Enabling Optimal Power Generation of Flow Cell Arrays in 3D MPSoCs with On-Chip Switched Capacitor Converters

Halima Najibi*, Jorge Hunter‡, Alexandre Levisse*, Marina Zapater*†, Miroslav Vasic‡, David Atienza*

*Embedded Systems Laboratory (ESL), EPFL, Switzerland
†REDS Institute, University of Applied Sciences Western Switzerland (HEIG-VD, HES-SO), Switzerland
‡Centro de Electrónica Industrial, UPM, Spain

Abstract— Flow cell arrays (FCAs) provide efficient on-chip liquid cooling and electrochemical power generation capabilities in three-dimensional multi-processor systems-on-chip (3D MPSoCs). When connected to power delivery networks (PDNs) of chips, the current flowing between FCA electrodes partially supplies logic gates and compensates over 20% $V_{dd}$ drop in high-performance 3D systems. However, operation voltages of CMOS technologies are generally higher than the voltage corresponding to the maximal FCA power generation. Hence, directly connecting FCAs to 3D MPSoC power grids results in sub-optimal performance. In this paper, we design an on-chip direct current to direct current (DC-DC) converter to improve FCA power generation in high-performance 3D MPSoCs. We use switched capacitor (SC) technology and explore different design space parameters to achieve minimal area requirement and maximal power extraction. The proposed converter enables a stable and optimal voltage between FCA electrodes. Furthermore, it allows us to dynamically control FCA connectivity to 3D PDNs, and switch off power extraction during chip inactivity. We show that activity-based dynamic FCA switching extends by over 5% drop below $V_{dd}$ and reduces area requirement down to 1 to the case they are directly connected to 3D PDNs. In addition, connecting multiple flow cells to a single optimized converter reduces area requirement down to 1.26%, while maintaining IR-drop below 5%. Finally, we show that activity-based dynamic FCA switching extends by over 1.8× and 4.5× electrolytes lifetime for a processor duty-cycle of 50% and 20%, respectively.

Index Terms—3D MPSoCs, FCA technology, on-chip power generation, 3D power delivery networks, DC-DC converter.

I. INTRODUCTION

Three-dimensional multi-processor systems-on-chip (3D MPSoCs) enable energy-efficient high density computing and provide ultra-wide communication bandwidth requirements for next-generation applications [1]. However, they are generally challenged by rising heat dissipation difficulty with the number of stacked dies, due to higher power consumption per surface unit and low thermal conductivity of bonding layers. Furthermore, 3D integration increases power delivery complexity in multi-processor architectures due to resistive losses in through-silicon-vias (TSVs) and metal wires, as well as 3D routing congestion constraints [2]. In this context, flow cell array (FCA) technology, first introduced in [3], promises to address 3D thermal and power challenges. FCAs consist of micro-fluidic channels etched in the silicon substrate of 3D MPSoC dies. They provide combined on-chip liquid cooling and power generation capabilities due to heat-accelerated electrolyte reactions. When connected to 3D MPSoC power networks, FCA-generated current partially powers logic gates and reduces voltage drop across metal lines, therefore limiting timing violations and system performance degradation [4].

As FCAs connect to 3D MPSoC power delivery networks (PDNs), their generated power depends on the voltage between flow cell electrodes. In particular, peak power generation is achieved with operation voltages between 0.55$V$ to 0.62$V$ for vanadium redox-based FCAs, with fluid temperatures of 30°C and 60°C, respectively [4] [5] [6]. However, $V_{dd}$ values in state-of-the-art high performance systems are generally between 0.75$V$ and 1.2$V$ [7]. Therefore, directly connecting FCA electrodes to 3D PDNs leads to sub-optimal power generation performance and prevents full exploitation of FCA benefits. In this context, we design in this paper an on-chip switched capacitor (SC) voltage regulator to serve as an interface between FCAs and 3D MPSoC PDNs. The proposed SC converter enables optimal operation of FCAs, by providing a stable voltage that leads to maximal power generation. Furthermore, it allows to disconnect FCAs in case of chip inactivity, hence preventing excess input power that is dissipated by delivery grid resistance and chip leakage. Our contributions can be summarized in the following:

- We design a direct current to direct current (DC-DC) converter to supply a stable voltage to FCAs in 3D MPSoCs, leading to optimal on-chip power generation performance. We use SC technology and explore different design-space parameters to achieve high power density and low area requirement [8] [9].
- We show that using the proposed voltage regulator, FCAs generate up to 123% higher power compared to the case where they are directly connected to the PDN of a high-performance 3D MPSoC, operating at 1.1$V$ $V_{dd}$ and achieving a fluid temperature up to 41 °C.
- We show that optimizing DC-DC converters to connect multiple FCA cells allows us to reduce the overall additional area requirement to less than 1.26%, keeping IR-drop across the chip under the 5% $V_{dd}$, which is a typical constraint of high-performance IC designs [4].
- We show that controlling FCA connectivity to 3D MPSoC power networks using SC converters prevents electrolyte reactions during chip inactivity. Hence dynamically switching on and off FCA power generation extends by up to 1.8× FCA reservoir lifetime for a 3D MPSoC duty-cycle of 50%, and by over 4.5× for 20% duty-cycle.
II. BACKGROUND AND RELATED WORK

A. 3D MPSoC Design with Flow Cell Arrays

3D MPSoCs are getting the attention of IC design engineers due to promising advantages in terms of computing density, heterogeneity, and communication bandwidth [1]. However, TSV-based 3D integration presents additional challenges related to thermal and power management. Power density increases with the number of stacked dies, generating heat that becomes difficult to dissipate due to the low thermal conductivity of silicon and other bonding materials [10]. In addition, power delivery is challenging as higher currents traverse power delivery TSVs and metal lines, causing a voltage drop that affects system performance. In this regard, the FCA technology addresses 3D MPSoC power and thermal challenges, by providing combined on-chip liquid cooling and electrochemical power generation [3] [4]. FCAs use micro-channels etched in the silicon substrate of 3D MPSoC dies, as shown in Figure 1. The channels are filled with an electrolytic liquid flow that adsorbs heat generated by the switching activity of chips. Furthermore, high channel temperatures increase the rate of the electrochemical reactions, which generates an electrical current that can be supplied to logic gates. Hence, FCAs effectively transform heat into available generated power in high-density and high-performance 3D MPSoCs [3].

FCA power generation depends on both the voltage between electrodes and liquid temperature. For vanadium-based redox flows used in this work, peak power generation is achieved around 0.6V, for different liquid temperatures, as shown in Figure 2. The authors in [4] propose a design and analysis methodology for 3D MPSoCs with integrated FCAs, using a fine-grained electro-thermal simulator [6]. For a lossless use of on-chip generated power, they directly connect FCA electrodes to nearby power delivery metal lines in the back-end-of-line (BEOL) of dies. They show that FCA-current can reduce the voltage drop across the power grid of a high-performance processor by 20%, hence improving voltage delivery network efficiency, at no extra cost in TSVs or grid density. In the same way, FCAs allow to relax power delivery system requirements of 3D MPSoCs with specific voltage constraints [4]. Although directly connecting FCAs to 3D MPSoC PDNs shows important improvements in power efficiency, their power generation capabilities are limited when operating at $V_{dd}$ of high-performance systems, typically around 1V. Hence, voltage regulation is critical to ensure optimal performance of FCAs, and full exploitation of their potential.

B. Voltage Regulator Design for Efficient IC Power Delivery

Voltage Regulators (VRs) are used for power management of modern 3D-ICs, providing voltage levels that are different than the standard printed circuit board (PCB) supply. Integrated VRs also enable fast voltage scaling to improve core performance while maintaining acceptable power consumption and voltage-drop levels [11]. In addition, VRs allow to decouple power supply from logic circuits and keep constant voltage levels in case of supply noise and transient load changes. The most common types of VRs used in ICs to step up voltage are switching regulators, which function by temporarily storing charge in magnetic or electric fields and then discharging it at a different voltage level. Two main types of switching regulators exist: inductor-based [12] and capacitor-based [13] [14].

Inductor-based switching regulators, used in IC power delivery networks, generally consist of buck-boost or buck-boost-derived topologies [15] [16]. They can achieve high power conversion efficiency rates and are easy to control. However, on-chip integration of inductors remains an important challenge of inductor-based VRs. Particularly, the low quality of air-core spiral inductors makes them unsuitable for ICs [15], and microfabricated inductors, although promising, are not well developed yet [16]. A successful implementation of an inductor-based VR is found in 4th generation Intel Core SoCs [12]. This integrated VR design achieves a maximum conversion efficiency of 90% and output power of 108W, but has a very large area requirement of 175mm², of which 160mm² are occupied by inductors alone.

Capacitor-based converters, known as SC converters, consist exclusively of capacitors and transistor switches. In comparison to inductor-based VRs, SC converters are easier to integrate and require significantly lower chip area. Common SC converter designs achieve efficiencies up to 80% [13] [14], sizing around 0.3mm² for an output power of 4.2mW. Furthermore, designs including new technologies such as deep trench capacitors [17] can achieve 85% efficiency with 30× lower area requirement. Due to the significantly lower cost in chip area, we focus in this work on capacitor-based technology to design a converter that meets voltage and power efficiency requirements of 3D MPSoCs with integrated FCAs.
Fig. 3: FCA connectivity to the 3D MPSoC power delivery grid via integrated DC-DC voltage regulator

**III. INTEGRATED SC DC-DC CONVERTER DESIGN FOR FCA VOLTAGE REGULATION**

This section describes the design methodology of a 1:2 SC converter to provide optimal FCA operation voltage. Thus, it connects FCA cells to 3D MPSoC power grids, each with their respectful voltage domain, as shown in Figure 3. It is primarily designed for minimal area, while aiming for the highest possible output power.

**A. SC Converter State-Space Model**

The typical 1:2 SC converter consists of four transistors and a flying capacitor $C_{fly}$, as shown in Figure 4. This topology can be simplified to the equivalent converter circuit model in Figure 5. It comprises of two resistors $R_{in}$ and $R_{f}$, modeling conduction and switching losses, respectively, and a capacitor $C_{eq}$ to account for first-order circuit dynamics. For a target output voltage $V_{out} = V_{dd}$ and FCA operation voltage $v_{in} = V_{FCA}$, the resistances $R_{v}$ and $R_{f}$ are determined as $R_{v} = \frac{v_{FCA} - v_{out}}{I_{in} - I_{out}}$ and $R_{f} = \frac{v_{out}}{I_{out}}$, where the Voltage Conversion Ratio (VCR) corresponds to the particular SC topology. The capacitive component $C_{eq}$ is calculated by analyzing circuit equations in the Laplace domain [18]. Thus, characterizing a given SC converter design requires to determine the electrical parameters of its circuit. To do so, a matrix-based methodology is developed [19], and improved taking into account the significant effect of on-chip bottom plate capacitance [20]. In this paper we use this generalized methodology to the specific use-case of FCAs, taking into account their electrical characteristics. Although the design and modeling steps are applicable to any SC converter topology, we apply it exclusively to a 1:2 VCR converter.

For any given SC converter, the following matrices and vectors are defined, where $n$ is the number of capacitors and

$$i, v \in \mathbb{R}^{n}$$

and $y \in \mathbb{R}$.

**Fig. 4: 1:2 SC Converter Circuit**

**Fig. 5: 1:2 SC Equivalent Converter Model**

$$i = -A^{-1}Bv - A^{-1}DU$$

(3)

Considering the capacitor’s fundamental equation $i = C\dot{v}$, $\dot{v}$ is expressed as:

$$\dot{v} = -C^{-1}A^{-1}Bv - C^{-1}A^{-1}DU$$

(4)

Which is a state-space equation of the form:

$$\dot{x} = Ax + Bu$$

(5)

This process is repeated for each switching phase $i$ of the SC converter. By solving the state-space equation 5 of each phase, we obtain a solution in the form:

$$v_{i}(t) = \Phi_{i}(t)v(0) + \Gamma_{i}(t)U$$

(6)

Assuming that in steady-state $v((k+1)T) = v(kT)$, we calculate $v(0)$, the equilibrium value of the voltage at the beginning of each cycle. Hence, we calculate $\Delta v$ and derive the charge across each capacitor over T. This procedure allows to approximate the expected converter output, and build the equivalent circuit model from Figure 5.

Additionally, with few modifications to the proposed procedure, first-order converter dynamics are calculated and incorporated into the circuit model. To reduce our analysis to the first-order dynamics, we compute the dominant eigenvalue from the state transition matrix (\(\Phi_{i}\)) which represents most of the dynamics of the system. The discrete model of the system can then be described as follows:

$$y[k+1] = \lambda_{max} y[k] + (1 - \lambda_{max})(g_{1}v_{in} + g_{2}v_{load})$$

(7)

This equation can be readily transformed into the Laplace domain from which the output impedance $Z_{out}$ is derived. It is represented as a parallel RC branch in our model, taking into account $R_{v}$ and $C_{eq}$. 

$$i_{in} = \frac{v_{in}}{R_{eq}}$$

$$v_{out} = \frac{v_{in}}{R_{eq}}$$

$$i_{out} = \frac{v_{out}}{R_{eq}}$$

$$v_{eq} = \frac{v_{in} + v_{out}}{2}$$

$$R_{eq} = \frac{R_{v} + R_{f}}{2}$$

$$C_{eq} = \frac{C_{fly} + C_{out}}{2}$$

$$Z_{out} = \frac{1}{i_{out}v_{out}}$$
B. Design Space Parameters

To design an SC converter, typical design parameters are transistor and capacitor sizing, and switching frequency:

- Transistor and capacitor sizing determines the SC converter circuit resistance and capacitance, which are important components responsible for conduction and switching losses, respectively.
- Switching frequency directly affects power loss in the circuit. In general, $P_{\text{loss}} \propto k f_{\text{clk}}$, where $k$ depends on the converter topology. However, in the case of high-performance ICs, this dependence does not hold for higher frequencies, as additional losses occur due to parasitic components [14].

The effects of each design parameter on switching and conduction losses are correlated. Hence, for a given design we apply the methodology in Section III-A to calculate the converter performance parameters for different configurations. To build the state-space model, we use transistor and capacitor models from a 32nm CMOS technology [17].

C. Converter Design Exploration

The implemented state-space model allows us to calculate converter performance for a given target output voltage. Indeed, modern processors typically operate at different $V_{dd}$ values depending on load and performance. Besides, design points that are optimized for a given output voltage are not necessarily optimal at slightly varying output voltages. In this context, we develop an algorithm that cycles through a wide range of combinations of design parameters, with over 10 million evaluated design points. Since the converter is expected to output several voltage levels, the performance results are weighted according to the target load profile to obtain averaged performance metrics.

Figure 6 shows different evaluated SC converter design points with respect to their total area requirement, their voltage conversion efficiency, and output power. This figure indicates a clear trade-off between efficiency, output power, and area. In particular, output power scales slower than area. In this work, we select the converter design represented by a star in Figure 6 for FCA voltage regulation. This design has the lowest area requirement and the highest output power density. Alternatively, depending on the design constraints, other configurations can be selected with higher efficiency or output power.

Fig. 6: Evaluated DC-DC converter design points

IV. Experimental Setup

To evaluate FCA performance using the proposed SC converter, we design a two-layer 3D MPSoC composed of a multicore computing layer and a memory layer. The memory layer contains four 2nd generation HBM memories with 4 DRAM layers. Each HBM has a base die size of 71mm², and consumes 15W [21]. We base the architecture and power profile of the layer on the 12-core IBM POWER8 processor [7], and consider its implementation in 32nm CMOS technology (used to build the converter state-space model). The processor die size is 649mm² and its power consumption is 190W, with a highly non-uniform powermap. Then, FCAs of 50μm width and 100μm height are etched in the silicon substrate of both dies with a pitch of 50μm, as shown in Figure 1. Each 200μm-long flow cell section is connected to a single SC converter, which is in turn connected to the dies power grid. Furthermore, TSVs are arranged in groups delivering power to independent subgrids. Their diameter and pitch are both fixed at 5μm. Finally, $V_{dd}$ is fixed at 1.1V, corresponding to the maximum processor performance.

During full utilization, the processor in the proposed 3D MPSoC has a higher power consumption than the memory. Therefore, we focus on it to evaluate FCA performance, while the memory layer activity contributes to the overall temperature increase. Thus, we use the four real measured POWER8 powermaps, shown in Figure 7, to evaluate the voltage regulator and FCA power generation efficiency in different load scenarios. All powermaps contain multiple high power density regions mainly concentrated in the computing cores. In powermap (1) from Figure 7, all cores are operating at nominal frequency. In powermaps (2), (3), and (4) six cores operate at maximal frequency and achieve peak power density, while the others are idle (e.g. awaiting data from the memory).

We perform a fine-grained simulation of the 3D MPSoC and assess the thermal and PDN performance. To do so, we use cell dimensions of $200 \times 100 \mu m^2$ and $50 \times 50 \mu m^2$ for the thermal and electrical simulation, respectively. In this way, we evaluate the FCA power generation, converter efficiency, and voltage map. We build a compact flow cell model corresponding to the voltage-power dependency in Figure 2, and a converter circuit model (Figure 5) to perform dynamic SPICE simulations. The cores are switched between 3 different activity levels: idle, nominal, and maximal frequency operation (according to powermaps in Figure 7). As dies have individual PDNs, the total number of FCAs and converters in the 3D MPSoC scales linearly with the number of layers, while their performance depends on the load and voltage level of each die independently.
V. RESULTS AND DISCUSSION

A. Optimized FCA Power Generation

We measure the FCA power generation and DC-DC converter efficiency when cores switch between different activity levels. Particularly, we select core 10 to demonstrate FCA power generation improvement using the proposed converter, as it shows different load and liquid temperature levels (Figure 7). Hence, Figure 8 shows HSPICE transient analysis results of the processor power grid. All three plots present FCA-generated current when directly connected to the power grid operating at $V_{dd}$, FCA-generated current when operating at the optimal voltage enabled by voltage regulation, and output current of the DC-DC converter, when used. Three nodes are selected from the POWER8 core, which correspond to different power densities: maximum (Figure 8a), medium (Figure 8b) and low (Figure 8c).

Our results show that FCAs generate for all the cases over four times higher current when operating at their optimal voltage, supplied via the DC-DC converter, with respect to the case they operate at $V_{dd}$ of the chip. Thus, FCAs generate up to 123% additional power due to voltage regulation. Furthermore, peak FCA power generation is achieved in case the load corresponds to powermap (3) for all three evaluated nodes, due to higher liquid temperature inside the channels, as it passes two high power density cores. Our results also indicate that DC-DC converter output power remains over 90% higher than FCA-generated power when no voltage regulation is used, corresponding to over 82% conversion efficiency.

Using the DC-DC converter allows us to boost the power generation capabilities of individual FCA cells in 3D MPSoCs regardless of load levels of their corresponding chip area, but requires a large number of converters, costing area and design complexity. In this context, we explore in section V-B different configurations in terms of the number of FCA cells per converter.

B. DC-DC Converter Area Optimization

To reduce the number and area requirements of DC-DC converters in 3D MPSoC dies, we explore different design options in terms of the number of FCA cells connected to each converter. To minimize FCA power dissipation in the metal wires to the converter, we group cells in squares so that each $N \times N$ connect to one converter. For each configuration, the input power density of the converter is proportional to the number of connected FCA cells. Therefore, the DC-DC converter is optimized accordingly, for maximal conversion efficiency, following the methodology described in Section III-C. Consequently, by taking into account the resistance of wires connecting FCA cells to converters, we perform a DC voltage analysis of the processor. To do so, we select powermap (3) in Figure 7, which contains the largest concentration of power hotspots, and therefore induces the most critical IR-drop.

Figure 9 presents the maximal IR-drop at the processor when connecting 1 to 36 FCA cells to each DC-DC voltage regulator. As multiple FCA cells connect to a single converter, their output currents traverse longer wires to reach all the loads they supply. Particularly, $10mV$ additional IR-drop occurs when connecting up to 36 flow cells to one converter. Figure 9 also shows the total area percentage requirement of optimized converters for each configuration, and their efficiency. This figure indicates that connecting 25 FCA cells to each converter saves over 2.5% of total chip area, while maintaining the maximal IR-drop value under 5% (typical voltage constraint in high-performance ICs [4]). In addition to area savings, decreasing the number of converters by $25\times$ reduces design and routing complexity in 3D MPSoCs with integrated FCAs.

C. Dynamic FCA Switching

FCA power generation occurs due to electrochemical reactions between electrolytes inside channels. Continuous extraction of power gradually decreases the concentration of reactants, as they are being consumed. Electro-thermal simulations demonstrate that, when liquid flows through FCA channels, up to 1.1% and 0.9% of its electrolytes react, in full load and idle chip scenarios respectively. Furthermore, electrolytes crossover that occurs along membrane-less flow cells causes...
controlling FCA connectivity to 3D MPSoC power grids, based on chip activity. Hence, our results have indicated that dynamically switching on and off the FCA power extraction increases between 1.8× and 4.5× the FCA reservoir lifetime for a processor duty-cycle of 50% and 20%, respectively.

VII. ACKNOWLEDGMENTS

This work has been partially supported by the ERC Consolidator Grant COMPUSAPIEN (GA No. 725657) and the EC H2020 WiPLASH (GA No. 863337).

VI. CONCLUSION

In this paper, we have evaluated the power generation performance of FCAs using high efficiency and low area SC converters. The proposed converters serve as an interface to connect FCA electrodes to 3D MPSoC PDNs. They allow operating flow cells at the voltage value that leads to optimal power generation capability, with \( V_{dd} \) supply of high-performance multi-core platforms. Thus, we have explored different design space parameters to meet the area and output power requirements of the DC-DC converter. Consequently, we have shown that voltage regulators enable up to 123% higher on-chip power generation, with respect to directly connecting FCAs to 3D MPSoC power grids. Furthermore, connecting multiple FCA cell segments to a single optimized converter allows us to limit the cost in total chip area to less than 1.26%. Finally, the proposed converters enable contamination of the oxidant and fuel, and decreases their concentration regardless of reaction rate. Generally, for high-speed flow cells, up to 40\( mA/cm^2 \) current is generated by electrolyte crossover in an open-circuit scenario, corresponding to 0.5% of generated current in optimal FCA power generation conditions [22]. As a result, using DC-DC converters, we can dynamically disconnect FCAs from 3D MPSoC PDNs, and hence prevent excess power extraction when the chip is idle or in low-power operation. In fact, when disconnected, only FCAs cooling capabilities are used, and fuel concentration degradation is limited to cross-contamination effects.

Based on the previous analysis, Figure 10 illustrates the lifetime of a fuel reservoir in different processor duty-cycle levels, with and without dynamically switching FCA power extraction (on and off), as well as the ratio between the two scenarios (in red). When connected to the processor power grid, FCAs operate at their optimal voltage, enabled by DC-DC conversion. FCA lifetime is normalized to the value corresponding to continuous power extraction and chip operation (i.e. 100% duty-cycle). Our results show that, with 50% duty-cycle of the processor, dynamic switching of FCA power generation extends by over 1.8× the lifetime of a fuel reservoir, as only crossover occurs and a large amount of electrolytes are not used up, when no extra power is needed. When the chip is only active 20% of time, dynamic switching extends by over 4.5× the FCA reservoir lifetime. In general, using DC-DC converters clearly improves FCA durability.

REFERENCES

[11] P. Vivet et al. A 220GOPS 96-Core Processor with 6 Chiplets 3D-Stacked on an Active Interposer Offering 0.6ns/mm Latency, 37/s/mm\(^2\) Inter-Chiplet Interconnects and 156mW/mm\(^2\) @ 82%-Peak-Efficiency DC-DC Converters. ISSCC, 2020.