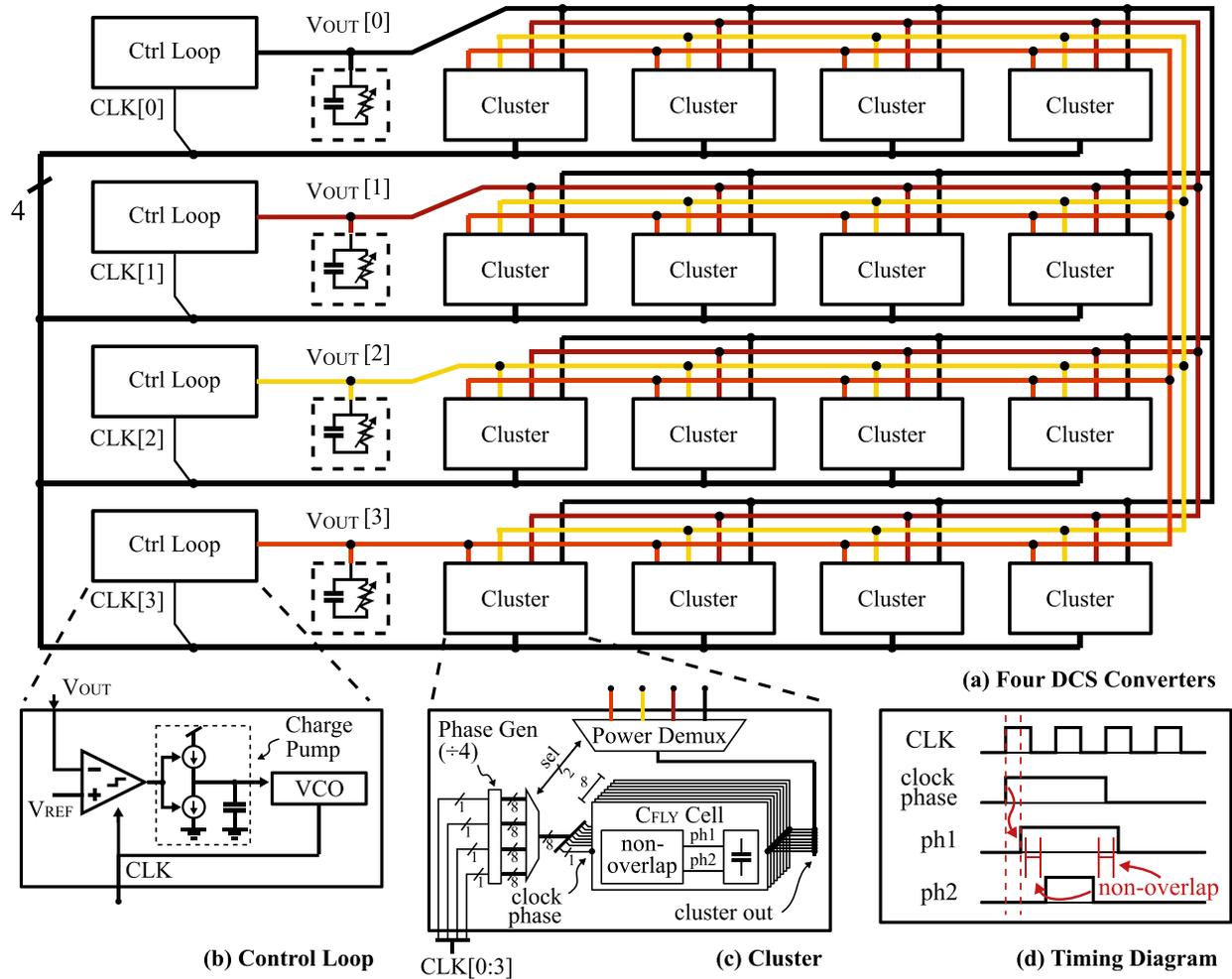BUKREYEV *et al.*: FOUR MONOLITHICALLY INTEGRATED SC DC–DC CONVERTERS

7

Fig. 8. Top-level structure of four DCS converters supporting four independent loads. 16 energy-storage clusters are dynamically connected via power demuxes to one of the loads with an associated control loop.
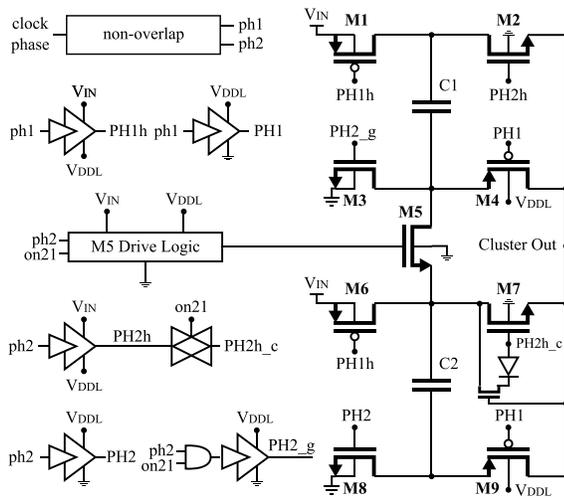

Fig. 9. Schematic diagram of a single flying capacitance cell capable of 3:2 and 2:1 conversion, selectable by *on*21. Incoming clock is converted into two non-overlapping phases *ph*1 and *ph*2 that, in combination with buffers and level-shifters, are used to drive switches M1–M9.

the prior work in [14]. Fig. 9 shows a single phase with corresponding switch drivers capable of supporting 2:1 and 3:2 conversion topologies, selectable by the *on21*

control signal. Clusters use thin-oxide MOS capacitors due to their high capacitance density, which also allows top metals to be used for low-resistance power routing. We use low-threshold thin-oxide devices to minimize converter loss. In our design, $V_{IN} = 2.3$ V and exceeds the breakdown voltage of the thin-oxide devices. To guarantee safe operation across the entire output voltage range, a supplementary off-chip voltage supply $V_{DDL} = 1.15$ V, is used to implement the switch drivers. The value of V_{DDL} was chosen to optimize converter efficiency across the full range of V_{out} while making sure that any unwanted conduction due to high V_{pk} stays low. All drive signals remain below 1.2 V and operate between the two fixed voltage supplies, which allows for fast edge transitions and deterministic operation.

Interestingly, the inclusion of V_{DDL} does not complicate power delivery to the chip. By symmetry, switches M3, M4, M8, and M9 operate between V_{DDL} and GND while M1, M2, M6, and M7 operate between V_{IN} and V_{DDL}. The current draw from V_{DDL} is roughly equal to the return current from V_{IN} to V_{DDL}, making V_{DDL} current consumption negligible. Switch M5 is driven by a reconfigurable inverter that operates between either V_{DDL} and ground in 2:1 mode or between V_{IN} and V_{DDL} in 3:2 mode. This drive asymmetry does cause

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

8   IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS–I: REGULAR PAPERS
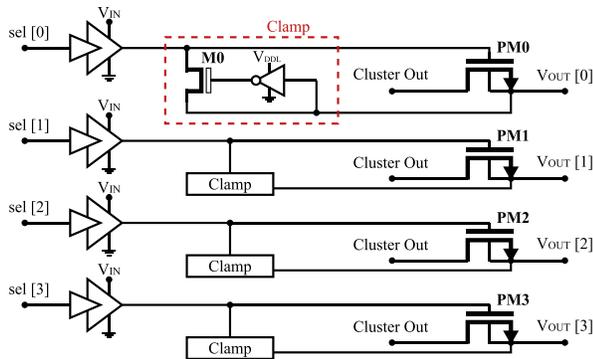

Fig. 10. Schematic diagram of a power demux that allows an associated cluster to be connected to any of the four loads. Voltage clamps prevent breakdown of the thin-oxide power switches during start-up.
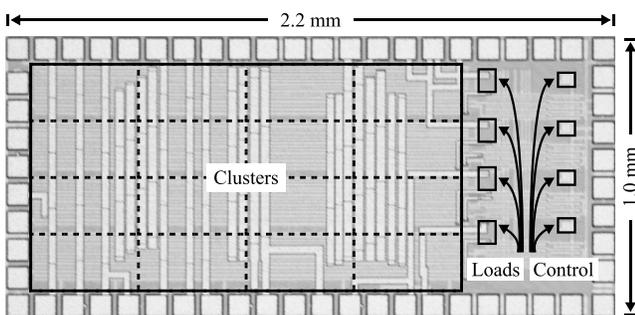

Fig. 11. Fabricated test chip with 16 clusters, 4 loads, and 4 control loops outlined. The remainder of the chip area is occupied by decoupling and testing circuits.

Fig. 12. Efficiency vs. I_L at 2.3 V input normalized to area for 3:2 and 2:1 converter topologies at 200 MHz switching frequency for 1 cluster.

use MIM capacitors in order to reduce routing complexity because they occupy the two thickest metal layers. Our chip serves as an evaluation vehicle for the DCS technique, and future designs could incorporate more sophisticated capacitor technologies.

In principle, DCS VRs can scale to an arbitrary number of loads. The work in [7] uses a system-level controller to manage 32 DCS clusters for an 8-core system, partitioned into two independent quads. To keep the overhead of power/clock routing low, DCS clusters and control loops could be partitioned into isolated domains when more than four loads are used. To this end, fixed α (switch width to C_{fly} ratio) greatly simplifies cluster design as any large converter can be evenly divided into an arbitrary number of clusters. A potential drawback of fixed α is reduced efficiency for $V_{out} < V_{pk}$ since it is optimized for V_{pk}.

F. Overhead Analysis

Each cluster has an area of 0.0766 mm^2. Switches, drivers, and C_{fly} occupy 90% of the cluster area, while DCS-specific circuitry occupies 9.4%, with the power demux (5.9%) and clock generation unit (2.6%) as the major components. Efficiency loss due to power demuxing is 0–1% corresponding to the increasing load current in Fig. 12. For testing, up to four cluster allocation configurations are programmed off-chip and stored in registers in each cluster, allowing configuration change on the fly for transient measurements. The programmability incurs 0.7% area overhead, but practical system reconfiguration could be handled by a system-level controller similar to that proposed in [7] and as described in Section III-A at no penalty to the cluster.

a net current draw from V_{DDL}, but it is between 0–1% of the total input power.

D. Cluster: Power Demux

A unique aspect of every cluster is the power demux shown in Fig. 10, which is used to direct the regulated output of a cluster to any of the four loads. Based on the load assignment, one of the power transistors (PM0-PM3) turns on to connect a cluster's C_{FLY} cells to the selected load. To minimize R_{on}/mm^2, each power transistor is a thin-oxide n-channel device. When enabled, a buffer drives the gate of PM0 to V_{IN}. This drive strategy presents an over-voltage hazard during start-up when load voltage is not yet sufficiently built up. We add a voltage clamp to the gate of PM0. If V_{out} is too low, an inverter turns on M0, which in turn prevents the gate voltage of PM0 from exceeding breakdown. For some combinations of V_{out}, a body diode conduction could occur in a power transistor, however, it is three orders of magnitude less than the current draw of the load.

E. Layout Considerations and Scaling

The majority of the test chip area (Fig. 11) is occupied by the MOS capacitors. Top metal layers are used for low-resistance power routing of the four loads to the clusters, as well as clock routing. To avoid inductive coupling, all signals are interleaved with power or ground. We did not
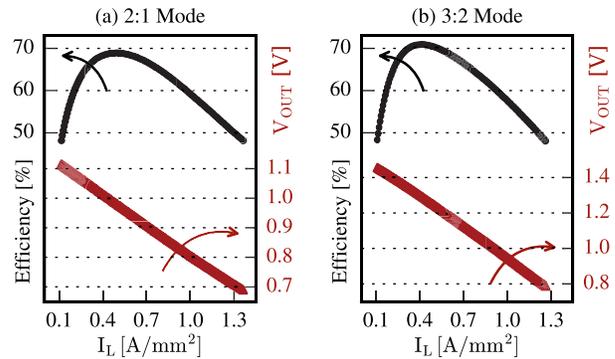
V. MEASURED RESULTS

We have fabricated four DCS converters with 16 clusters that support four independent loads in 65 nm CMOS. Fig. 11 shows the fabricated test chip with the active area of key blocks outlined. The remaining non-active area is used for debugging circuits and decoupling capacitance.

A. On-Chip Load and Testing Circuitry

Each on-chip load is comprised of a 3-bit thermometer-encoded NMOS transistor bank and a parallel MOS capacitor

TABLE I
V_{span} VS. NUMBER OF ACTIVE LOADS FOR FIXED-AREA AND DCS VRS

Regulator Sys.	$N_{clusters}$	V_{span}	V_{nspan}
Fixed, $N_A = 1$–4	4	0.864–1.219 V	0.291
DCS, $N_A = 1$	1–13	0.742–1.367 V	0.457
DCS, $N_A = 2$	1–7	0.742–1.303 V	0.431
DCS, $N_A = 3$	1–5	0.742–1.256 V	0.441
DCS, $N_A = 4$	1–4	0.742–1.219 V	0.392

N_A = number of active loads.

$N_{clusters}$ = number of clusters available to each load.

V_{nspan} facilitates comparison with future works on DVFS-enabled loads.

of $\approx$377 pF. Each transistor bank has an isolated ground routed to a pad for detailed off-chip DC characterizations. Varying the number of connected parallel transistors from zero to seven allows rapid load current steps.

Each control loop selects one of the four off-chip V_{REF} biases via an analog mux. At runtime, any control loop can rapidly switch between any of the available V_{REF}. Cluster allocation, load configuration, and V_{REF} selection are all controlled by an on-chip serial peripheral interface and can be reconfigured at runtime, emulating a global DVFS controller.

B. Conversion Efficiency

Fig. 12 shows that our DCS converters achieve peak efficiencies of 68.9% (505 mW/mm²) for 2:1 mode and 70.9% (550 mW/mm²) for 3:2 mode including the power demux loss. Our SPICE-level simulations indicate that the dominant source of loss is the bottom-plate capacitance ($\sim$ 15%) followed by the switch on-resistance ($\sim$ 5%). DCS converters are not significantly affected by charge-share losses. Worst-case transition based on reported values in Table I would correspond to 1.376 V and 0.742 V clusters connected together. Given cluster capacitance of $\approx$384 pF, the energy loss is 38.6 pJ for the connected clusters. The reported 2:1 mode peak efficiency would be modified per charge-share losses to 70.67% over 100 ns window and to 70.85% over 500 ns window. Based on a survey of state-of-the-art SC regulators, our results are in-line with the efficiency vs. power density trend for our process [24], [25], while enabling additional functionality. Our results indicate that adding the capacitance sharing capability extends the efficient range of load regulation and reduces transient response time at little-to-no cost to converter efficiency.

C. V_{Span} Measurements and Area Savings

A large range of efficient operation is critical for DVFS applications. DVFS-enabled loads have predictable power patterns, unlike unpredictable loads commonly assumed by other works [2]–[5], [17]. Since each DVFS mode is characterized by different voltage and current requirements, reporting load regulation alone does not adequately gauge the ability of a VR to deliver power to a DVFS-enabled load. To better characterize VRs for DVFS applications, we define V_{span} of a VR as:

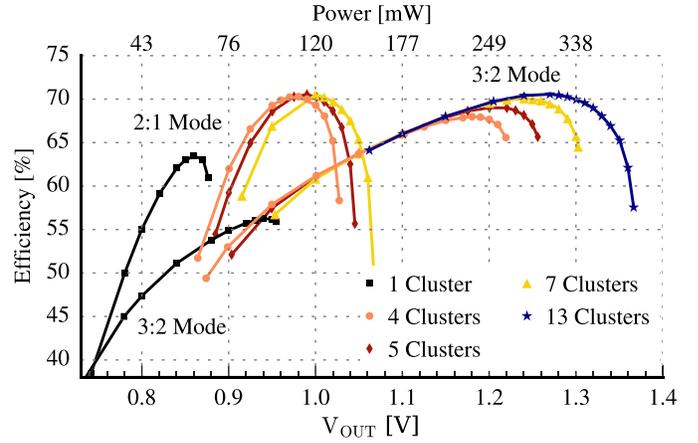$$V_{span} = V_{dvfs,max} - V_{dvfs,min}, \qquad (21)$$

Fig. 13. Efficiency vs. Voltage and Power – Different combinations of clusters and conversion topologies are capable of achieving different V_{span} boundaries. Measured points correspond a DVFS load with cubic power scaling. Data for 2:1 and 3:2 conversion modes is shown.
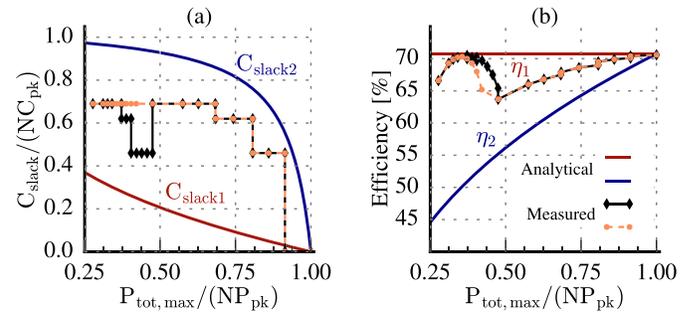

Fig. 14. Comparison of capacitance slack analysis (Fig. 3b) to measured data in Fig. 13 (orange circle and black diamond markers). (a) Capacitance slack has a stair-step pattern due to the discrete cluster count. (b) Corresponding measured vs. theoretical efficiency. It is possible to trade off efficiency for higher C_{slack}: if efficiency is prioritized (black diamonds), slack has a dip due to conversion topology switch; if slack is prioritized (orange circles), efficiency is lower.

where $V_{dvfs,max}$ and $V_{dvfs,min}$ represent the maximum and minimum achievable output voltages at high efficiency (better than ideal linear regulator) for a DVFS-enabled load. A VR with a large V_{span} can efficiently support both the high-power and low-power operating modes of a DVFS-enabled load. The key difference between V_{span} and traditional output voltage range is that the former has cubic power scaling factored in. To highlight how effectively a VR utilizes available input voltage, we define V_{nspan} as V_{span} normalized to the highest output voltage:

$$V_{nspan} = \frac{V_{span}}{V_{dvfs,max}}. \qquad (22)$$

This metric facilitates VR comparison across different technology nodes with different input voltages.

To evaluate V_{span} improvements of DCS converters, we measure four DVFS loads, each with cubic power scaling corresponding to a single core of a low-power embedded processor in 65 nm CMOS. Our on-chip loads serve as a good approximation for digital loads since the converter's switching period is longer than the period of a typical digital clock.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

10   IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS–I: REGULAR PAPERS
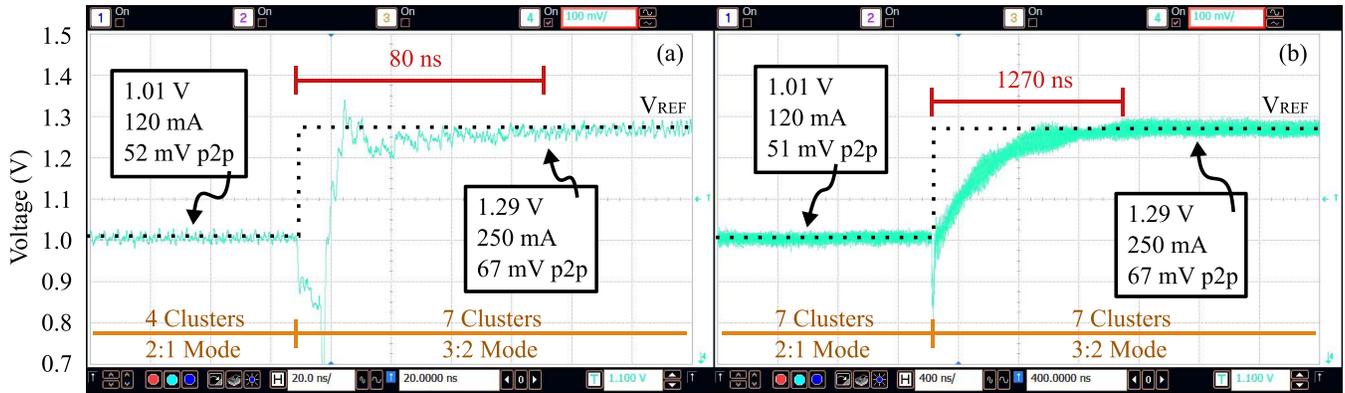

Fig. 15. DCS VR vs. fixed-area VR closed-loop step-up transition. (a) DCS VR transition from low voltage/current to high voltage/current with cluster change from 4 clusters at 2:1 mode to 7 clusters at 3:2 mode. (b) Equivalent transition for a VR with fixed-area of 7 clusters. Dotted line represents V_{REF}.

Fig. 13 shows how DCS converters can dynamically extend a load's voltage range at both the low end, by allocating a single cluster, and at the high end, by allocating up to 13 clusters (assuming each load must have at least one cluster). The range of voltage regulation achievable without DCS is shown in Fig. 13 under *4 Clusters* where each VR has fixed 4-cluster energy-storage area and, consequently, a lower V_{span}.

Given fixed energy-storage area, Table I shows how the DCS technique impacts the range of efficient voltage regulation for a varying number of active loads. In every scenario, DCS converters can select the optimum amount of C_{FLY} and achieve operation modes not possible for the non-DCS equivalent. DCS enables $1.70\times$ higher voltage regulation range by reducing C_{FLY} at the low-power end (1–3 clusters) as well as borrowing capacitance at the high-power end (5–13 clusters). In order for fixed SCVRs to support the full voltage and power range possible with DCS, each per-load VR must have an equivalent energy-storage area of 13 clusters [7].

Fig. 14a shows how measured results in Fig. 13 correspond to the theory from Section II. For a given V_{pk}, DCS converters reduce C_{fly} overhead by up to 70% depending on the power constraint compared to per-core fixed-capacitance VRs designed for equivalent DVFS-enabled loads. The stair-step nature of the plot is due to the discrete cluster count. Fig. 14b shows how measured efficiency lies between the two analytical boundaries. It is possible to trade off DCS VR efficiency for higher capacitance slack. Results shown in black diamonds in Fig. 14 aim at maintaining highest efficiency, or following the peak efficiency curves of Fig. 13. For that reason, there is a decrease in capacitance slack that corresponds to ≈ 1.05 V since 2:1 mode is more efficient at seven clusters than 3:2 mode at four. On the other hand, if aggressive area reduction is desired, results in orange circles show that efficiency could be traded for higher C_{slack}.

D. Transient Measurements

We measured transient performance of our DCS converters using internal loads and a 2 GHz active probe. To evaluate the benefits of DCS over fixed-area VRs, we performed select closed-loop transients corresponding to transitions along the efficiency curves shown in Fig. 13. Fig. 15 demonstrates a transition from 2:1 mode 1.01 V at 120 mA to 3:2 mode 1.29 V at 250 mA. The fixed-area VR in Fig. 15b must use seven clusters for the entire transition in order to supply the highest power demand. On the other hand, the DCS VR in Fig. 15a can use only four clusters at the 1.01 V operating point, which allows three clusters to be used by other DCS VRs. The transient response benefits of capacitance sharing are also apparent. The DCS VR settles within 80 ns, which is more than an order of magnitude faster than the 1270 ns settling time for the fixed-area VR. The fast transition is possible due to the dual loop control described in Section III-C. Specifically, coarse-grain capacitance allocation significantly reduces the frequency adjustment range for DCS converters (see Fig. 6b). In addition, the closed-loop control gain can be set higher for the DCS VR due to the lower minimum C_{fly}, which further improves the response time.

Fig. 16 shows DCS VR transitions for the 2:1 conversion mode between 0.85 V at 73 mA with one cluster and 1.01 V at 120 mA with four clusters. The settling time for both up and down transitions is on the order of 35 ns. The fast response time of the DCS VRs in this case can be attributed to near-perfect capacitance transitions that require only very small frequency adjustments, as discussed in Section III-C.

Fig. 17 demonstrates capacitance sharing between two DCS VRs. In Fig. 17a, two loads start with four clusters and each supply 120 mA at about 1 V (note that the voltages are set to slightly different levels for visual clarity). One of the loads then scales to 0.86 V at 80 mA in 2:1 mode and donates three of its clusters to enable the other load to reach 1.3 V at 250 mA in 3:2 mode. Fig. 17b demonstrates the reverse transition. In both cases, the DCS VRs settle to target voltages within 100 ns. DCS VRs are capable of supporting natural current variation at any output voltage. Fig. 18 shows that for a 1.01 V output voltage and a 80 mA to 120 mA current step on a single load, the DCS VR dips 100 mV and recovers within 100 ns.

A testing setup limitation causes artificial voltage undershoot for current step-up transitions and voltage overshoot for current step-down transitions. Both cluster allocation and load configuration settings are stored inside programmable registers that are controlled by the same on-chip serial interface,

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

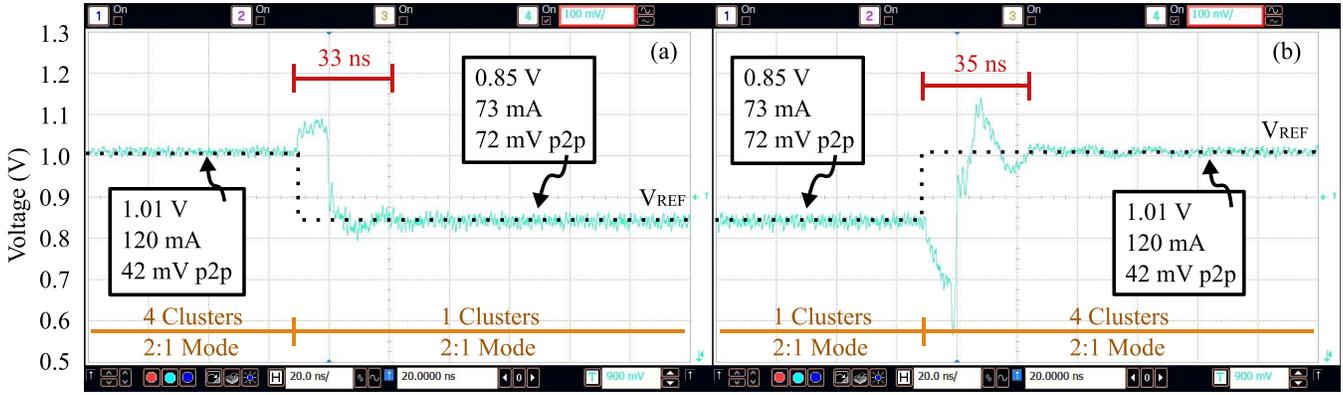BUKREYEV et al.: FOUR MONOLITHICALLY INTEGRATED SC DC–DC CONVERTERS

11

Fig. 16. DCS VR closed-loop 2:1 mode transitions. (a) Step-down transition from 4 clusters at high voltage/current to 1 cluster at low voltage/current. (b) Reverse transition. Dotted line represents V_{REF}.
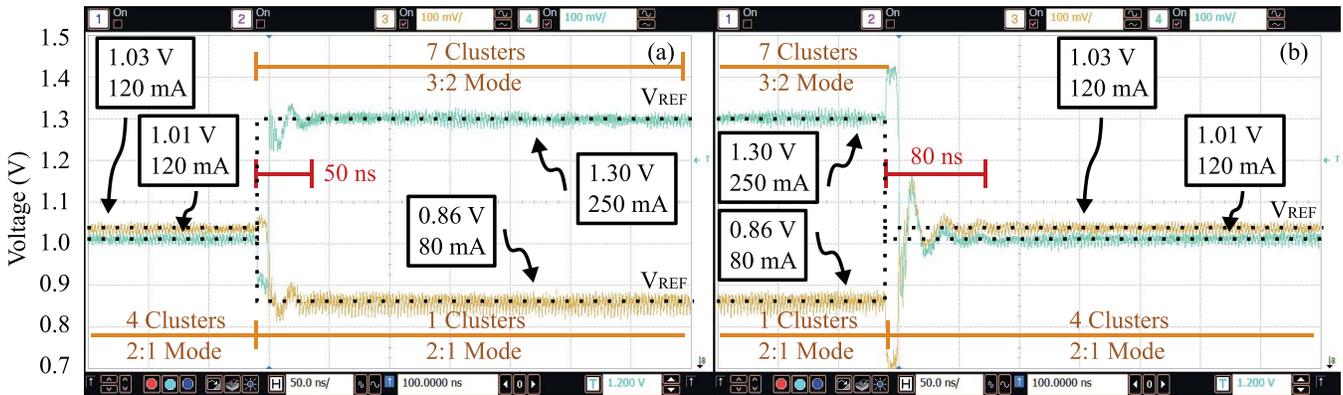

Fig. 17. DCS VR cluster reallocation transitions. (a) Each load starts with 4 clusters at medium current load. The controller then assigns one load to high voltage and current while the other to low voltage and current. The low-voltage load donates 3 of its clusters to the high-power load. (b) The reverse transition. Dotted line represents V_{REF}.
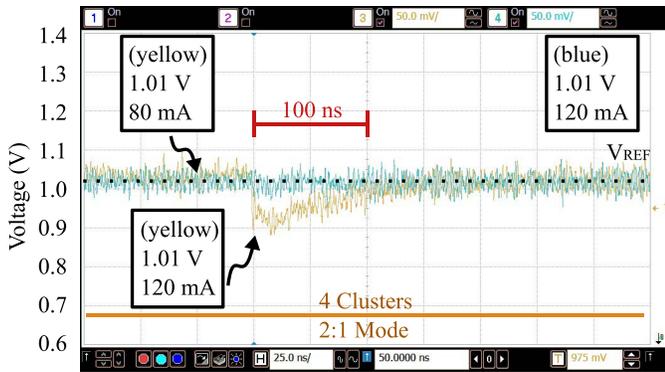

Fig. 18. DCS VR current step at 1 V. Each load has 4 clusters. While one load remains at 120 mA, the other transitions from 80 mA to 120 mA. Dotted line represents V_{REF}.

which requires them to be switched simultaneously. However, transistor bank reconfiguration leads cluster reallocation by up to 10 ns. As a result, load current changes slightly before clusters reconfigure. This artificial voltage overshoot and undershoot due to the testing circuitry miss-timing can be clearly seen in Fig. 16 and is present in all cluster reconfiguration transients. This issue could be addressed by adding a timed reconfiguration delay to the on-chip loads and/or clusters.

E. Comparison With Prior Work

Because area saving is a unique feature of DCS converters, it is difficult to compare it directly to previous work. We compare our work with previously reported fully integrated SCVRs that target similar applications across a variety of processes and summarize our findings in Table II. While there is a diverse variety of SCVRs, we limit our comparison to other converters that are suitable for low-power embedded applications or SoCs that could support DVFS-enabled loads.

The addition of the DCS technique improves transient response time by approximately 10× compared to fixed-area regulators without DCS. This topology also enables use of capacitance slack in DVFS-enabled systems to reduce the overall capacitance area by up to 70% with no cost in performance. In addition, this technique improves voltage regulation range (V_{span}) under cubic power scaling by 1.76× at the cost of 10.4% area increase compared to a fixed-area baseline. For VRs supplying multiple DVFS-enabled loads, operating at maximum output voltage is not practical/possible for all regulators at the same time due to the power-thermal budget of the loads. Our key observation is that the added reconfiguration capability incurs little-to-no performance cost while providing significant area and

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

12

IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS–I: REGULAR PAPERS

TABLE II

COMPARISON TABLE. DCS CONVERTERS ACHIEVE COMPARABLE EFFICIENCY, HIGHER POWER DENSITY, AND WIDER EFFICIENT OUTPUT RANGE COMPARED TO CONVERTERS IN SIMILAR PROCESSES [2], [17]. RESULTS ARE COMPARABLE WITH VRs IN SOI PROCESS THAT ARE LESS IMPACTED BY BACK-PLATE PARASITICS [14]. MEASURED DATA INDICATES LITTLE/NO PERFORMANCE COST TO ADDED RECONFIGURATION CAPABILITY

Metric	[23]	[2]	[14]	[9]	[17]	This Work, no DCS	This Work with DCS
Process	28 nm FDSOI	65 nm CMOS	32 nm SOI	22 nm trigate	45 nm CMOS	65-nm CMOS	
Capacitor Type	MIM	MIM + MOS	MOS	MIM	MOS	MOS	
Interleaved Phases	8	20	32	8	1	8	
Feedback Control	Linear VCO	Cap. Dithering	none	L.B. Hysteretic	Cap. Modulation	Cap. Modulation & Linear VCO	
Switching Freq.	not specified	95 MHz	ext. adjusted	up to 250 MHz	30 MHz	25–400 MHz	
Conversion Ratios	4:3, 3:2, 2:1, 3:1	2:1	3:2, 2:1, 3:1	5:4, 3:2, 2:1, 1:1	3:2	3:2, 2:1	
Input Voltage	1.8 V	2.3 V	2.0 V	1.23 V	1.8 V	2.3 V	
Peak Efficiency η	72.5%	70.8%	79.8%	84.2%	69%	70.9%	
P.D. (W/mm^2) at η	0.31	0.187	0.86	0.126	0.05	0.607	0.550
Converter Area (mm^2)$*$	0.457	0.73	0.378	0.103	0.16	1.11	1.226
Output Voltage	0.2–1.1 V	1.0 V	0.5–1.15 V	0.45–1.0 V	0.8–1.0 V	0.864–1.219 V	0.742–1.367 V
Response time	100–200 ns	Recovery 20 ns	none specified	Simulated †	100–1000 ns	up to 2400 ns	30–250 ns
V$_{nspan}$‡						0.291	0.457

$*$ Switch and capacitance area.

† Limited by test equipment.

‡V_{nspan} cannot be calculated from other works, but is provided for future comparison.

voltage range benefits, in addition to improved transient response.

VI. CONCLUSION

We have demonstrated four DCS converters that are well suited for on-chip integration with multiple DVFS-enabled loads. We show that for cubic power scaling under a power constraint, DCS can reduce required energy-storage area by up to 70%. For a given fixed area, DCS VRs can increase V_{span} while improving transient response times. In combination with the power output, V_{nspan} can be used to holistically evaluate the performance of a regulator for DVFS applications.

REFERENCES

[1] Y. Bai, V. W. Lee, and E. Ipek, "Voltage regulator efficiency aware power management," in *Proc. Int. Conf. Archit. Support Program. Lang. Oper. Syst. (ASPLOS)*, Apr. 2017, pp. 825–838.

[2] S. Bang *et al.*, "A fully-integrated 40-phase flying-capacitance-dithered switched-capacitor voltage regulator with 6 mV output ripple," in *Proc. VLSIC*, Jun. 2015, pp. C336–C337.

[3] T. Van Breussegem and M. Steyaert, "A fully integrated gearbox capacitive DC/DC-converter in 90 nm CMOS: Optimization, control and measurements," in *Proc. COMPEL*, Jun. 2010, pp. 1–5.

[4] N. Butzen and M. Steyaert, "A 94.6%-efficiency fully integrated switched-capacitor DC-DC converter in baseline 40 nm CMOS using scalable parasitic charge redistribution," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Jan./Feb. 2016, pp. 220–221.

[5] L. Chang, R. K. Montoye, B. L. Ji, A. J. Weger, K. G. Stawiasz, and R. H. Dennard, "A fully-integrated switched-capacitor 2:1 voltage converter with regulation capability and 90% efficiency at 2.3 A/mm^2," in *Proc. Symp. Very Large-Scale Integr. Circuits*, Jun. 2010, pp. 55–56.

[6] S. Gangopadhyay, S. B. Nasir, and A. Raychowdhury, "Integrated power management in IoT devices under wide dynamic ranges of operation," in *Proc. ACM/EDAC/IEEE Design Autom. Conf.*, Jun. 2015, pp. 1–6.

[7] W. Godycki, C. Torng, I. Bukreyev, A. Apsel, and C. Batten, "Enabling realistic fine-grain voltage scaling with reconfigurable power distribution networks," in *Proc. MICRO*, 2014, pp. 381–393.

[8] "Intel Turbo Boost technology in Intel Core microarchitecture (Nehalem) based processors," Intel Corp., Mountain View, CA, USA, White Paper 320354-001, Nov. 2008.

[9] R. Jain *et al.*, "A 0.45–1 V fully-integrated distributed switched capacitor DC-DC converter with high density MIM capacitor in 22 nm tri-gate CMOS," *IEEE J. Solid-State Circuits*, vol. 49, no. 4, pp. 917–927, Apr. 2014.

[10] R. Jevtić *et al.*, "Per-core DVFS with switched-capacitor converters for energy efficiency in manycore processors," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 23, no. 4, pp. 723–730, Apr. 2015.

[11] J. Jiang, Y. Lu, W.-H. Ki, S.-P. U, and R. P. Martins, "A dual-symmetrical-output switched-capacitor converter with dynamic power cells and minimized cross regulation for application processors in 28 nm CMOS," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2017, pp. 344–345.

[12] 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. Int. Symp. High-Perform. Comput. Archit.*, Feb. 2008, pp. 123–134.

[13] H.-P. Le, J. Crossley, S. R. Sanders, and E. Alon, "A sub-ns response fully integrated battery-connected switched-capacitor voltage regulator delivering 0.19 W/mm^2 at 73% efficiency," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2013, pp. 372–373.

[14] H.-P. Le, S. R. Sanders, and E. Alon, "Design techniques for fully integrated switched-capacitor DC-DC converters," *IEEE J. Solid-State Circuits*, vol. 46, no. 9, pp. 2120–2131, Sep. 2011.

[15] T. N. Miller, X. Pan, R. Thomas, N. Sedaghati, and R. Teodorescu, "Booster: Reactive core acceleration for mitigating the effects of process variation and application imbalance in low-voltage chips," in *Proc. Int. Symp. High-Perform. Comput. Archit.*, Feb. 2012, pp. 1–12.

[16] T. S. Muthukaruppan, M. Pricopi, V. Venkataramani, T. Mitra, and S. Vishin, "Hierarchical power management for asymmetric multi-core in dark silicon era," in *Proc. Design Autom. Conf.*, May 2013, Art, no. 174.

[17] Y. K. Ramadass, A. A. Fayed, and A. P. Chandrakasan, "A fully-integrated switched-capacitor step-down DC-DC converter with digital capacitance modulation in 45 nm CMOS," *IEEE J. Solid-State Circuits*, vol. 45, no. 12, pp. 2557–2565, Dec. 2010.

[18] E. Rotem, A. Naveh, A. Ananthakrishnan, E. Weissmann, and D. Rajwan, "Power-management architecture of the Intel microarchitec-ture code-named sandy bridge," *IEEE Micro*, vol. 32, no. 2, pp. 20–27, Mar./Apr. 2012.

[19] L. G. Salem, J. G. Louie, and P. P. Mercier, "Flying-domain DC–DC power conversion," *IEEE J. Solid-State Circuits*, vol. 51, no. 12, pp. 2830–2842, Dec. 2016.

[20] M. D. Seeman, "A design methodology for switched-capacitor DC-DC converters," Ph.D. dissertation, Dept. EECS, Univ. California, Berkeley, Berkeley, CA, USA, May 2009.

[21] H.-P. Le, M. Seeman, S. R. Sanders, V. Sathe, S. Naffziger, and E. Alon, "A 32 nm fully integrated reconfigurable switched-capacitor DC-DC converter delivering 0.55 W/mm^2 at 81% efficiency," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2010, pp. 210–211.

[22] T. Souvignet, B. Allard, and X. Lin-Shi, "Sampled-data modeling of switched-capacitor voltage regulator with frequency-modulation con-trol," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 62, no. 4, pp. 957–966, Apr. 2015.

[23] T. Souvignet, B. Allard, and S. Trochut, "A fully integrated switched-capacitor regulator with frequency modulation control in 28-nm FDSOI," *IEEE Trans. Power Electron.*, vol. 31, no. 7, pp. 4984–4994, Jul. 2016.

[24] M. Steyaert *et al. DCDC Performance Survey*. Accessed: Sep. 27, 2017. [Online]. Avaliable: http://homes.esat.kuleuven.be/~steyaert/DCDC_Survey/DCDC%20converters%20Survey.xlsx

[25] G. Villar-Piqué, H. J. Bergveld, and E. Alarcón, "Survey and benchmark of fully integrated switching power converters: Switched-capacitor versus inductive approach," *IEEE Trans. Power Electron.*, vol. 28, no. 9, pp. 4156–4167, Sep. 2013.

[26] X. Wang, "Intelligent power allocation," ARM, Cambridge, U.K., White Paper ARM DTO 0052A, 2017.

[27] B. Zimmer *et al.*, "A RISC-V vector processor with tightly-integrated switched-capacitor DC-DC converters in 28 nm FDSOI," in *Proc. VLSIC*, Jun. 2015, pp. C316–C317.

Ivan Bukreyev received the dual B.S. degree in electrical and computer engineering and physics from the University of Florida, Gainesville, FL, USA, in 2013. He is currently pursuing the Ph.D. degree in electrical and computer engineering with Cornell University. His current research focuses on an analog circuits and computer architecture co-design approach to reconfigurable switched-capacitor voltage regulators for embedded processors. His other research interests include long-range low-BER synchronization for IoT nodes using peer-to-peer communication.
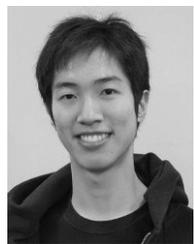
Christopher Torng received the B.S. degree in electrical and computer engineering from Cornell University, Ithaca, NY, USA, in 2012. He is currently pursuing the Ph.D. degree in electrical and computer engineering with Cornell University. His research interests span across computer architecture and VLSI, with a focus on high-performance and energy-efficient heterogeneous systems based on both hardware accelerators and the potential of integrated voltage regulation.

Waclaw (Wacek) Godycki received the B.S. degree in electrical engineering from Columbia University, New York, NY, USA, in 2004, and the Ph.D. degree in electrical engineering from Cornell University, Ithaca, NY, USA, in 2014. Between the degrees, he spent four years with the Advanced Process Development Group, Analog Devices, Boston, MA, USA. Since 2014, he has been with Eridan Communications, Santa Clara, CA, USA, where he has been involved in highly efficient polar amplifiers for radio communications and dynamic, and tracking power supplies. His research interests include wide-band, efficient RF and mixed-signal circuits, communication standards, and ultra-fast tracking supplies.

Christopher Batten (S'05–M'10) received the B.S. degree in electrical engineering from the University of Virginia, Charlottesville, VA, USA, in 1999, the M.Phil. degree in engineering from Cambridge University, Cambridge, U.K., in 2000, and the Ph.D. degree in electrical engineering and computer science from the Massachusetts Institute of Technology, Cambridge, MA, USA, in 2010. In 2010, he joined the Department of Electrical and Computer Engineering, Cornell University, Ithaca, NY, USA, where he is currently an Associate Professor. His research interests include computer architecture, VLSI, parallel programming frameworks, and hardware design methodologies. He was a recipient of the Cornell Engineering Research Excellence Award (2015), AFOSR YIP (2015), the Intel Early Career Faculty Award (2013), NSF CAREER (2012), and DARPA YFA (2012). His teaching has been recognized with the Ruth and Joel Spira Award for Excellence in Teaching (2016) and the Michael Tien '72 Excellence in Teaching Award (2013).

Alyssa Apsel joined Cornell University in 2002, where she is currently an Associate Professor of electrical and computer engineering. The focus of her research is on low power and reconfigurable RF design and power-aware mixed signal circuits and design for highly scaled CMOS and modern electronic systems. She received best paper awards at ASYNC 2006 and IEEE SiRF 2012, had a MICRO Top Picks paper in 2006, received the College Teaching Award in 2007 and the National Science Foundation CAREER Award in 2004, and was selected by *Technology Review Magazine* as one of the Top Young Innovators in 2004. She has served as Associate Editor of various journals, including the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS I AND II, as the Chair of the Analog and Signal Processing Technical Committee of ISCAS 2011, as the Deputy Editor in Chief of *Circuits and Systems Magazine*, and on the Board of Governors of the IEEE CAS.