Optimizing the Energy Efficiency of Power Supply in Heterogeneous Multicore Chips with Integrated Switched-Capacitor Converters

Lu Wang[1], Leilei Wang[1], Dejia Shang[1], Cheng Zhuo[2] and Pingqiang Zhou[1]
[1] School of Information Science and Technology, ShanghaiTech University, Shanghai, P.R.China
[2] College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou, P.R.China
{wanglu2,wangll,shangdj,zhoupq}@shanghaitech.edu.cn; czhuo@zju.edu.cn

***Abstract*—Energy efficiency is a major concern in heterogeneous multi-core chips. Due to the switching-capacitor converter (SCC) has wide output voltages and high potential ratio efficiency, they are widely used in multi-core chips. In this paper we propose the optimization of Metal-Insulator-Metal (MIM) capacitance resource allocation and converter ratio selection for SCCs to improve the power efficiency by transforming the mixed integer nonlinear programming (MINLP) problems into a series of convex problems. The experimental results show that our approach can achieve a 9%-13% improvement in power efficiency and can be applied to more complicated heterogeneous multicore scenarios.**

I. Introduction

We are entering a new era of heterogeneous multicore integration – by fully utilizing the unique features of different types of cores, such as CPU cores, GPU cores and accelerators [1]–[6], a heterogeneous multicore chip can significantly enhance the chip performance and improve the overall energy efficiency. What's more, in a multicore chip (especially a heterogeneous one), to achieve even better energy efficiency, the chip is typically divided into several power domains. Each domain is featured with individually adjustable Vdd level that can be supplied by the recently proposed integrated voltage regulators, such as LDOs [7], inductive switching regulators [8], [9] and switching-capacitor converters (SCCs) [10], [11].

Among these three types of regulators, due to their nice features such as wide output voltage, high potential of energy conversion efficiency and high power density, SCCs have been extensively researched and adopted in the recent processors [12]–[14]. As shown in Fig. 1, a step-down SCC converts the high input voltage V_{in} to a lower output voltage V_{out} through a network of flying capacitors and switches. We can also observe that different conversion ratios can be achieved with different typologies. For a SCC with a certain topology, the switches are controlled by periodic clocks with a switching frequency of f_{sw}, and the internal structure operates in two non-overlapping phases: a charging phase ϕ_1 and a discharging phase ϕ_2. During phase ϕ_1, the flying capacitors in the network are connected to the input to get charged, while in phase ϕ_2 those capacitors are delivering power to the output load. Obviously, voltage ripple ΔV exists at the output due to the two-phase operations.

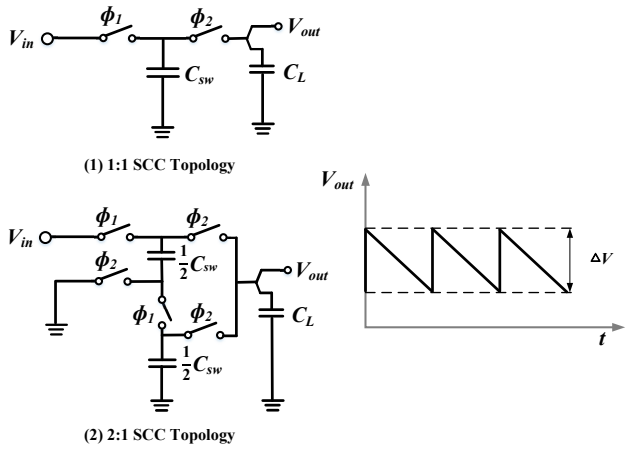

Fig. 1. SCC topology with different conversion ratios.

Energy conversion efficiency is a major concern in the design of a SCC structure [15]–[17]. A SCC delivers energy from the input to the output along with several inherent losses, including switching loss, conduction loss and load power loss (see Table II), which together account for a non-negligible proportion of the total energy provided by the input. As a result, many prior works focus on optimizing the design of a SCC circuit to improve its power efficiency, by tuning its switch width, operating frequency and size of flying capacitors [15], [18]–[23] etc. However, there have been only a few works at the system level that focus on the optimization of SCCs for better energy efficiency in a multicore chip. The authors in [11] proposed to adapt the SCC's switching frequency f_{sw} to the output load so as to improve the energy efficiency. The allocation of limited die area between flying capacitance and decoupling capacitance is explored in [23], with an objective of maximizing energy conversion efficiency while maintaining reasonable supply voltage drop in the power grid. The work [17], [24] first developed a system level efficiency model, then used the model to maximize the energy efficiency in a single power domain, by optimizing the number

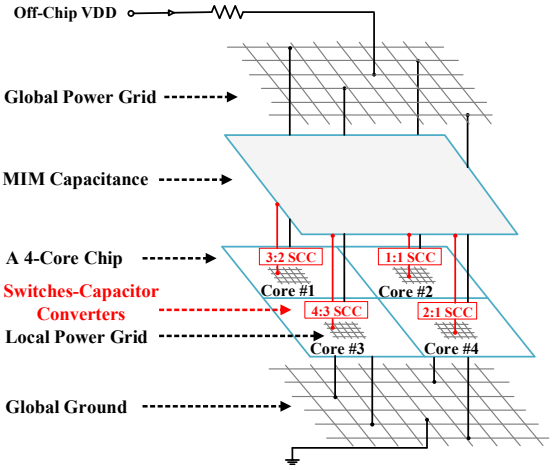

Fig. 2. An example chip consisting of 4 heterogeneous cores with each core powered by an individual SCC. The MIM capacitance resource is shared among the SCCs distributed over the chip.

and distribution of the SCCs, together the size of flying capacitors for each SCC.

Our work is motivated by the following two observations:

1) To the best of our knowledge, there has been no prior work on the capacitance resource allocation for SCCs in a heterogeneous chip with different power domains. Recently MIM capacitance residing between two top metal layers has been utilized to implement the flying capacitors in SCCs [10], [16]. In a heterogeneous chip, such MIM capacitance resource is shared among all the SCCs in different power domains, as shown in Fig. 2. Table I shows an example of two-core heterogeneous chip, with each core powered by an individual SCC. We compare the total energy efficiency of the SCCs in two cases. For Case1, the total amount of 10nF MIM capacitance is allocated to the two SCCs in a simple way – the amount for each SCC is proportional to the area of the core it powers. In contrast, for Case2, we optimized the capacitance allocation by our proposed method. As a result, we can see the energy efficiency can be improved by about 15%. Therefore, it is evident that CAD solutions are desired to allocate the MIM resource among the SCC regulators in a heterogeneous chip.

TABLE I

AN EXAMPLE CHIP WITH TWO HETEROGENEOUS CORES, EACH CORE LOAD'S POWER DEMAND IS SPECIFIED BY ITS MINIMUM SUPPLY VOLTAGE (V_{core}) AND MAX CURRENT DEMAND (I_{core}). TOTAL MIM CAPACITANCE OF 10NF IS SHARED BY TWO SC CONVERTERS (ONE FOR EACH CORE, WITH DIFFERENT CONVERSION RATIOS) IN TWO DIFFERENT WAYS.

$\#SC$	V_{core} (V)	I_{core} (mA)	$Ratio$ $(Vdd = 1.2V)$	Case1		Case2	
				$C(nF)$	η	$C(nF)$	η
SC_1	0.6	0.1	3:2	7.0	**71.75%**	1.7	**85.56%**
SC_2	0.9	0.6	1:1	3.0		8.3	

2) For a processor core with specified power demand, the efficiency of its power supply is also affected by the conversion ratio of its SCC regulators. Fig. 3 presents the efficiency curve of two conversion ratios 2:1 and 3:2 ratio, with an input V_{in} of 1.2V and varying load voltage from 0.2V to 0.8V. We can observe that there exists a large overlap region of the output voltage between these two conversion ratios, where a load voltage can be supplied by either of the conversion ratios. However, the energy efficiency can differ significantly. For example, when the load voltage is 0.58V, the efficiency difference $\Delta\eta$ between two ratios is more than 15%. Hence, for a specific core load, especially when its voltage demand is within the overlapping region between two adjacent conversion ratios, the ratio selection for its SCCs is critical to the efficiency of the power supply system.

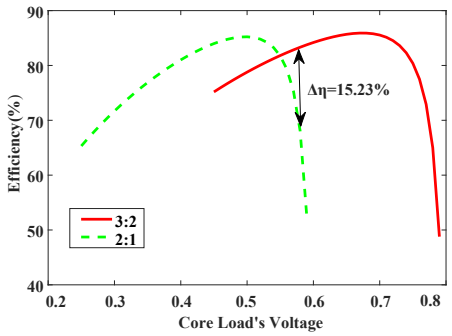

Fig. 3. For a given core load, the ratio choice of its SCCs has big impact on the efficiency of the power supply system. The core's load current is set to 0.1A.

Motivated by the aforementioned observations, in this work, we study the optimization of MIM capacitance allocation among SCCs and the SCCs' ratio selection with the objective of maximizing energy efficiency in a heterogeneous multicore chip. We formulate the optimization problem as an MINLP problem and then solve it by transforming the problem into a series of subproblems. Results show that our proposed optimization could improve the power efficiency by around 9%-13%. *Note that our work addresses the capacitance allocation problem in the early chip planning stage when the core loads are typically characterized by their maximum current demands.*

The rest of this paper is organized as follows. In Section II, we formulate our optimization problem and propose the solving method. We then show the experimental results in Section III. Finally, we make conclusions in Section IV.

II. EFFICIENCY OPTIMIZATION

In this section, we first formulate our efficiency optimization problem as a MINLP problem in Section II-A, then we show how to transform the MINLP problem into a series of simplified convex optimization problems in Section II-B.

A. Problem Formulation

Let's assume that there are M cores in a given heterogeneous chip, and each core is powered by an individual SCC regulator. Note that it is easy to extend our work to handle the case that a core is powered by more than one SCCs. Further, we assume that there are N possible conversion ratios for each

TABLE II
THE MAIN LOSS MECHANISMS FOR A SC CONVERTER [15], [17], [19].

Loss Mechanisms	Formula	Definition of Related Parameters
Conduction Loss	$P_{cond} = \frac{M_{sw} I_{out}^2 R_{on}}{\sigma \gamma C_{sw} f_{sw}}$	I_{out}: The load current of a SCC R_{on}: Equivalent resistance of a switch when it is on C_{sw}: Total flying capacitance of a SCC f_{sw}: The stitching frequency of a SCC M_{sw}: A topology-dependent parameter σ: A fitting parameter γ: A topology-dependent parameter N_{sw}: Number of switches in a SCC, a topology-dependent parameter V_{dd}: The supply Vdd to turn on the switches C_{gate}: The per-unit-width gate capacitance of the switches ΔV: The output voltage ripple of a SCC M_{topo}: A topology-dependent parameters N_{phase}: The number of interleaving stages in a SCC
Gate-drive loss	$P_{sw} = N_{sw} f_{sw}^2 V_{dd}^2 C_{gate} \sigma \gamma C_{sw}$	
Load power loss	$P_{ripple} = \frac{1}{2} I_{out} \Delta V = \frac{I_{out}^2}{2 M_{topo} f_{sw} N_{phase} C_{sw}}$	

SCC. Then we can use a binary variable $x_{m,n}$ to indicate that the m-th SCC selects the n-th conversion ratio, where $m \in [1, M]$ and $n \in [1, N]$. Therefore, we have $M \times N$ binary variables of $x_{m,n}$s. If the m-th SC converter selects n-th ratio, $x_{m,n}$ is set to be 1, otherwise, it is set to be 0. Since each SCC can only use one conversion ratio at a time point, for the m-th SCC, we have $\sum_{n=1}^{N} x_{m,n} = 1$.

According to Table II, which shows the main loss mechanisms of a SC converter, for the m-th SCC the total power loss can be expressed as

$$\begin{aligned} P_{loss}^m(x_{m,n}) &= P_{cond}^m(x_{m,n}) + P_{sw}^m(x_{m,n}) + P_{ripple}^m(x_{m,n}) \\ &= \frac{M_{sw}^m(x_{m,n}) \cdot {I_{out}^m}^2 R_{on}}{\gamma^m(x_{m,n}) \cdot \sigma f_{sw} C_{sw}^m} \\ &\quad + N_{sw}^m(x_{m,n}) \cdot \gamma^m(x_{m,n}) \cdot f_{sw}^2 V_{dd}^2 C_{gate} \sigma C_{sw}^m \\ &\quad + \frac{{I_{out}^m}^2}{M_{topo}^m(x_{m,n}) \cdot 2 f_{sw} N_{phase} C_{sw}^m} \\ &= e_1^m(x_{m,n}) \cdot C_{sw}^m + \frac{e_2^m(x_{m,n})}{C_{sw}^m} \end{aligned} \tag{1}$$

where

$$e_1^m(x_{m,n}) = N_{sw}^m(x_{m,n}) \cdot \gamma^m(x_{m,n}) \cdot f_{sw}^2 C_{gate} \sigma V_{dd}^2 \tag{2}$$

$$e_2^m(x_{m,n}) = {I_{out}^m}^2 \left(\frac{1}{2 M_{topo}^m(x_{m,n}) N_{phase} f_{sw}} + \frac{M_{SW}^m(x_{m,n}) R_{on}}{\gamma^m(x_{m,n}) \sigma f_{sw}}\right) \tag{3}$$

Note that the m-th SCC has N possible conversion ratios and each ratio corresponds to a different set of topology-related parameters such as M_{sw}^m, γ^m, N_{sw}^m and M_{topo}^m (see Table III), so in Equations (1) to (3), these parameters are linear functions of the binary variables $x_{m,n}$s. For instance, N_{sw}^m in fact can be expressed as a linear combination of $x_{m,n}$s: $N_{sw}^m = \sum_{n=1}^{N} (N_{sw}^{m,n} \cdot x_{m,n})$, where $N_{sw}^{m,n}$ is the number of switches in the n-th topology (i.e., conversion ratio).

TABLE III
TOPOLOGY-DEPENDENT PARAMETERS OF FOUR CONVERSION RATIOS IN THE LITERATURE [15], [17]. NOTE THE INPUT VOLTAGE OF THE SCCS, V_{in}, IS ASSUMED TO BE 1.2V.

Conversion Ratio	V_{nl}	N_{sw}	M_{topo}	M_{sw}	γ
2:1	0.6V	4	2	2	2
3:2	0.8V	7	9/8	2	1
4:3	0.9V	10	8/9	7/3	2/3
1:1	1.2V	2	1/2	1	1

Therefore, to optimize the overall energy conversion efficiency of the power supply system in a heterogeneous multicore chip, we minimize the total power loss of the SCCs

$$P_{loss}^{total} = \sum_{m=1}^{M} \sum_{n=1}^{N} x_{m,n} \cdot P_{loss,n}^m(x_{m,n}) \tag{4}$$

The optimization is subject to the following constraints:

1) each SCC only chooses one valid conversion ratio, i.e.,

$$\sum_{n=1}^{N} x_{m,n} = 1, \forall m = 1, 2, ..., M \tag{5}$$

2) the total flying capacitance of all the SCCs should not exceed C_{tol}, the total available MIM capacitance in the chip, i.e.,

$$\sum_{m=1}^{M} C_{sw}^m \leq C_{tol} \tag{6}$$

3) finally, the demands of the core loads should be met [17]. This can be expressed as

$$C_{sw}^m \geq \frac{I_{out}^m}{M_{topo}(x_{m,n}) \cdot f_{sw} N_{phase} (V_{nl}^m(x_{m,n}) - V_{core}^m)} \tag{7}$$

where $m = 1, 2, ..., M$. Note that V_{nl} represents the output voltage of a SCC when no load is attached to it and is thus topology-dependent.

The above optimization problem has both continuous variables C_{sw}^ms and binary variables $x_{m,n}$s. Besides, there exists nonlinear terms in both the objective function (see Equation (1)) and the constraints (see Equation (7)). Therefore, the formulated optimization problem is actually a MINLP problem.

B. Solve the problem

It's well known that MINLP is hard to solve [17], so in our work we propose an approach to transform the problem in Section II-A into a series of convex problems with only

continuous variables. Next we will introduce how to eliminate the binary variables $x_{m,n}$s in the MINLP formulation.

Remember that we introduce $x_{m,n}$s to characterize the fact that, for each SCC powering one given core load, it has $K_m = N$ choices of the conversion ratios. So if we enumerate all the possible N ratios for each SCC in a naive way, we will break down the MINLP problem in Section II-A into N^M subproblems. Although it is easy to verify that each subproblem is a simple convex optimization problem with only continuous variables C_{sw}^ms, the number of the subproblems (i.e., the solution space for enumeration) is huge for big M and N. Therefore, the key problem is how to reduce the number of possible ratio choices K_m for each SCC.

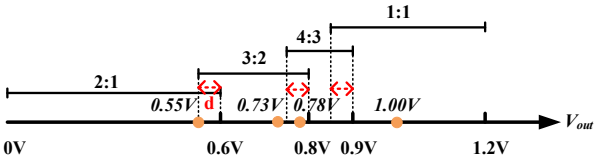

Fig. 4. The output voltage range is divided into several regions, where each region stands for different possible ratio could be employed. And the yellow points on the axis represent four different load voltages.

For the load voltage range [0, 1.2V] shown in Fig. 4, we mark the no-load output voltage V_{nl} of each SCC conversion ratio (i.e., 0.6V, 0.8V, 0.9V and 1.2V). Therefore, each ratio covers a certain output voltage range, starting beyond its left neighbor V_{nl} and ending up to its own V_{nl} (see Fig. 4). After that the voltage range is divided into several sub-regions according to the different SCC ratios that could be employed in that sub-region, and as a result part regions are overlapping regions while the others can be assumed to be attributed to only one ratio. Here, we define the parameter d to indicate the length of the overlapping region.

With the aid of Fig. 4, we can explain our two key ideas to reduce the possible ratio choices of m-th SCC, say K_m (whose original value is 4 here):

1) Given a core load voltage V_{core}, the valid ratio for its SCC needs to guarantee that $V_{nl} > V_{core}$. For example, for the load voltage V_{core} =0.73V, we can only choose the converter ratio whose V_{nl} is larger than 0.73V (see Fig. 4). So the possible ratio could be 3:2, 4:3 and 1:1. As a result, we reduce the possible ratios K_m to 3 for this SCC.

2) For a certain V_{core}, if it falls in the overlapping region of two adjacent ratios, then we set $K_m = 2$, otherwise, we set $K_m = 1$. Therefore, for the SCC with a load voltage 0.73V, the possible ratios K_m is 1 and its ratio is 3:2.

Let us use an example to show how our idea works. In a 4-core chip, each core has a SCC to supply its load voltage 0.73V, 0.55V, 1.00V and 0.78V respectively (see Fig. 4). According to our ideas, the possible ratios for each SCC are {3:2}, {2:1, 3:2}, {1:1} and {3:2, 4:3} respectively, which is also shown in Fig. 5. We define one ratio configuration as one kind of ratio combination of each SCC, and the total ratio configuration $K_{total} = \prod_{m=1}^{M} K_m$. Since the total ratio configuration K_{total} is $4 * 4 * 4 * 4 = 256$ at beginning, we now reduce it to $1 * 2 * 1 * 2 = 4$.

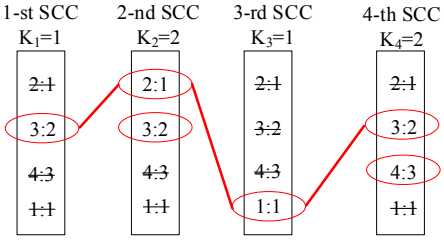

Fig. 5. In this 4 SCC case with each converter owning 1, 2, 1 and 2 ratio candidates respectively, the total number of ratio configuration is 4. And the linked ratios here denote the first ratio configuration.

In each specific ratio configuration, the conversion ratio is determined for each SC (see Fig. 5). As a result all the $x_{m,n}$s variables are also determined and can be eliminated. Hence, for a specific ratio configuration, the MINLP problem in Eq. 4 can be reduced to a convex problem which only has variables of C_{sw}s. And we could easily solve it with CVX [25].

Since we have K_{total} ration configurations, we would have K_{total} convex problems. After we transform the MINLP problem into K_{total} subproblems, we can solve the subproblems one by one. Assuming each subproblem has a optimized loss $P_{loss,i}, i = 1, 2, ..., K_{total}$, the optimized power loss of the original problem is the minimal value among them, i.e., $P_{loss} = min\{P_{loss,1}, P_{loss,2}, ..., P_{loss,K_{total}}\}$. The optimized ratio and flying capacitance of each SCC are the optimized result of the subproblem, which gets this P_{loss}.

III. Experimental Results

In this section, we present results to show the effectiveness of our optimization work.

A series of multi-core benchmarks are tested, including a heterogeneous 4-core, 8-core and 16-core. We get the voltages and currents in each core of our benchmarks from a reasonable scaling of the value in [17].

Given the core load information, we assume that each core's SCC has four possible conversion ratios to choose. The ratio configurations and related parameters for the SCCs are shown in Table III and Table IV. The convex optimization problems in the experiments are solved by CVX [25].

TABLE IV
Parameters of SCCs.

Param.	V_{dd}	N_{phase}	f_{sw}	C_{gate}	R_{on}	σ
Value	1.2V	8	200MHz	3fF/μm	130$\Omega \cdot \mu$m	512μ/(μF·MHz)

In our work, we try to maximize the efficiency of the power supply in a multicore chip by two ideas: 1) optimizing the allocation of on-chip MIM capacitance resource among the SCCs powering the heterogeneous cores, and 2) selecting the best conversion ratio for each SCC. We summarize our results in Table V. Columns 2-4 lists all the SCCs in each benchmark, together with the voltage and current demands for the cores. Since there is no prior work that we can directly compare with, to show the effectiveness of our optimization work, we compare our approach with the following three approaches that apply either zero or only one of our ideas:

TABLE V
COMPARISON OF FOUR OPTIMIZATION APPROACHES.

$\#Core$	$\#SC$	V_{core} (V)	I_{core} (A)	$Method\ 00$				$Method\ 01$				$Method\ 10$				$Our Proposed Approach$				
				C^m_{tol} (nF)	$ratio$	P^m_{loss} (mW)	η $(\%)$	C^m_{tol} (nF)	$ratio$	P^m_{loss} (mW)	η $(\%)$	C_{tol} (nF)	C^m_{opt} (nF)	$ratio$	P^m_{loss} (mW)	η $(\%)$	C_{tol} (nF)	C^m_{opt} (nF)	$ratio$	P^m_{loss} (mW)	η $(\%)$
4	SC_1	0.55	0.08	5	2:1	5.08	73.4	5	2:1⋆ 3:2	5.08	81.2	13	0.88	2:1	11.03	81.1	13	1.04	2:1⋆ 3:2	9.47	86.3
	SC_2	0.68	0.10	2	3:2	15.32		2	3:2	15.32			1.55	3:2	19.08			1.85	3:2	16.41	
	SC_3	0.82	0.30	3	4:3	145.62		3	1:1⋆ 4:3	57.37			6.05	4:3	74.82			4.67	1:1⋆ 4:3	37.37	
	SC_4	0.90	0.35	3	1:1	77.90		3	1:1	77.90			4.52	1:1	52.13			5.44	1:1	43.59	
8	SC_1	0.55	0.15	8	2:1	9.53	76.8	8	2:1⋆ 3:2	9.53	82.6	41	1.94	2:1	17.92	81.5	41	2.25	2:1⋆ 3:2	15.84	85.9
	SC_2	0.65	0.18	6	3:2	18.93		6	3:2	18.93			3.29	3:2	29.80			3.83	3:2	26.22	
	SC_3	0.65	0.18	5	3:2	21.35		5	3:2	21.35			3.29	3:2	29.80			3.83	3:2	26.22	
	SC_4	0.72	0.20	4	3:2	30.65		4	3:2⋆ 4:3	30.65			3.65	3:2	33.11			4.25	3:2⋆ 4:3	29.14	
	SC_5	0.78	0.25	5	4:3	62.89		5	4:3	62.89			5.97	4:3	53.74			6.95	4:3	47.22	
	SC_6	0.85	0.35	5	4:3	120.42		5	1:1⋆ 4:3	47.30			8.35	4:3	75.23			6.33	1:1⋆ 4:3	37.78	
	SC_7	0.85	0.35	5	4:3	120.42		5	1:1⋆ 4:3	47.30			8.35	4:3	75.23			6.33	1:1⋆ 4:3	37.78	
	SC_8	0.90	0.40	3	1:1	101.57		3	1:1	101.57			6.15	1:1	50.34			7.23	1:1	43.18	
16	SC_1	0.55	0.08	5	2:1	5.08	75.2	5	2:1⋆ 3:2	5.08	82.0	67	0.96	2:1	10.15	81.5	67	1.13	2:1⋆ 3:2	8.87	86.3
	SC_2	0.68	0.10	2	3:2	15.32		2	3:2	15.32			1.65	3:2	18.09			1.92	3:2	15.89	
	SC_3	0.82	0.30	3	4:3	145.62		3	1:1⋆ 4:3	57.37			6.67	4:3	68.63			5.08	1:1⋆ 4:3	34.46	
	SC_4	0.90	0.35	3	1:1	77.90		3	1:1	77.90			5.00	1:1	47.31			5.93	1:1	40.21	
	SC_5	0.55	0.08	5	2:1	5.08		5	2:1⋆ 3:2	5.08			0.96	2:1	10.15			1.13	2:1⋆ 3:2	8.87	
	SC_6	0.68	0.10	2	3:2	15.32		2	3:2	15.32			1.70	3:2	17.61			2.00	3:2	15.33	
	SC_7	0.82	0.30	3	4:3	145.62		3	1:1⋆ 4:3	57.37			6.67	4:3	68.63			5.08	1:1⋆ 4:3	34.47	
	SC_8	0.90	0.35	3	1:1	77.90		3	1:1	77.90			5.00	1:1	47.31			5.93	1:1	40.21	
	SC_9	0.55	0.15	8	2:1	9.53		8	2:1⋆ 3:2	9.53			1.81	2:1	19.02			2.12	2:1⋆ 3:2	16.63	
	SC_{10}	0.65	0.18	6	3:2	18.93		6	3:2	18.93			3.17	3:2	30.79			3.77	3:2	26.56	
	SC_{11}	0.65	0.18	5	3:2	21.35		5	3:2	21.35			3.17	3:2	30.79			3.77	3:2	26.56	
	SC_{12}	0.72	0.20	4	3:2	30.65		4	3:2⋆ 4:3	30.65			3.40	3:2	35.21			4.00	3:2⋆ 4:3	30.66	
	SC_{13}	0.78	0.25	5	4:3	62.89		5	4:3	62.89			5.56	4:3	57.19			6.53	4:3	49.72	
	SC_{14}	0.85	0.35	5	4:3	120.42		5	1:1⋆ 4:3	47.30			7.78	4:3	80.06			5.93	1:1⋆ 4:3	40.21	
	SC_{15}	0.85	0.35	5	4:3	120.42		5	1:1⋆ 4:3	47.30			7.78	4:3	80.06			5.93	1:1⋆ 4:3	40.21	
	SC_{16}	0.90	0.40	3	1:1	101.57		3	1:1	101.57			5.71	1:1	54.06			6.77	1:1	45.95	

- Method00: in this approach, we do not apply any ideas proposed in this work, namely jointly optimization of MIM capacitance allocation or ratio selection optimization. Instead, the total available MIM capacitance is allocated to each core's SCC proportional to the core's area. Column 5 of Table V lists the amount of MIM capacitance allocated to each SCC by this method. As for the conversion ratio, given a core load's voltage demand V_{core}, we first identify all the valid ratios from Table III that have $V_{nl} > V_{core}$, then pick the ratio with smallest V_{nl} from the valid ratio set for this core's SCC. Column 6 of Table V lists the picked ratio for each SCC in each benchmark.
- Method01: in this approach, we only apply the idea of ratio selection optimization proposed in our work. Column 10 of Table V lists all the possible ratio choices for each SCC and the selected ratio by this method is mark with a '*'. For the MIM capacitance allocation, we allocate the total available MIM capacitance to each core's SCC proportional to the core's area. Column 9 of Table V lists the amount of MIM capacitance allocated to each SCC by this method.

- Method10: in this approach, we only apply the idea of jointly optimization of MIM capacitance allocation proposed in our work. The total available MIM capacitance shared by the SCCs in each benchmark is listed in column 13 of Table V. We then show the optimized MIM capacitance allocation in column 14 for this method. For the ratio selection, given a core load's voltage demand V_{core}, we first identify all the valid ratios that have $V_{nl} > V_{core}$, then pick the ratio with smallest V_{nl} from the valid ratio set for this core's SCC. Column 15 of Table V lists the picked ratio for each SCC in each benchmark.

For all the four methods, we also show the results of total power loss for each SCC (P^m_{loss}) and the overall efficiency of the power supply for the chip (η) for each method in Table V. Our approach has good performance in the following aspects:

- *Power efficiency improvement*. For the 4-core benchmark, comparing Method 00 and Method 01, we see that our ratio selection idea can improve the power efficiency from 73.4% to 81.2%. Comparing with Method 00, Method 10 reduces the total power loss of SCCs thus improves the power efficiency by 7.7%. Overall, our jointly optimization of MIM capacitance allocation and ratio selection optimization improve the power efficiency by more than 10%. Similar results can be observed for the 8-core benchmark and 16-core benchmark in Table V.
- *Subproblem size reduction*. For the three benchmarks with different numbers of cores, each SCC has four ratio candidates to select, the original numbers of subproblems are 4^4, 4^8, 4^{16}. After applying our approach, the numbers of subproblems are respectively reduced to 2^2, 2^4 and 2^8. Hence the K_m reduction technique (see Section II-B) significantly reduces the size of the subproblems.
- *Methodology scalability*. The runtime to solve the 4-core benchmark is 1.31 seconds, and the time for the 8-core and 16-core benchmarks are respectively 7.08 and 161.34 seconds. This shows that our methodology can be applied to larger multicore scenarios. It's also worth to mention that for each benchmark in our experiments, half of the SCCs' load voltages are within the overlapping regions between two adjacent conversion ratios. But in reality less cores may have their load voltages in the overlapping region, so our approach is promising to solve large benchmarks.

IV. Conclusion

In order to improve the power efficiency of the heterogeneous multi-core chip, in this paper we study the optimization of MIM capacitance allocation among SCCs and the SCCs' ratio selection to maximize energy efficiency in heterogeneous multi-core chips. We formulate the optimization problem as an MINLP problem and then solve it by transforming the problem into a series of convex subproblems. Results show that our proposed optimization could improve the power efficiency by around 9%-13%.

References

[1] A. Branover *et al.*, "AMD Fusion APU: Llano," *IEEE Micro*, vol. 32, no. 2, pp. 28–37, 2012.

[2] M. Yuffe *et al.*, "A fully integrated multi-CPU, GPU and memory controller 32nm processor," in *Proceedings of the IEEE International Solid-State Circuits Conference*, 2011, pp. 264–266.

[3] J. Burt, "Intel begins shipping Xeon chips with FPGA accelerators," *eWeek*, 2016.

[4] T. Liang *et al.*, "Paas: A system level simulator for heterogeneous computing architectures," in *Proceedings of the International Conference on Field-Programmable Logic and Applications*, 2017, pp. 1–8.

[5] "Apple A11," available at https://en.wikipedia.org/wiki/Apple_A11.

[6] "Hisilicon Kirin 970," available at https://en.wikichip.org/wiki/hisilicon/kirin/970.

[7] E. J. Fluhr *et al.*, "5.1 POWER8 TM: A 12-core server-class processor in 22nm SOI with 7.6 Tb/s off-chip bandwidth," in *Proceedings of the IEEE International Solid-State Circuits Conference*, 2014, pp. 96–97.

[8] B. Bowhill *et al.*, "The Xeon® processor E5-2600 v3: A 22nm 18-core product family," in *Proceedings of the IEEE International Solid-State Circuits Conference*, 2015, pp. 1–3.

[9] S. M. Tam *et al.*, "Skylake-SP: A 14nm 28-core Xeon® processor," in *Proceedings of the IEEE International Solid-State Circuits Conference*, 2018, pp. 34–36.

[10] P. Meinerzhagen *et al.*, "An energy-efficient graphics processor featuring fine-grain DVFS with integrated voltage regulators, execution-unit turbo, and retentive sleep in 14nm tri-gate CMOS," in *Proceedings of the IEEE International Solid-State Circuits Conference*, 2018, pp. 38–40.

[11] R. Jevtić *et al.*, "Per-core DVFS with switched-capacitor converters for energy efficiency in manycore processors," *IEEE Transactions on VLSI Systems*, vol. 23, no. 4, pp. 723–730, 2015.

[12] S. Bang *et al.*, "A low ripple switched-capacitor voltage regulator using flying capacitance dithering," *IEEE Journal of Solid-State Circuits*, vol. 51, no. 4, pp. 919–929, 2016.

[13] J. Jiang *et al.*, "A dual-symmetrical-output switched-capacitor converter with dynamic power cells and minimized cross regulation for application processors in 28nm CMOS," in *Proceedings of the IEEE International Solid-State Circuits Conference*, 2017, pp. 344–345.

[14] T. M. Andersen *et al.*, "A 10 W on-chip switched capacitor voltage regulator with feedforward regulation capability for granular microprocessor power delivery," *IEEE Transactions on Power Electronics*, vol. 32, no. 1, pp. 378–393, 2017.

[15] Y. K. Ramadass, "Energy processing circuits for low-power applications," Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge, Massachusetts, 2009.

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

[17] P. Zhou *et al.*, "Distributed on-chip switched-capacitor DC-DC converters supporting DVFS in multicore systems," *IEEE Transactions on VLSI Systems*, vol. 22, no. 9, pp. 1954–1967, Sep. 2014.

[18] M. D. Seeman, "A design methodology for switched-capacitor DC-DC converters," Ph.D. dissertation, University of California, Berkeley, 2011.

[19] H.-P. Le *et al.*, "Design techniques for fully integrated switched-capacitor DC-DC converters," *IEEE Journal of Solid-State Circuits*, vol. 46, no. 9, pp. 2120–2131, Sep. 2011.

[20] J. D. Vos *et al.*, "A sizing methodology for on-chip switched-capacitor DC/DC converters," *IEEE Transactions on Circuits and Systems I*, vol. 61, no. 5, pp. 1597–1606, May 2014.

[21] V. S. Sathe *et al.*, "Analysis and optimization of CMOS switched-capacitor converters," in *Proceedings of the ACM International Symposium on Low Power Electronics and Design*, 2015, pp. 327–334.

[22] T. M. Andersen *et al.*, "Modeling and Pareto optimization of on-chip switched capacitor converters," *IEEE Transactions on Power Electronics*, vol. 32, no. 1, pp. 363–377, 2016.

[23] J. S. Seo *et al.*, "Flying and decoupling capacitance optimization for area-constrained on-chip switched-capacitor voltage regulators," in *Proceedings of the Design, Automation & Test in Europe*, 2017, pp. 1269–1272.

[24] P. Zhou, B. Kim, W. Choi, C. H. Kim, and S. S. Sapatnekar, "Optimization of on-chip switched-capacitor DC-DC converters for high-performance applications," in *Proceedings of the IEEE/ACM International Conference on Computer-Aided Design*, 2012, pp. 263–270.

[25] M. Grant *et al.*, "CVX: Matlab software for disciplined convex programming," 2008.