

An Analytical Study of Power Delivery Systems for Many-Core Processors Using On-Chip and Off-Chip Voltage Regulators

Xuan Wang, *Student Member, IEEE*, Jiang Xu, *Member, IEEE*, Zhe Wang, *Student Member, IEEE*, Kevin J. Chen, *Senior Member, IEEE*, Xiaowen Wu, *Student Member, IEEE*, Zhehui Wang, *Student Member, IEEE*, Peng Yang, Luan H.K. Duong, *Student Member, IEEE*,

Abstract—Design of power delivery system has great influence on the power management in many-core processor systems. Moving voltage regulators from off-chip to on-chip gains more and more interest in the power delivery system design, because it is able to provide fine-grained dynamic voltage scaling. Previous works are proposed to implement power efficient on-chip voltage regulators. It is important to analyze the characteristics of the entire power delivery system to explore the tradeoff between the promising properties and costs of employing on-chip voltage regulators, especially the on-chip buck converters. In this work, we present a novel analysis and design optimization platform of power delivery system called PowerSoC. It employs an analytical model to provide an accurate and fast evaluation of important characteristics, e.g. power efficiency, output stability and dynamic voltage scaling, for the entire power delivery system consisting of on-chip/off-chip buck converters and power delivery network. Based on our model, geometric programming is utilized to find the optimal design for different power delivery systems and explore the tradeoff of using on-chip converters. Compared with SPICE simulations, our model achieves a simulation time reduction of six to seven orders of magnitude within 5% model error for the characteristic evaluation of different power delivery systems. By using PowerSoC, various architectures of power delivery systems are optimized for power efficiency under constraints of output stability, area, etc. Simulation results show compared with the conventional architecture, the hybrid one using both on-chip and off-chip converters achieves 1.0% power efficiency improvement and 66.4% area reduction of converters on average. We conclude the hybrid architecture has potential for efficient dynamic voltage scaling, small area and the adaptability of the change of power delivery network parasitic, but careful account for the overhead of on-chip converters is needed.

Index Terms—power delivery system, on-chip voltage regulator, analytical modeling, optimization

I. INTRODUCTION

Due to the power dissipation limitation of transistors, many-core processor system becomes promising to improve system performance and power efficiency instead of feature size scaling alone. Power delivery system is a key subsystem within it, which has important influence towards its performance and power consumption [1]. Voltage regulators are the essential

components of power delivery systems, and traditionally are contained in board-level modules with large inductors or capacitors. However, putting voltage regulators closer to processor becomes a trend of the power delivery system design. Since the off-chip voltage regulators, e.g. buck converters, are implemented in the board-level with large inductors and capacitors, the inductors and capacitors slow down the responding time of their feedback control. The costs and sizes of the capacitors and inductors also severely limit their usage for fine-grained power domain regulation [2]. The interconnect effect between the voltage regulator and processor, e.g. the conduction loss, is considered to be a severe limitation of the power delivery system design [3]. Meanwhile, the growth of dark silicon area provides an opportunity to implement voltage regulators on-chip for better voltage tuning using the dark silicon [4]. All these make it interesting of developing fully integrated on-chip voltage regulators.

The on-chip voltage regulators are usually implemented and integrated with the processor in the same chip package [5]. The voltage regulator designs range from buck converters to switched-capacitor converters to low-dropout linear regulators. Linear regulators have a maximum efficiency limit given by the ratio of output voltage to input voltage, and they suffer from low efficiency at high ratios [6] [7]. In contrast, switching converters can maintain high efficiency across a wide range of output voltages. There are two types of switching converters commonly used, buck and switched-capacitor converters. The switched-capacitor converters can not provide dynamic voltage scaling of arbitrary output voltage, but some discrete voltage levels [8]. However, the buck converter relies on an inductor to generate a step-down voltage on the output capacitor. The buck converter creates a square-wave voltage of varying duty cycles at the output of the power FETs. By adjusting the duty cycle, the output voltage can be dynamically adjusted with a wide range. Hence, the buck converter is widely used in research area and commercial products, because it can maintain power efficiency across a wide range of conversion ratios. Moreover, the on-chip buck converters can reduce the current flowing through the package and alleviate the power losses of power delivery network and off-chip voltage regulators. It is possible for the hybrid power delivery system employing on-chip converters to cover the power losses induced by the on-chip converters, due to the power loss reduction of power delivery network and off-chip converters.

This work is partially supported by GRF620911, GRF620512, and FS-GRF15EG04.

The authors are with the Hong Kong University of Science and Technology, Hong Kong, China. E-mail: {eexwang, jiang.xu, eekjchen}@ust.hk

Copyright (c) 2015 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending an email to pubs-permissions@ieee.org.

In recent years, there has been a surge of interest to implement on-chip buck converters [5] [9] [10] [11] [12]. An on-chip buck converter, operating with high switching frequencies, can obviate bulky inductor and capacitor, and allow them to be integrated on the chip or on the package. It decreases the power delivery path between the converter and the processor decreases, which alleviates the conduction loss of power delivery network. It is also able to provide fast voltage transitions for fine-grained power management. An on-chip converter can easily be divided into multiple parallel copies with little additional overhead, readily providing multiple on-chip power domains. The fast voltage transition is achieved due to the small capacitors and inductors. However, these potential benefits are tempered by low-quality integrated inductor, increased susceptibility to load current steps and induced power and area overhead. Previous works focus on the implementation of on-chip voltage regulators to improve the power efficiency. It is also important to characterize the power delivery system, including on/off-chip voltage regulators and the passive on-chip/board parasitics, and to explore the tradeoff between promising characteristics and costs of using on-chip voltage regulators for many-core processors.

In this paper, in order to explore the tradeoff of employing on-chip voltage regulators, an analysis and design optimization platform of the power delivery system called PowerSoc is proposed to investigate important characteristics, e.g. power efficiency and transient voltage drop, based on our earlier work [13]. It is publicly available with documentation at [14]. It employs an analytical model to achieve a fast system-level evaluation with comparable accuracy, compared with SPICE simulations. Based on our platform PowerSoC, the characteristics of different architectures of power delivery systems are optimized and compared under design constraints of area overhead, transient voltage drop, etc. The hybrid architecture shows the potential of efficient dynamic voltage scaling, small area and the adaptability of the change of power delivery network parasitic, but careful account for the overhead of on-chip voltage regulators is needed.

The rest of the paper is organized as follows. Section II reviews the related work. In Section III, we present an analytical model of PowerSoC to evaluate the important characteristics of the power delivery systems, e.g. the power efficiency and the output voltage stability. The model maps over a multidimensional design variable space, and the design optimization strategy of PowerSoC using geometric programming (GP) is described in Section IV. In Section V, our model is validated by comparing to SPICE simulations, and it achieves comparable accuracy with less simulation time. Based on the model and optimization strategy, important characteristics of different power delivery systems are compared, and some observations are discussed to help effectively utilize on-chip voltage regulators. Section VI concludes the paper.

II. RELATED WORK

Voltage regulator is an essential component of the power delivery systems in most computing systems. Because of the potential of providing fast voltage scaling and multiple on-chip power domains, there was a surge of interest of on-chip

voltage regulators. Hazucha *et al.* proposed a fully integrated linear regulator to achieve a fast response time with small capacitor area [6]. Ramadass *et al.* presented a fully integrated switched-capacitor converter with additional step-down ratios to extend the range of output voltage [8]. Meanwhile, a lot of literatures focus on the fully integrated buck converters, because of high power efficiency across a wide range of output voltages. On-chip integrated buck converter with different inductor implementation schemes are proposed to provide high quality inductors and improve the power efficiency [9] [10]. Sun *et al.* proposed fully integrated buck converter in SiGe BICMOS process, focusing on the converter power stage to improve the power efficiency [15]. Sturcken *et al.* presented a complete 2.5D chip prototype containing fully integrated buck converter circuitry and a realistic digital load [16]. Due to the wide design space, e.g. the selection of channel width of power bridge and switching frequency, modeling and optimization are used to assist converter designs. Lee *et al.* proposed a GP compatible model of on-chip converters with an on-die integrated inductor model to find the optimal converter design [11]. Schrom *et al.* presented a model of monolithic integrated buck converter, and derived an analytical solution of converter designs for maximum power efficiency [17]. Those works mainly focus on the implementation and evaluation of individual fully integrated voltage regulators. However, the on-chip voltage regulator doesn't come for free. They will induce additional power loss, area consumption and increased susceptibility to load current steps. There is a need of a system-level evaluation of the pros and cons of employing on-chip voltage regulators for the power delivery system design.

Some works proposed the case studies of different hybrid power delivery systems employing on-chip buck converters or linear regulators. Kim *et al.* presented a detailed analysis of a 4-core chip-multiprocessor system with power delivery schemes using on-chip integrated buck converters [2]. It demonstrated the potential benefits of improving the system power consumption by providing fine-grained power management and fast voltage scaling. Gjanci *et al.* proposed a hybrid two-stage power delivery system with off-chip buck converters and a tree structure of on-chip linear regulators, which was able to be efficient, simple and small area costly [18]. Yan *et al.* presented an application-aware scheduling strategy to dynamically utilize the on-chip converters for the dynamic voltage scaling sensitive applications with limited area overhead of on-chip converters [19]. Those hybrid power delivery systems are designed and evaluated for targeted platforms. In order to provide efficient design space exploration for arbitrary power domain distribution, modeling and optimization strategy is usually introduced to find optimum power delivery system design and maximize the system power efficiency.

Design space exploration strategies of different hybrid power delivery systems are presented, using the different models and optimization methods [7] [20] [21]. Vaisband *et al.* presented an optimization strategy of the hybrid power delivery system with off-chip switching converters and on-chip linear regulators to maximize the power efficiency using a power efficient clustering method [7]. Zeng *et al.* proposed a simulation-based optimization strategy to find the optimal number of

on-chip linear regulators and input voltages to maximize the power efficiency, by exploiting GPU-accelerated SPICE evaluations and pattern search optimization strategy [20]. Amelifard *et al.* proposed an optimization strategy of the voltage regulator assignment from a collection of existing commercial regulators to build a two-level tree structure and maximize the power efficiency, which is solved as a combinatorial problem using dynamic programming [21]. Those works mainly focus on the design space exploration of the hybrid power delivery systems employing on-chip linear regulators. In this work, we propose an efficient design space exploration framework of the entire power delivery system consisting of on/off-chip voltage regulators and power delivery network to explore the trade-off of employing on-chip voltage regulators, especially buck converters. It provides a detailed and GP-compatible analytical model of the entire power delivery system with models of each component, which achieves an accurate and fast estimation of a variety of important characteristics compared with SPICE simulations, and reduces the computation complexity of the design space exploration through convex formulation. Based on our framework, the characteristics of different architectures of power delivery systems are investigated to quantitatively evaluate the benefits of the hybrid power delivery system employing on-chip buck converters.

III. MODELING OF POWER DELIVERY SYSTEM

With the development of on-chip voltage regulators, engineers have more choices to build up the customized power delivery systems for many-core processors. On-chip voltage regulators are recommended for power delivery of many-core processors, because they can provide multiple power domains and fast voltage scaling more easily. The multi-stage power delivery system using both off-chip and on-chip voltage regulators becomes promising. The conventional design using only off-chip voltage regulators directly steps the power supply voltage down to the core voltage, while a two-stage power delivery system is illustrated in Figure 1. Given an inherent degradation in power efficiency for large conversion ratios, the first stage of off-chip voltage regulators performs the initial step-down to an intermediate voltage. The intermediate power supply then drives the second stage, on-chip voltage regulators, which further steps the voltage down to the core voltage. The number of the on-chip voltage regulators varies with different granularity to provide at most one power domain per core. Different architectures constituting the framework will be investigated to evaluate the potential benefits of on-chip voltage regulators, especially the on-chip buck converters. The quantitative analysis and comparison of different power delivery systems will be discussed in Section V. The on-chip voltage regulators are usually implemented and integrated with the processor in the same package. They can be implemented in the same die of the processor, or in multiple dies integrated in the same chip package, e.g. 3D integration [5] [15] [16]. For the quantitative analysis in Section V, we assume that the on-chip voltage regulators are implemented on different dies, and integrated with the processor in the same chip package using 3D integration with face-to-back bonding. It allows the

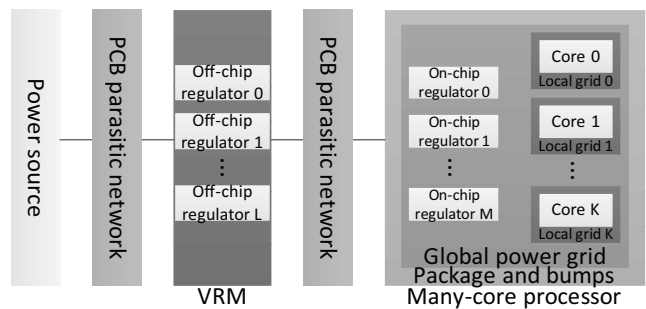


Fig. 1: Overview of a two-stage power delivery system process technology to be optimized separately for different on-chip voltage regulators to improve the power efficiency.

A. Modeling of Switching Mode Voltage Regulators

Tight steady state and dynamic tolerance requirements for core voltage set a big challenge for powering high performance processors. Voltage regulator is an essential component of different power delivery systems. The characteristics of the voltage regulators have significant influence towards the entire power delivery system, e.g. power efficiency, transient response. Interleaved multi-phase buck converter becomes popular to supply high-current processors because of the lower input and output current ripple and fast transient response [22]. The primary building blocks of a multi-phase buck converter is shown in Fig. 2. Each phase of the buck converter is implemented with fixed switching frequency and pulse width modulation. The output voltage is adjusted by modulating the duty cycle of a constant-frequency pulse. The fixed switching frequency will eliminate the undesirable noise in certain frequency band. The pulse width modulation with type-III feedback compensation network is adopted in this paper [12]. Similar phases of the buck converters will be operated in parallel with a common output capacitor. By applying a $360^\circ/N$ phase difference between sawtooth waves of the adjacent phases, the output current ripple can be canceled out while maintaining the fast transient response. N is the number of the parallel phases. Because we focus on high performance many-core processors, continuous mode operation of converters is assumed when analyzing the characteristics.

Power efficiency is one of the important features of voltage regulators, which directly influences the power efficiency of the power delivery system. There are some important power losses which are usually considered in the literatures about the buck converter modeling, e.g. the switching loss of the power bridge and corresponding driver circuits, resistive loss of the power bridge, resistive loss of the inductor and the power of the control circuits [11] [17]. The estimation of those power losses of one phase is presented as follows:

$$P_{driver} = (C_{bridge} + C_{driver})V_{driver}^2 f_{sw} \quad (1)$$

$$P_{Ron} = (DR_{ds,h} + (1-D)R_{ds,l})(I_{ind}^2 + \frac{\Delta I_{ind}^2}{12}) \quad (2)$$

$$P_{R_{ind}} = R_{ind}(I_{ind}^2 + \frac{\Delta I_{ind}^2}{12}) \quad (3)$$

$$P_{control} = I_{control}V_{driver} \quad (4)$$

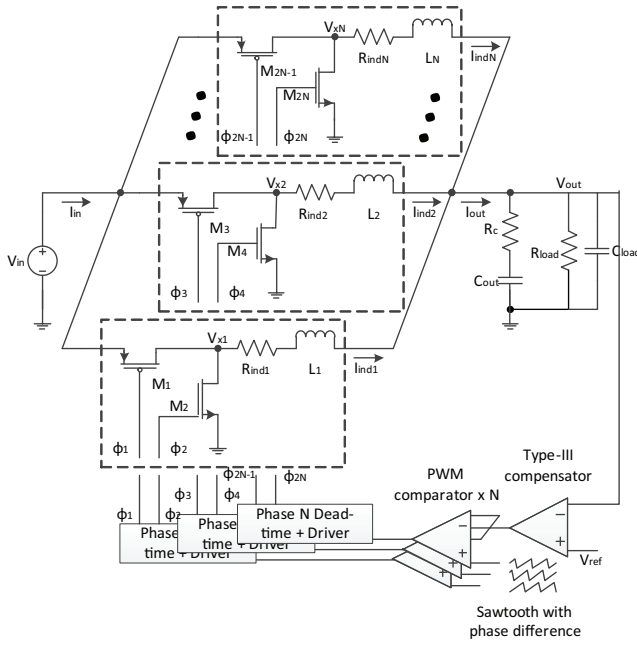


Fig. 2: Simplified block diagram of multi-phase interleaved buck converters

where C_{bridge} and C_{driver} is the effective switched capacitance of the power bridge and drivers, and f_{sw} is the switching frequency. C_{bridge} is estimated by including the gate and diffusion capacitance of the power bridge transistors, and C_{driver} is derived by summarizing the switched capacitance of all the stages in the entire buffer chain of driver circuit. The size of the drivers for the power bridge is designed to be fan-out of 4. Although a larger fan-out can provide additional area and power saving in the driver, the transition of driving signal would be compromised for low duty-cycle signals. V_{driver} is the supply voltage of the drivers and control logic. It may be different from the input voltage V_{in} for off-chip converters and on-chip converters with high input voltage to reduce the power loss of the driver circuit. Those buck converters use two n-channel MOSFETs as the power bridge, and additional circuits, e.g. level shift and bootstrap circuit, are implemented in the driver circuit [23]. $R_{ds,h}$ and $R_{ds,l}$ are the on-resistance of the high-side and low-side transistors of the bridge, and R_{ind} is the equivalent series resistance of the inductor. D is the duty ratio of the gate signal. I_{ind} and ΔI_{ind} are the average value and the peak-to-peak value of the inductor current. $I_{control}$ stands for the supply current of the control circuit of each phase. The device models of the components, e.g. the transistors and inductors, may change with different implementation techniques or accuracy requirements. We will discuss the details in Section IV. The other variables can be derived based on principles of buck converters in Eq. 5 and 6.

$$D = \frac{V_{out} R_{ind} + DR_{ds,h} + (1 - D)R_{ds,l} + R_{load}}{V_{in} R_{load}} \quad (5)$$

$$\Delta I_{ind} = \frac{(V_{in} - V_{out})D}{f_{sw} L_{ind}} \quad (6)$$

Besides the power losses above, there are some other power losses that are usually neglected, e.g. static power loss and switching power loss. However, it is observed that some of

them may be significant during the design space exploration. Hence, they are considered in our analytical model to improve the accuracy. Static power loss is induced by leakage current of the transistors. The leakage current is usually at least three orders of magnitude less than other losses in the proper design. However, the leakage power will increase exponentially with the supply voltage, and may influence the design optimization results. The static power loss is included in our model, and the leakage current of the transistors per unit width is estimated by numerical fitting from SPICE simulations as a function of supply voltages. Simple circuits, e.g. the inverters with different transistor channel widths and supply voltages, are used to estimate the relationship between the leakage current and supply voltage based on the SPICE simulations.

$$V_{in} I_{in} = V_{out} N I_{ind} + N(P_{driver} + P_{R_{on}} + P_{R_{ind}} + P_{stat} + P_{sw} + P_{control}) \quad (7)$$

Switching power loss is induced by the current flowing through the transistors of power bridge during its transition [24]. It mainly consists of the direct path loss and ringing loss. Most switching schemes incorporate dead-time control to ensure both transistors are not conducting for any period of time, and make direct path loss negligible. The dead-time control is considered to improve the power efficiency [25]. However, the ringing loss may be significant for the conventional synchronous buck converter designs [24]. This happens at the instant the power bridge transistors is switching, because the load current and drain-source voltage of the transistor may not go to zero at the same time. The current spike through the power bridge transistors will fluctuate until the ringing is completely damped. It will induce additional power loss due to the ringing current spike. Small testing circuits are built, and the ringing loss is estimated based on SPICE simulations. It is calculated by integrating the energy pumped from input source until the current oscillation is completely damped with the deduction of the load power during this transition period. Hence, we develop the analytical model to estimate the power consumption of converters in Eq. 7.

Besides the power efficiency of the converters, there are other important characteristics which affect the cost and the reliability, e.g. the area and output voltage stability. Output ripple is one criterion of the output voltage stability. In order to calculate the influence of the output capacitor and the number of working phases, some concepts concerning the multiple interleaving should be considered. The output current ripple cancellation due to interleaving technique depends on the number of phases N , and improves with more phases in parallel. Assuming that all of the ripple components of the output current flows through the filter capacitor C_{out} , the worst case peak-to-peak output voltage ripple with multiple phases is derived [26]. C_{load} is the parasitic capacitance contributed by the load circuit of the supported power domain, e.g. the processor cores [27]. The second term represents the ripple current cancellation effect of the interleaving technique.

$$\Delta V_{out,ripple} = \frac{\Delta I_{ind}}{8f_{sw}(C_{out} + C_{load})} \frac{0.25}{D(1 - D)} \frac{1}{N^2} \quad (8)$$

Load transient response is another important criterion to maintain the stability of supply voltage, which determines how much the voltage fluctuates in response to a current change. If the converters is integrated on-chip, the output voltage of the on-chip converter drops much more in response to the load current step [2]. This is because the on-chip capacitor is much smaller than the total decoupling and filter capacitance used for off-chip converters, and large load current steps can rapidly drain out the limited charge stored on the capacitor before the converter loop can respond. Hence, there is a need to evaluate the transient voltage drop for the stability of output voltages. The maximum voltage drop is derived in Eq. 9 [28]. The worst case transient voltage drop tends to approach the open-loop value, when the feedback loop response of the control logic is sluggish. α is an user-specified empirical factor to bring the open-loop estimation into agreement with the actual voltage drop. It changes with the response time of the control strategy of converters. Load-current feedforward technique is able to reduce the response time with considerable overhead by summing the voltage-error signal at the output of compensator and a signal proportional to the load current [28]. In this paper, we adopt the voltage mode feedback control at switching frequencies of hundreds of MHz, and α is about 0.6 using the numerical fitting of SPICE simulations. The estimation is comparable to the results in [9].

$$\Delta V_{out,tr} = \alpha \cdot \Delta I_{out,tr} \cdot \sqrt{\frac{L_{ind}}{C_{out} + C_{load}}} \quad (9)$$

Voltage scaling performance is another important characteristic of converters, which influence the effectiveness of dynamic voltage scaling based power management [2]. When the output voltage scales to a new voltage level, it scales gradually. To ensure sufficient timing margins for processors, the upscaling first attempts voltage scaling and waits until the voltage is stabilized. Once the voltage is stabilized, the processor changes the clock frequency. Downscaling is the opposite. The clock frequency is changed first, and the voltage is scaled latter. This sequence is commonly used in the modern processor [27]. At system level, there are two important features about the dynamic voltage scaling. One is the settling time of voltage scaling, which affects the timescale granularity of dynamic voltage scaling. The other is the energy loss, including the converter and underclocking related energy overhead. The optimum settling time is derived according to the minimum time control law [29]. D_{min} is the average of the duty cycles of the initial and final state of the voltage scaling, and ΔD is the difference between these duty cycles. The voltage mode feedback control can not achieve the optimum settling time. β is used as an user-specified empirical factor to fit the practical settling time of the converters. It is about 2.5 using the voltage mode feedback control at switching frequencies of hundreds of MHz according to SPICE simulations. The settling time is proportional to the product of L_{ind} and C_{out} , while the output ripple decreases with larger ones. The tradeoff between these two features is achieved using the optimization strategy in Section IV.

Voltage scaling will also induce energy overhead, which is divided into two parts [27]. One is the converter induced

energy overhead, when a large surge current flows into and out of C_{out} via the inductor and transistors of power bridge. If the output voltage of the final state, $V_{out,fin}$, is higher than the initial one, $V_{out,int}$, the stored energy on the capacitor and the conduction loss of the inductor and power bridge is provided by the supply V_{in} . If $V_{out,fin}$ is lower than $V_{out,int}$, no work is done by V_{in} . Hence, the converter induced energy overhead is presented as the first term in Eq. 11, assuming the voltage upscaling and downscaling occur evenly. The other is the underclocking related loss, due to the mismatch of the supply voltage and clock frequency of the processor during the voltage scaling. Taking upscaling as an illustration, the clock frequency increases until the voltage scaling settles. During this period, the processor is supplied by an unnecessary high voltage. The underclocking related loss is estimated as the second term of Eq. 11, assuming that the output voltage scales linearly. As shown in Eq. 11, C_{out} plays an important role for both the converter related and underclocking energy overhead. Hence, the voltage scaling energy overhead of on-chip converters also benefits from the reduced filter capacitance.

$$T_{scale} = \beta \sqrt{\frac{2L_{ind}(C_{out} + C_{load})\Delta D}{D_{min}(1 - D_{min})(1 + 0.5\Delta D)}} \quad (10)$$

$$E_{scale} = 0.5(C_{out} + C_{load})V_{in}|V_{out,fin} - V_{out,int}| + C_{load} \min\{f_{sw,int}, f_{sw,fin}\} \times \int_0^{T_{scale}} (V_{out}(t)^2 - \min\{V_{out,fin}, V_{out,int}\}^2) dt \quad (11)$$

The area consumption is also an important issue of the power delivery system design, and the trend slightly changes. Because modern ICs are usually pin limited, the PCB boards are expensive and dark silicon area is observed, it is possible to move a few off-chip converters into the chip. It implies that we need to consider both the on-chip and off-chip area during the power delivery system design. In this paper, we estimate the area consumption of the converters including the power bridge, the driver circuit, the control logic, and the inductor and capacitor of the output filter. The area of the power bridge and the driver circuit is determined by the channel length and width of the transistors, containing the area of the gate, drain and source. The channel length of transistors is related to the fabrication technology, and the channel width will be selected with the optimization technique proposed in Section IV. The area of the drivers is derived by summarizing the areas of all the stages in the entire buffer chain. The area consumption of the control logic per phase is estimated according to [9]. We assume that the output filter capacitor is implemented on-die with the deep-trench and thick oxide MOS capacitance, and the capacitance density is estimated referring to [10]. The on-package integrated inductor is selected to improve the power efficiency of the on-chip buck converter because of its high quality factor [9]. According to the products of Coilcraft [30], the ratio between the inductance and the area is derived.

B. Modeling of Power Delivery Network

In order to capture the overall property of the power delivery system, we also need to pay attention to the passive parasitics of the power delivery network, e.g. the parasitic resistance,

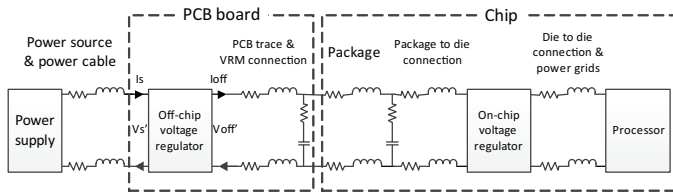


Fig. 3: Model of power delivery network with on-chip voltage regulators

capacitance and inductance of the printed board trace and package. A detailed model of the hybrid power delivery network with off-chip and on-chip voltage regulators is presented in Fig. 3. A ladder RLC network is utilized to capture the parasitics of the power delivery network. The on-board power supply is modeled as a fixed voltage source. The PCB board includes the PCB trace and off-chip decoupling capacitors. The power delivery network of the chip has the following components: the package and corresponding decoupling capacitors, package to die connection, die to die connection, and the lumped VDD and GND power grids of the processor. The package to die connection, e.g. C4 bumps, are modeled as lumped RL pairs that connect the on-chip voltage regulators to the power pins of the package. The on-chip voltage regulators then delivery the power to the processor through the die to die connection and the power grids of the processor. Because we assume the different dies are integrated using 3D integration with face-to-back bonding, the parasitics of the die to die connection are mainly contributed by the through silicon vias (TSVs). The design of the TSVs refers to [31], and the lumped resistance of TSVs is derived based on the density and the resistance per TSV. The other parasitics of the power delivery network, e.g. the inductance and capacitance of the PCB trace, package and power grids of the processor, are adopted from [32], and will be scaled to be consistent with the power consumptions of processors.

IV. DESIGN OPTIMIZATION OF POWER DELIVERY SYSTEMS

The analytical model of the power delivery system is derived including the buck converters and the parasitics of the power delivery network. Those properties of power delivery systems are determined by the selection of the converter design, e.g. channel width of the power bridge, inductance and the capacitance of the output filter. The conversion ratio of the converters in a multi-stage power delivery system is also considered. The intermediate voltage level, V_b , tradeoff the power efficiency of different stages by affecting the conversion ratios. Generally speaking, the power efficiency of converters decreases as the conversion ratio increases. Selecting higher V_b will also reduce conduction loss of power delivery network, by reducing the supply current through the package. Because of the wide design space, an optimization strategy is needed to find the optimal design variables.

In this paper, a method using convex optimization is adopted to find the optimal converter design that maximizes the power efficiency under the constraints of output ripple, transient response, area, etc. Convex optimization based on GP has been widely employed to optimize various mixed-mode circuits [33]. We utilize GP to find the optimal design variables

for different architectures of power delivery systems. Without losing the generality, the optimization strategy of the hybrid architecture combining both on-chip and off-chip converters is illustrated. The design variables and the conversion ratios of on/off-chip buck converters are optimized to maximum the system power efficiency, while provided with the supply voltage of the power delivery system V_{in} , load distribution of the processor power domains $V_{out,ij}$ and $I_{out,ij}$, and design specs. It is assumed that the power delivery system consists of multiple off-chip converters, and an off-chip converter j supports M_j number of on-chip converters. The assignment of on/off-chip converters is based the principle of load balance to make sure each on/off-chip converter supports same amount of power consumption. The design specs constrain performance requirement and the boundary of the design variables, e.g. the maximum output ripples of on-chip and off-chip converters $V_{ripple,on,max}$ and $V_{ripple,off,max}$, the maximum transient voltage drop $V_{tr,on,max}$ and $V_{tr,off,max}$, the maximum settling time of voltage scaling $T_{scale,on,max}$ and $T_{scale,off,max}$, the area constraint $A_{on,max}$ and $A_{off,max}$ and the boundaries of design variables of different components.

In order to apply GP, the device models have to be compatible with GP. Our model leverages the transistor models similar to [33] and inductor model in [17]. For the transistors of the power bridge, since they are in the linear region, R_{ds} is expressed as $\frac{k_{R,h}}{W_h}$ and $\frac{k_{R,l}}{W_l}$, where W is a width of a device, and $k_{R,h}$ and $k_{R,l}$ are process and driver voltage dependent resistance per unit width at minimum gate length for high-side and low-side transistors of the bridge. The capacitance is modeled as $k_{C_{gs}} \cdot W$ and $k_{C_{ds}} \cdot W$, where $k_{C_{gs}}$ and $k_{C_{ds}}$ are process dependent parameters for gate-to-source capacitances and drain/source-to-body capacitances. The static power loss is expressed as $k_{P_{stat}} \cdot W$, and $k_{P_{stat}}$ is the process and supply voltage dependent leakage power per unit width. The switching power loss is assumed to be linearly scaled with the load current at the time of commutation [34]. It is expressed as $k_{P_{sw}} \cdot I_{out} \cdot f_{sw} \cdot k_{P_{sw}}$ is the process and supply voltage dependent ringing loss per unit load current per cycle. All parameters are found for each process technology and voltage level by numerical fitting to SPICE simulation results of simple testing circuits. τ_e is the ratio between the resistance and inductance of the inductors [17]. It is estimated according to the quality factor, which is assumed to be a constant within a certain range of switching frequency.

In order to improve the accuracy, some of the device models are still incompatible with GP, and can not be directly used in the GP optimization, e.g. the voltage level dependent parameters of transistor model and the parasitic resistance parameter of the inductor. We adopt the decomposition method to deal with it. Given a combination of V_b and a certain range of switching frequency, the corresponding parameters of transistor model and inductor model become constant, and the formulation of the power delivery system design shown in Alg. 1 becomes a GP problem. The power efficiency of the power delivery system can be optimized using the existing convex solver [33]. The maximum power efficiencies for different combinations of V_b and different frequency ranges will be derived using the convex solver, and then be compared

Algorithm 1 Optimization of the hybrid architecture

Require: workload distribution of power domains, intermediate voltage levels, parameters of device models, boundaries of design variables of converters, design specs

- 1: minimize $\frac{P_{in}}{P_{out}}$
- 2: subject to
- 3: $W_{min,on} \leq W_{h/l,on,ij} \leq W_{max,on}, \forall i, j$
- 4: $f_{sw,on,min} \leq f_{sw,on,ij} \leq f_{sw,on,max}, \forall i, j$
- 5: $L_{ind,on,min} \leq L_{ind,on,ij} \leq L_{ind,on,max}, \forall i, j$
- 6: $\frac{W_{h,on,ij}}{W_{l,on,ij}} = \frac{\mu_{l,on,ij}}{\mu_{h,on,ij}}, \forall i, j$
- 7: $\frac{V_{out,ij}}{V_{b,j}} \frac{R_{load,on,ij} + R_{ind,on,ij} + R_{pin,on,ij} + R_{ds,on,ij}}{R_{load,on,ij}} \leq \frac{D_{on,ij}}{D_{on,ij}}, \forall i, j$
- 8: $\frac{(V_{b,j} - V_{out,ij}) \cdot D_{on,ij}}{f_{sw,on,ij} \cdot L_{ind,on,ij}} = \Delta I_{ind,on,ij}, \forall i, j$
- 9: $(C_{bridge,on,ij} + C_{driver,on,ij}) V_{b,j}^2 f_{sw,on,ij} + (R_{ds,on,ij} + R_{ind,on,ij} + R_{pin,on,ij}) (I_{out,ij}^2 + \frac{1}{12} \Delta I_{ind,on,ij}^2) + P_{stat,on,ij} + P_{sw,on,ij} + V_{b,j} \cdot I_{control,on,ij} \cdot N_{on,ij} + Z_{grid,ij} \cdot I_{out,ij}^2 + V_{out,ij} \cdot I_{out,ij} \leq V_{b,j} \cdot I_{on,ij}, \forall i, j$
- 10: $\frac{\Delta I_{ind,on,ij}}{8 f_{sw,on,ij} (C_{out,on,ij} + C_{load,ij})} \frac{0.25 V_{b,j}}{D_{on,ij} (V_{b,j} - V_{out,ij})} \frac{1}{N_{on,ij}^3} \leq \frac{V_{ripple,on,max}}{V_{tr,on,max}}, \forall i, j$
- 11: $\alpha_{on} \cdot \Delta I_{out,tr,on,ij} \cdot \sqrt{\frac{L_{ind,on,ij}}{C_{out,on,ij} + C_{load,ij}}} \leq V_{tr,on,max}, \forall i, j$
- 12: $\beta_{on} \sqrt{\frac{2 L_{ind,on,ij} (C_{out,on,ij} + C_{load,ij}) \Delta D_{on,ij}}{D_{min,on,ij} (1 - D_{min,on,ij}) (1 + 0.5 \Delta D_{on,ij})}} \leq T_{scale,on,max}, \forall i, j$
- 13: $A_{bridge,on} \sum_i \sum_j (W_{h,on,ij} + W_{l,on,ij}) + A_{control,on} \sum_i \sum_j N_{on,ij} + A_{ind,on} \sum_i \sum_j L_{ind,on,ij} N_{on,ij}^2 + A_{cap,on} \sum_i \sum_j C_{out,on,ij} \leq A_{on,max}$
- 14: $\sum_i I_{on,ij} \leq I_{off,j}, \forall j$
- 15: $W_{min,off} \leq W_{h/l,off,j} \leq W_{max,off}, \forall j$
- 16: $f_{sw,off,min} \leq f_{sw,off,j} \leq f_{sw,off,max}, \forall j$
- 17: $L_{ind,off,min} \leq L_{ind,off,j} \leq L_{ind,off,max}, \forall j$
- 18: $\frac{W_{h,off,j}}{W_{l,off,j}} = \frac{\mu_{l,off,j}}{\mu_{h,off,j}}, \forall j$
- 19: $\frac{V_{b,j}}{V_{in}} \frac{R_{load,off,j} + R_{ind,off,j} + R_{ds,off,j}}{R_{load,off,j}} \leq D_{off,j}, \forall j$
- 20: $\frac{(V_{in} - V_{b,j}) \cdot D_{off,j}}{f_{sw,off,j} \cdot L_{ind,off,j}} = \Delta I_{ind,off,j}, \forall j$
- 21: $(C_{bridge,off,j} + C_{driver,off,j}) V_{b,j}^2 f_{sw,off,j} + (R_{ds,off,j} + R_{ind,off,j}) (I_{off,j}^2 + \frac{1}{12} \Delta I_{ind,off,j}^2) + P_{stat,off,j} + P_{sw,off,j} + V_{driver,off,j} \cdot I_{control,off,j} \cdot N_{off,j} + (Z_{package} + Z_{PCB}) \cdot I_{off,j}^2 + V_{b,j} \cdot I_{off,j} \leq V_{in} \cdot I_{in,j}, \forall j$
- 22: $\frac{\Delta I_{ind,off,j}}{8 f_{sw,off,j} C_{out,off,j}} \frac{0.25 V_{in}}{D_{on,j} (V_{in} - V_{b,j})} \frac{1}{N_{off,j}^3} \leq \frac{V_{ripple,off,max}}{V_{tr,off,max}}, \forall j$
- 23: $\alpha_{off} \cdot \Delta I_{out,tr,off,j} \cdot \sqrt{\frac{L_{ind,off,j}}{C_{out,off,j}}} \leq V_{tr,off,max}, \forall j$
- 24: $\beta_{off} \sqrt{\frac{2 L_{ind,off,j} C_{out,off,j} \Delta D_{off,j}}{D_{min,off,j} (1 - D_{min,off,j}) (1 + 0.5 \Delta D_{off,j})}} \leq T_{scale,off,max}, \forall j$
- 25: $A_{bridge,off} \sum_j (W_{h,off,j} + W_{l,off,j}) + A_{control,off} \sum_j N_{off,j} + A_{ind,off} \sum_j L_{ind,off,j} N_{off,j}^2 + A_{cap,off} \sum_j C_{out,off,j} \leq A_{off,max}$
- 26: $\sum_j \sum_i (V_{out,ij} \cdot I_{out,ij}) = P_{out}$
- 27: $\sum_j I_{in,j} \cdot V_{in} \leq P_{in}$

to find the optimal power efficiency of the system within the entire design space and the corresponding design variables.

The formulation of optimizing on-chip converters are described in Lines 3-13 referring to Section III. The formulation is simplified as one-channel converter instead of the interleaved multiple phase one, because most of the properties of an interleaved converter, except the output ripple and control overhead, can be derived according to an equivalent one-channel converter [22]. Lines 3-5 define the boundaries of the multidimensional variable space of on-chip converters. The power consumption of each on-chip converter is calculated in Lines 7-9. Line 7 estimates the duty cycle, including additional conductive loss from the dedicated bumps connecting the on-package inductors. The conduction loss of the die to die connection and power grids of the processor is included in Line 9, where Z_{grid} includes the resistance of the die to die connection and power grids. The total amount of the die to die connection and power grids is fixed, and each on-chip converter will possess part of it, according to the proportion of its load power. The power consumption of on-chip converters is conducted in Line 9, including the power losses mentioned in Section III. There are some modifications of the equations from equality to inequality in order to apply GP. However, the equality of the equations achieves, if the power losses are minimized. Lines 10-13 shows the constraints of the design specs, e.g. the output ripple, transient voltage drop and area overhead. The influence of the interleaving technique is shown in the calculation of output ripple, and power and area overhead of control logic. The area overhead is constrained in Line 13, where $A_{bridge,on}$ is the area consumption of the bridge and driver circuit per unit width of the power bridge, $A_{control,on}$ is the area overhead of the control logic per phase, and $A_{ind,on}$ and $A_{cap,on}$ are the area overhead of the inductor and capacitor per unit inductance and capacitance accordingly.

After the optimization of the on-chip converters, the input currents of the on-chip converters will be used as the load currents of the off-chip converters. The conduction loss of the PCB trace and package is included in Line 21, where $Z_{package}$ includes the resistance of the package and package to die connection. The formulation of optimizing off-chip converters is described in Lines 14-25, which is similar to that of on-chip converters. The power efficiency of the system will be derived in Lines 25 and 26. Alg. 1 shows the GP optimization formulation with a fixed combination of V_b and a constant τ_e .

V. MODEL VALIDATION AND QUANTITATIVE ANALYSIS OF DIFFERENT POWER DELIVERY SYSTEMS

The analytical model of different components of a power delivery system has been proposed. Based on our model, design variables of different power delivery systems can be optimized under the constraints of output ripple, transient voltage drop and area overhead. It is necessary to validate the accuracy of our model, before the system level analysis is conducted to evaluate the characteristics of different architectures of power delivery systems. In this paper, a comprehensive analysis of different architectures will be carried out to explore the design space of power delivery systems. We will evaluate the conventional architecture using only off-chip buck converters, labeled as *off-chip*. On the contrary, the architecture that

TABLE I: Parameters of on/off-chip buck converters.

	off-chip converter	on-chip converter (1.5 μm)	on-chip converter (130nm)
$k_{R,h} (\Omega \cdot \mu m)$	37600	9740	976
$k_{R,l} (\Omega \cdot \mu m)$	37600	9740	488
$k_{C_{gs}} (nF \cdot m^{-1})$	0.48	2.54	1.54
$k_{C_{ds}} (nF \cdot m^{-1})$	0.20	1.09	0.66
$k_{P_{stat}} (W \cdot m^{-1})$	$4.4 \cdot 10^{-3}$	1.75	$5.8 \cdot 10^{-3}$
$k_{P_{sw}} (nJ \cdot A^{-1})$	56.7	1.4	0.03
$A_{bridge} (\mu m)$	2.7	15.7	0.86
$\tau_e (\mu H \cdot \Omega^{-1})$	50	0.1	0.1
$A_{ind} (mm^2 \cdot nH^{-1})$	$12.5 \cdot 10^{-3}$	0.5	0.5
$A_{cap} (mm^2 \cdot nF^{-1})$	$0.5 \cdot 10^{-3}$	0.02	0.02
$I_{control} (mA)$	2	2	2
$A_{control} (mm^2)$	0.03	0.03	0.03

puts all the converters on-chip will be evaluated, labeled as *on-chip*. It is expected to significantly reduce the power loss of the power delivery network due to the high conversion ratio, but suffer from high power losses due to low quality of on-chip components. It is assumed that the on-chip converters are implemented on different dies, and integrated with the processor using 3D integration. It is able to implement the on-chip converters on different dies using proper technology nodes to improve the power efficiency for different input voltage levels. The architecture using both off-chip and on-chip converters, labeled as *hybrid*, will also be investigated. It increases the flexibility of power delivery system design to combine the advantages of the previous two architectures.

There are 3 kinds of buck converters required to implement different power delivery system architectures. The off-chip converter is used to perform the voltage conversion for *off-chip* and *hybrid*. It utilizes the power trench MOSFETs of Fairchild as the power bridge, and transistors of 1.5 μm CMOS process for the control logic and drivers [35] [36]. For the architecture using only on-chip converters, there is a need of on-chip converters to provide the voltage conversion directly from the on-board power supply. We implemented its power bridge and the driver circuit using 1.5 μm due to high input voltage level. The on-chip converters driven by the intermediate voltage in a multi-stage architectures, are built with the transistors in 130nm [35]. The technology is selected to tradeoff between the conductivity and leakage power of different technology nodes at the interested range of input voltages. For example, the unified leakage power of transistors at 90nm is about 100 times higher than that of 130nm at the supply voltage of 2V. Hence, we extract the parameters of transistors at different technology nodes, and find the proper one according to the requirements of different converters.

A. Simulation Setup

The parameters of the device models of different components will be introduced, before the model validation and the comparison of different architectures of power delivery systems is conducted. The parameters of the MOSFETs of the power bridge of off-chip converters are estimated according to the power trench MOSFET of Fairchild [36], and that of the control logic and drivers are extracted using 1.5 μm CMOS

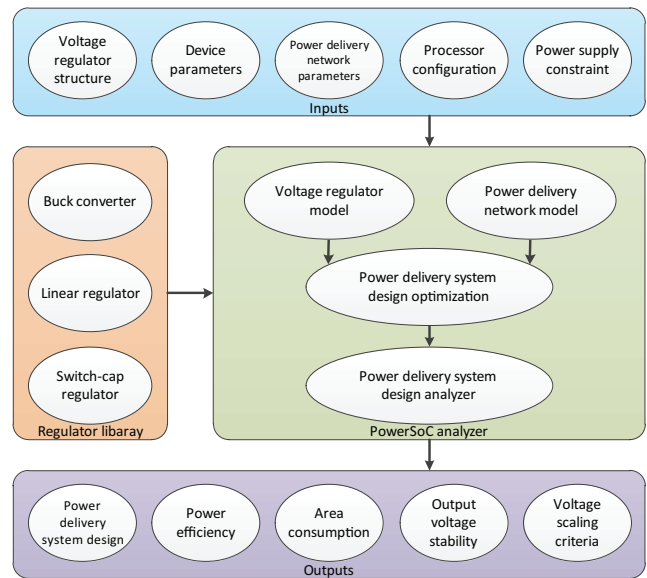


Fig. 4: Overview of PowerSoC

process [35]. The architecture with only on-chip converters uses 1.5 μm CMOS process [35] to implement both power bridge and drivers, because the input voltage of the converters is high. The transistors of on-chip converters driven by intermediate voltage are implemented using 130nm CMOS process [35]. The bulky inductor and capacitor of off-chip converters is estimated according to [37] [38], due to the high requirement of the maximum current of inductor for off-chip converters. The parameters of the inductor and capacitor for on-chip converters are derived based on [9], because it is assumed that the inductor is integrated on package for on-chip converters. The area and power overhead of the control logic is estimated according to [9] [11]. The parameters of the component model are summarized in Table I.

The power delivery system is required to step down the supply voltage from 12V to 1V. The supply voltage of the drivers of off-chip converters is 5V. The parasitic resistance of Z_{grid} is 0.11m Ω or 0.09m Ω with or without the TSVs, and the resistance of $Z_{package}$ and Z_{PCB} are 1.25m Ω and 0.09m Ω for the processor with 256W maximum power consumption [31] [32]. The parasitic capacitance of the processor is around 96nF [27]. In order to maintain the stability of the output voltage, the maximum peak-to-peak output ripple is 10% of its output voltage. The maximum transient voltage drop is also 10% of the output voltage due to a load stop of 50% of the load current [9]. The settling time is measured due to the voltage scaling of 40% output voltage, e.g. from 0.6V to 1V. The slew rate requirement of off-chip converters is 80mV \cdot μs^{-1} [39], and that of on-chip converters is 50mV \cdot ns $^{-1}$ [2]. The area of on-chip converters is 110mm 2 for the processor with 256W maximum power consumption, and that of off-chip converters is 2700mm 2 [40].

B. Power Supply On-Chip (PowerSoC)

In this work, we present the PowerSoC, which is short for Power Supply On-Chip. It is an analysis and design optimization platform of power supply targeted to fast and accurately evaluate the important characteristics of the complex power

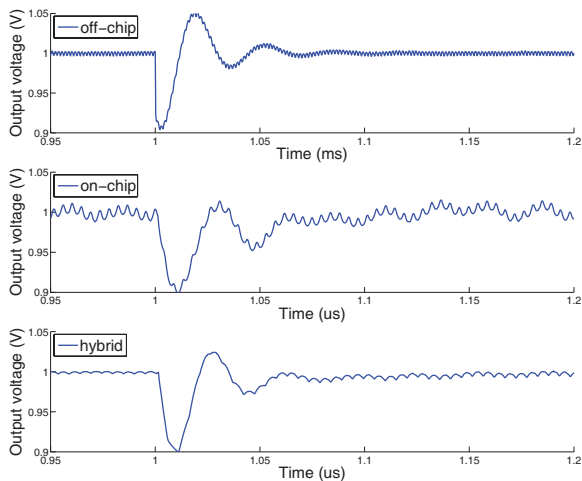


Fig. 5: SPICE simulations of transient voltage drop due to 2A load step for different architectures of power delivery systems including both on/off-chip voltage regulators and power delivery network for multi-core processors. PowerSoC is a C++ based program to analyze the important characteristics of power delivery system, e.g. power efficiency, area and transient response, and optimize the complex power delivery system design to tradeoff among those important characteristics [14]. The overview of PowerSoC is illustrated in Fig. 4. The inputs to the platform include the device parameters of the on/off-chip voltage regulators including the transistors, capacitors and inductors, the parameters of the power delivery network, the configuration and design constraints of multi-core processors and power supply system. The voltage regulator library includes several typical designs of different kinds of voltage regulators. Based on the detailed models of the voltage regulators and power delivery network, a GP based optimization strategy will be used to efficient explore the design space and guide the power supply system design. PowerSoC models the steady state and transient response of different components of the power supply system.

C. Analytical Model Validation

The simulation setups are discussed, and it is necessary to validate the accuracy of our model, compared with SPICE simulations. A comprehensive comparison between our model and SPICE simulation is conducted for a full system validation including both converters and power delivery network. It includes the comparison among different architectures of power delivery systems. The single-stage conventional architecture achieves the highest power efficiency, which directly convert the supply voltage to the core voltage level. However, the architecture using only on-chip converters has to use two-stage scheme to improve the power efficiency of on-chip converters by reducing the conversion ratio, and to get the highest power efficiency. The first stage on-chip converters provide an initial voltage step-down to an intermediate voltage. The intermediate voltage then drives the second stage on-chip converters to the core voltage. The hybrid architecture also adopt two-stage scheme with off-chip converters for the initial voltage

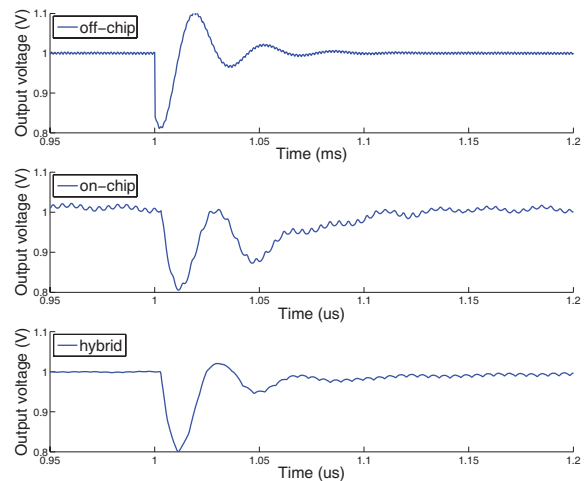


Fig. 6: SPICE simulations of transient voltage drop due to 4A load step for different architectures of power delivery systems

TABLE II: Model comparison between the analytical model and SPICE simulation of different architectures of power delivery systems optimized at workload of 4W.

Architecture	R_{load} (Ω)	SPICE Effi- ciency	Model Effi- ciency	SPICE V_{ripple} (mV)	Model V_{ripple} (mV)	SPICE V_{tr} (mV)	Model V_{tr} (mV)	SPICE T_{scale} (ns)	Model T_{scale} (ns)
off-chip	0.25	75.9%	75.8%	7.3	7.7	98	100	$27 \cdot 10^3$	$26 \cdot 10^3$
off-chip	0.125	75.4%	75.9%	9.4	9.9	99	100	$27 \cdot 10^3$	$25 \cdot 10^3$
on-chip	0.25	50.1%	49.9%	11.5	11.3	102	100	24	23
on-chip	0.125	49.6%	50.1%	9.4	9.2	104	100	34	33
hybrid	0.25	76.3%	76.2%	3.6	3.4	100	100	25	25
hybrid	0.125	76.0%	76.5%	1.1	1.0	99	100	34	33

conversion instead. The parameters of different component models are shown in Table I. The parasitic of power delivery network is linear scaled with the power of processors.

As a case study, the power deliver system is optimized at the workload of 4W or 8W, when converting from 12V to 1V. As mentioned in Section III, the voltage regulators are implemented using the buck converter with pulse width modulation, type-III feedback compensation network and two interleaved phases. The design variables of the converters are derived according to optimization procedure discussed in Section IV. With the converter schematic and optimized design variables, the SPICE models of different architectures are built, and the characteristics of power delivery network are evaluated based on the SPICE simulations. The SPICE simulation results of the output voltage responses of different architectures optimized at 4W are shown in Fig. 5 and Fig. 6, while applying a load step from 2A to 4A or from 0A to 4A with the slew rate of $2A/ns$. The transient voltage drops of different architectures are around $100mV$ or $200mV$ expected as the analytical model evaluations. Table II summarizes the comparison between SPICE simulation and our model at different workload, in terms of the power efficiency, output ripple, transient voltage drop due to load step of 50% load current and settling time of voltage scaling of 40% output voltage. The optimized results based on our model are well matched with SPICE simulations within 5% difference.

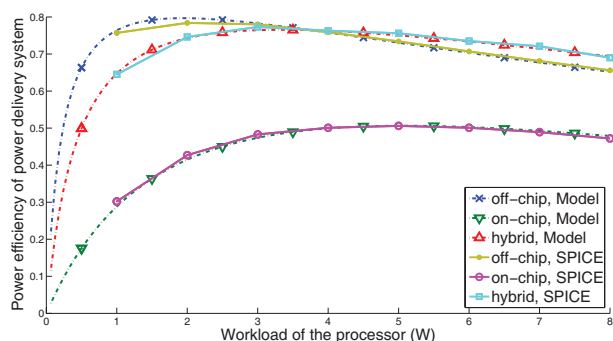


Fig. 7: Model validation of power efficiency curve of different architectures of power delivery systems optimized at 4W

Besides the important features at the point of the optimized workload, it is also important to predict the trend of power efficiency curve of the power delivery system. Power efficiency curve is usually shown in commercial voltage regulator products, because the real workload will fluctuate and the power efficiency should maintain at a high level within a range of workload. The power delivery system is optimized at the average workload instead of maximum power consumption [41]. The average power consumption of the processor is estimated as half of the maximum one [42]. Fig. 7 shows a comparison of the power efficiency curves between SPICE simulation and our model. The power delivery systems are optimized at the workload of 4W, and the SPICE models are built accordingly. SPICE simulations are conducted to obtain the power efficiencies of different architectures at different workload, e.g. every 1W, which are shown as markers in the solid lines in Fig. 7. The dashed lines show the power efficiency curves of different architectures based on our model. Those curves consist of 128 points at different workload for each line, and their markers are used to differentiate the overlapped curves between our model and SPICE simulations. In Fig. 7, the power efficiency curves of our models are well matched with that of SPICE simulations within 2% difference for different architectures of power delivery systems.

Our model is able to provide accurate evaluation of different power delivery designs. The device model of the components, e.g. the transistors and inductors, are derived based on the model parameter estimation. By integrating the device model, our model is built to provide the system level characteristic evaluation of the system. The device model parameters are estimated using the numerical fitting of SPICE simulations of simple testing circuits to improve the accuracy. The SPICE simulation time of the testing circuits is tolerant, e.g. around 3 minutes for the derivation of the unified leakage power of transistors at 130nm, and the parameter estimations are conducted once for each technology. On the other hand, the system level evaluation based on our model will be repeatedly used during the design space exploration to find the optimal design variables. Our model is able to accelerate the evaluation of different designs compared with the traditional circuit simulation engine, e.g. SPICE. Our model is implemented in Matlab, and the SPICE simulation is carried out using HSPICE from Synopsys. We run them on Intel Xeon Processor W5580, and the runtime comparison is investigated. For the conven-

TABLE III: Performance comparison of design optimization strategies for different power delivery systems.

Test bench	Power efficiency (%)		Runtime (sec)	
	This work	APPS	This work	APPS
hybrid 1*2-domain	76.7	76.5	30	89
hybrid 1*4-domain	76.7	76.2	59	150
hybrid 1*8-domain	76.7	75.7	122	438
hybrid 4*8-domain	76.6	75.6	196	2176

tional architecture of power delivery system, the runtime of SPICE simulation will take around 6.8 hours to capture the power efficiency curve with 128 points of workload shown in Fig. 8, while the runtime of the Matlab code is about 1 ms. The simulation time of our model stays almost the same, while that of SPICE simulations increases with more complicated power delivery system. The speedup using our model will be more impressive in the case of the hybrid architecture involving both on/off-chip buck converters, which will cost about 217 hours to derive the entire power efficiency curve.

Our framework provides not only a detailed model of power delivery systems to provide accurate and fast evaluation compared with SPICE simulations, but also a GP-compatible model to facilitate efficient design space exploration. A search-based optimization method is usually used to explore the design space and optimize the hybrid power delivery system designs [20] [21]. On the other hand, we formulate the design optimization as a convex decomposition problem. Each subproblem with fixed power delivery system topology is formulated as a GP problem and solved efficiently using the convex solver [33]. For a fair comparison of the optimization method, we replace the SPICE model in [20] with the proposed analytical model for fast characteristic evaluation. Our analytical model is used in the asynchronous parallel pattern search (APPS) formulation as [20], by assigning the design constraints as the penalty functions. We generate 10 sets of the initial design values randomly from the optimum design with the same order of magnitude to accelerate the convergence and evaluate the average performance of the search-based optimization method. The search-based optimization process is carried out using the APPSPACK 5.0.1 on Intel Xeon Processor W5580, and the comparisons of optimization result and runtime are summarized in Table III. The proposed convex-based optimization strategy achieves efficient design space exploration with good optimization results and small execution time, when the problem scales increase. Our work outperforms the search-based strategy with 1.3% power efficiency improvement and 10 times runtime reduction, when the number of design variables reaches 84 for hybrid 4*8-domain. Our framework achieves a high power efficiency and good scalability of the design optimization, because convex formulation is able to find the global optimal solution with great efficiency even for problems with hundreds of variables [33].

D. Performance Evaluation

With the improvement of semiconductor technologies, more processing units can be integrated on a single chip. Multiprocessor system-on-chip becomes promising to satisfy the growing computation demands from high performance applications.

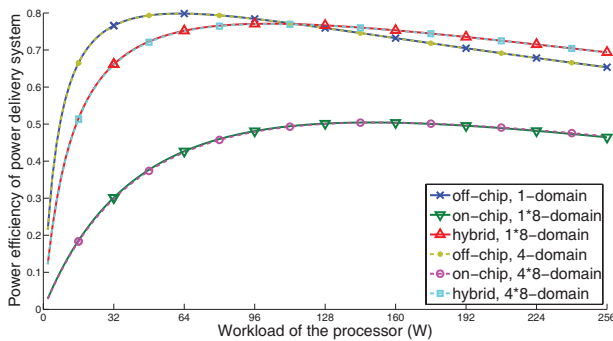


Fig. 8: Power efficiency curve of different architectures of power delivery systems optimized at workload of 128W

TABLE IV: Summary of characteristics of different architectures at 128W workload.

	off-chip 1-domain	on-chip 1*8-domain	hybrid 1*8-domain
<i>power efficiency</i>	75.9%	50.2%	76.7%
A (mm^2)	2700	248	907
T_{scale} (ns)	$28 \cdot 10^3$	39	34
E_{scale} (μJ)	121	0.21	0.18
$\Delta V_{out,ripple}$ (mV)	33	78	85
$\Delta V_{out,tr}$ (mV)	86	100	100

The core voltages and currents of high performance processors are approaching 1V and 130A [22]. As a case study, the power delivery system is designed to support a 64-core homogeneous processor with 128W average power consumption. The design variables of different architectures are optimized at the workload of 128W to maximize the power efficiency under the constraints of output ripple, transient voltage drop and area. Different configurations of the same architecture will also be evaluated, e.g. the conventional architecture with 1 off-chip converter, labeled as *off-chip 1-domain*, and the hybrid architecture with 1 off-chip converter and 8 on-chip converters per off-chip converter, labeled as *hybrid 1*8-domain*. For the architecture using only on-chip converters, the area constraint is alleviated, because it is difficult to put all the on-chip converters into the small area.

The characteristics of the optimized power delivery systems are summarized in Table IV. The hybrid architecture achieves 1.0% power efficiency improvement and 66.4% area reduction of buck converters, compared with the conventional architecture *off-chip*. The on-chip converter is able to provide higher voltage at the package and reduce the supply current through it. It decreases the conduction loss of package and PCB, and the power losses of off-chip converters, due to the smaller load current. These gains compensate the power losses from on-chip converters, and make the power efficiency of *hybrid* and *off-chip* comparable. The area of hybrid architecture is also reduced due to the decreasing load current of off-chip converters. It alleviates design requirements of off-chip converters, e.g. the output capacitance, and decreases the area of 69%. The power efficiency of *on-chip* is 34% lower than others. The power losses of first stage on-chip converter becomes dominate. It is difficult to maintain comparable power efficiency without the help of off-chip converters.

As for the dynamic voltage scaling, the architecture *hybrid*

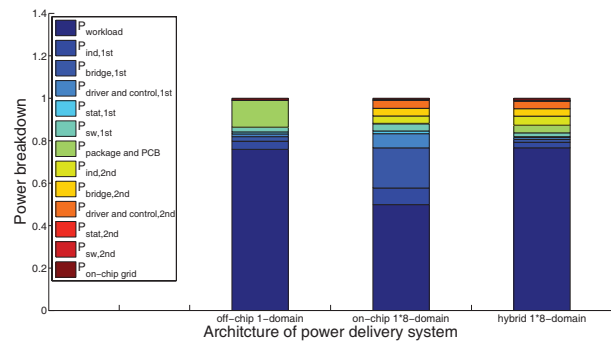


Fig. 9: Power consumption breakdown of different architectures of power delivery systems at workload of 128W

and *on-chip* facilitate fast dynamic voltage scaling by using on-chip converters. Due to the small filter inductance and capacitance, the settling time of on-chip converters is three orders of magnitude less than off-chip converters. It will also reduce the penalty of each voltage scaling. The energy overhead of voltage scaling of on-chip converters is also several hundred times less. The small filter capacitance of on-chip converters brings out the advantages of dynamic voltage scaling, while it will induce large transient voltage drop and make the output voltage less stable. The output ripple of different architectures increases because the optimized design selects single phase to reduce the control overhead. As shown in Table IV, the output ripple and transient voltage drop of on-chip converters are larger than that of off-chip converters.

Besides the important features at the optimized workload, it is also important to show the power efficiency curve, because the real workload will fluctuate around the average value. The power efficiency curves of the optimized power delivery systems are shown in Fig. 8. Each curve consist of 128 points at different workload of every 2W estimated using our model, and the markers are just used to differentiate the overlapped curves. The hybrid architecture achieves a flat power efficiency curve within a large range of workload. It successfully maintain the power efficiency at the workload above 128W, while the power efficiency curve of *off-chip* is significantly affected by the conduction loss of power delivery network. The optimization of off-chip converters can not mitigate the quadratic increase of the conduction loss, as the workload increases. Hence, the power efficiency of *off-chip* is higher at light workload, but significantly decreases with increasing workload. It keeps the conventional architecture from supporting the instantaneous power consumption during burst mode. The architecture *on-chip* is outperformed within the entire range of workloads. The power efficiency curves of different configurations of the same architecture are similar, because it supports a homogeneous processor. Some of the optimized design variables, e.g. the transistor width and the inductance, are linearly scaled with the workload.

The simulation result is explained with the help of the power breakdown shown in Fig. 9, where different kinds of power losses correspond to the notations in Section III. The conduction loss from the interconnect between the converter and processor, e.g. the package and PCB, plays an important role for the conventional architecture. It costs 52.7% of the

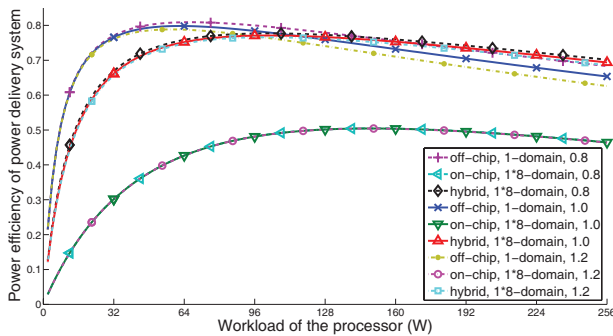


Fig. 10: Power efficiency curve of different architectures for different resistive parasitics of power delivery network

total power loss of the system, and can not be alleviated by the optimization of off-chip converters. The conduction loss will increase quadratically with the increase of workload, and significantly decrease the power efficiency at heavy load. On the other hand, for the architecture using only on-chip converters, the power loss of power delivery network is negligible, because of the small current flowing through the package. However, the power losses of the first stage on-chip converter becomes dominate. The old technology, e.g. $1.5\mu\text{m}$ technology, has to be used to maintain the low leakage power at the high input voltage level, and the power losses are further enlarged due to the high switching frequency for the on-chip implementation of small filter capacitor and inductor.

The hybrid architecture take advantages of these two architectures to facilitate the fine-grained dynamic voltage scaling without inducing significant power losses. Compared with the conventional architecture, the power losses induced by the on-chip converters are covered by the power loss reduction from the conduction loss of power delivery network and the power losses of off-chip converters. The on-chip converter is able to provide higher voltage at the package and reduce the supply current through it. It decreases the conduction loss of package and PCB, from $21.4W$ to $6.0W$. Meanwhile, the power losses of off-chip converters also decrease from $17.6W$ to $11.8W$, due to lighter workload. Hence, different kinds of power losses are evenly distributed in the optimized hybrid architecture. It achieves a flat power efficiency curve, and has the potential for maintaining the power efficiency at heavy workload during the burst mode. The hybrid architecture is also able to adapt to the change of the parasitics of power delivery network. It may provide an solution to deal with the conduction loss of the package by adjusting the intermediate voltage level in some cases, e.g. less pin count or more I/O pins are required. The influence of the resistive parasitic of power delivery network is investigated in Fig. 10. The hybrid architecture adjusts the intermediate voltage from $2.0V$ to $2.3V$ under the change of the parasitic from 80% to 120% . It maintains almost the same performance, e.g. 1.0% difference on average. Compared with the conventional architecture, the hybrid one maintains 1.0% power efficiency improvement on average at $128W$ workload.

The area consumption is also important for next generation processors, e.g. the Haswell processors [5]. Compared with the conventional architecture, the hybrid one alleviates design requirements of off-chip converters by decreasing its

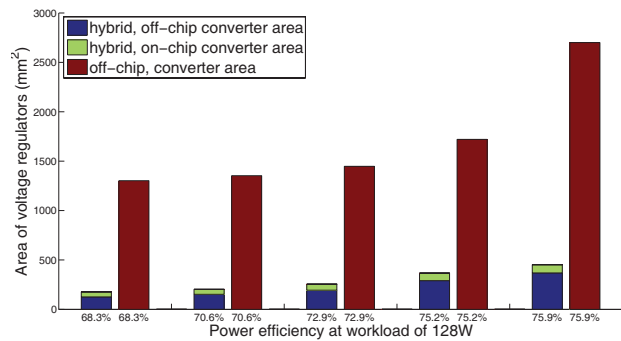


Fig. 11: Design tradeoff between power efficiency and area for different architectures

load. Hence, the area of off-chip converters in the hybrid architecture significantly decreases, which consists with the observation from Haswell [5]. Fig. 11 shows the minimum area of converters in different architectures, with different power efficiency constraints at $128W$ workload. The area of converters in hybrid architecture decreases about 83.2% on average, with the same power efficiencies as the conventional architecture. The area of converters increases sharply with tighter power efficiency requirement, because more effort is needed to make improvement at higher power efficiency level, and it requires larger power bridge, inductor and capacitor.

Switching converters and linear regulators are the most commonly used designs for voltage regulation. Switching converters, e.g. buck converters, are preferred due to the high power efficiency over a wide range of conversion ratios. On the other hand, linear regulators achieve a compact area, but face the power efficiency limitation given by the conversion ratios. Linear regulators are preferable mainly for small conversion ratios. Multiple works are proposed to evaluate the hybrid power delivery system architecture using off-chip converters and on-chip linear regulators [7]. It is assumed that the area of a single linear regulator that provides a specific current and voltage to a load consumes approximately the same area and power consumption as the summation of multiple small linear regulators that provide same total current [7]. The design of the small linear regulators refers to [6]. It achieved a $100mA$ rating linear regulator with a minimum dropout voltage of $0.2V$, consisting of pre-driver, operational amplifier, output stage and on-chip decoupling capacitor. The $0.6nF$ on-chip decoupling capacitor is used to maintain a comparable load transient response, and the total area of the linear regulator is about $0.019mm^2$ using the on-chip capacitance density shown in Table I. The current efficiency of the linear regulator is assumed to be 1. In this case, if supporting a power domain with $1A$ maximum current, 10 of this small linear regulators are needed, and area cost is about $0.19mm^2$.

The model of linear regulators is integrated into our optimization framework, and the design of the architecture using off-chip converters and on-chip linear regulators is optimized at the workload of $128W$ to maximize the power efficiency. To maximize the power efficiency of the power delivery system, all the linear regulators operate at the minimum voltage drop with $1.2V$ input voltage. The power efficiency curves of different optimized systems are shown in Fig. 12. Our

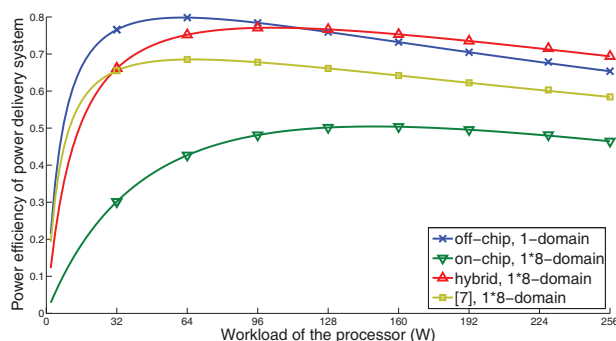


Fig. 12: Power efficiency curve comparison of different power delivery systems optimized at workload of 128W

strategy outperforms the conventional architecture, because of reduced power losses of power delivery network and off-chip converters. The conventional architecture outperforms the architecture using on-chip linear regulators, due to the resistive power loss of linear regulators. Compared to conventional architecture at 128W workload, when employing on-chip linear regulators, the power efficiency of off-chip converters increases from 87.9% to 89.2% due to the increase of output voltage from 1V to 1.2V, while the power efficiency of on-chip linear regulators is around 83.3%. The power efficiencies of three architectures are 76.7%, 75.9% and 66.1% at 128W workload. Without the advantage of the current reduction through the package for on-chip buck converters, the power loss induced by the on-chip linear regulators will degrade the system power efficiency. In order to support the processor with 256W maximum power consumption, the area of the on-chip linear regulators is about $48mm^2$, while our strategy implements the on-chip converters using $94mm^2$. For the off-chip converters, our strategy outperforms the architecture using on-chip linear regulators with the area reduction from to $2228mm^2$ to $813mm^2$. Hence, the hybrid architecture using on-chip linear regulators is desirable to provide fine-grained power management with tight on-chip area constraints. On the other hand, the hybrid architecture using on-chip converters decreases the current through the package to alleviate the conduction loss of power delivery network and the design of off-chip converters. It achieves high power efficiency of power delivery system and small area of off-chip converters.

VI. CONCLUSION

In this paper, we present an analysis and design optimization platform of power delivery system called PowerSoC. It employs an analytical model to fast evaluate important characteristics of power delivery system with comparable accuracy. Based on our model, GP is used in PowerSoC to optimize different system designs, and to explore the tradeoff between the promising characteristics and costs of employing on-chip voltage regulators, especially on-chip buck converters. Compared with the conventional architecture, the hybrid architecture using on-chip converters has potential for efficient dynamic voltage scaling and small area, and it is adaptable with the change of the parasitic of power delivery network within 1.0% power efficiency difference on average. However, careful account of the overhead of on-chip converters is

necessary, and currently it is difficult to maintain comparable power efficiency without the help of off-chip converters.

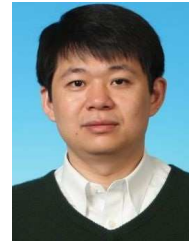
REFERENCES

- [1] J. A. Winter, D. H. Albonesi, and C. A. Shoemaker, "Scalable thread scheduling and global power management for heterogeneous many-core architectures," in *PACT '10*. New York, NY, USA: ACM, 2010.
- [2] W. Kim, M. Gupta, G.-Y. Wei, and D. Brooks, "System level analysis of fast, per-core dvfs using on-chip switching regulators," in *High Performance Computer Architecture, 2008. HPCA 2008. IEEE 14th International Symposium on*, feb. 2008, pp. 123–134.
- [3] J. Sun, M. Xu, F. Lee, and Y. Ying, "High power density voltage divider and its application in two-stage server vr," in *Power Electronics Specialists Conference, 2007. PESC 2007. IEEE, 2007*, pp. 1872–1877.
- [4] H. Esmaeilzadeh, E. Blem, R. St.Amant, K. Sankaralingam, and D. Burger, "Dark silicon and the end of multicore scaling," in *Computer Architecture (ISCA), 2011 38th Annual International Symposium on*, June 2011, pp. 365–376.
- [5] E. Burton, G. Schrom, F. Paillet, J. Douglas, W. Lambert, K. Radhakrishnan, and M. Hill, "Fivr fully integrated voltage regulators on 4th generation intel core socs," in *Applied Power Electronics Conference and Exposition (APEC), 2014 Twenty-Ninth Annual IEEE*, March 2014.
- [6] P. Hazucha, T. Karnik, B. Bloechel, C. Parsons, D. Finan, and S. Borkar, "Area-efficient linear regulator with ultra-fast load regulation," *Solid-State Circuits, IEEE Journal of*, vol. 40, no. 4, april 2005.
- [7] I. Vaisband and E. Friedman, "Heterogeneous methodology for energy efficient distribution of on-chip power supplies," *Power Electronics, IEEE Transactions on*, vol. 28, no. 9, pp. 4267–4280, Sept 2013.
- [8] Y. Ramadass, A. Fayed, and A. Chandrakasan, "A fully-integrated switched-capacitor step-down dc-dc converter with digital capacitance modulation in 45 nm cmos," *Solid-State Circuits, IEEE Journal of*, vol. 45, no. 12, pp. 2557–2565, dec. 2010.
- [9] P. Hazucha, G. Schrom, J. Hahn, B. Bloechel, P. Hack, G. Dermer, S. Narendra, D. Gardner, T. Karnik, V. De, and S. Borkar, "A 233-mhz 80%-87% efficient four-phase dc-dc converter utilizing air-core inductors on package," *Solid-State Circuits, IEEE Journal of*, vol. 40, no. 4, pp. 838–845, april 2005.
- [10] J. Wibben and R. Harjani, "A high-efficiency dc-dc converter using 2 nh integrated inductors," *Solid-State Circuits, IEEE Journal of*, vol. 43, no. 4, pp. 844–854, april 2008.
- [11] J. Lee, G. Hatcher, L. Vandenberghe, and C.-K. Yang, "Evaluation of fully-integrated switching regulators for cmos process technologies," *Very Large Scale Integration (VLSI) Systems, IEEE Transactions on*, vol. 15, no. 9, pp. 1017–1027, 2007.
- [12] P. Wu, S. Tsui, and P. Mok, "Area- and power-efficient monolithic buck converters with pseudo-type iii compensation," *Solid-State Circuits, IEEE Journal of*, vol. 45, no. 8, pp. 1446–1455, 2010.
- [13] X. Wang, J. Xu, Z. Wang, K. Chen, X. Wu, and Z. Wang, "Characterizing power delivery systems with on/off-chip voltage regulators for many-core processors," in *Design, Automation and Test in Europe Conference and Exhibition (DATE), 2014*, March 2014, pp. 1–4.
- [14] "Power system on-chip (PowerSoC). [online]."
<http://www.ust.hk/~ceeu>
- [15] J. Sun, D. Giuliano, S. Devarajan, J.-Q. Lu, T. Chow, and R. Gutmann, "Fully monolithic cellular buck converter design for 3-d power delivery," *Very Large Scale Integration (VLSI) Systems, IEEE Transactions on*, vol. 17, no. 3, pp. 447–451, march 2009.
- [16] N. Sturcken, E. O'Sullivan, N. Wang, P. Herget, B. Webb, L. Romankiw, M. Petracca, R. Davies, R. Fontana, G. Decad, I. Kymissis, A. Peterchev, L. Carloni, W. Gallagher, and K. Shepard, "A 2.5d integrated voltage regulator using coupled-magnetic-core inductors on silicon interposer delivering 10.8a/mm²," in *Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2012 IEEE International*, feb. 2012.
- [17] G. Schrom, P. Hazucha, F. Paillet, D. S. Gardner, S. Moon, and T. Karnik, "Optimal design of monolithic integrated dc-dc converters," in *Integrated Circuit Design and Technology, 2006. ICICDT '06. 2006 IEEE International Conference on*, 2006, pp. 1–3.
- [18] J. Gjanci and M. Chowdhury, "A hybrid scheme for on-chip voltage regulation in system-on-a-chip (soc)," *Very Large Scale Integration (VLSI) Systems, IEEE Transactions on*, vol. 19, no. 11, pp. 1949–1959, Nov 2011.
- [19] G. Yan, Y. Li, Y. Han, X. Li, M. Guo, and X. Liang, "Agileregulator: A hybrid voltage regulator scheme redeeming dark silicon for power efficiency in a multicore architecture," in *High Performance Computer Architecture (HPCA), 2012 IEEE 18th International Symposium on*, Feb 2012, pp. 1–12.

- [20] Z. Zeng, X. Ye, Z. Feng, and P. Li, "Tradeoff analysis and optimization of power delivery networks with on-chip voltage regulation," in *Design Automation Conference (DAC), 2010 47th ACM/IEEE*, June 2010.
- [21] B. Amelifard and M. Pedram, "Optimal design of the power-delivery network for multiple voltage-island system-on-chips," *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, vol. 28, no. 6, pp. 888–900, June 2009.
- [22] R. Miftakhutdinov, "Optimal design of interleaved synchronous buck converter at high slew-rate load current transients," in *Power Electronics Specialists Conference, 2001. PESC. 2001 IEEE 32nd Annual*, vol. 3, 2001, pp. 1714–1718 vol. 3.
- [23] Z. Zhang, W. Eberle, Z. Yang, Y.-F. Liu, and P. Sen, "Optimal design of current source gate driver for a buck voltage regulator based on a new analytical loss model," in *Power Electronics Specialists Conference, 2007. PESC 2007. IEEE, 2007*, pp. 1556–1562.
- [24] Y. Ren, M. Xu, J. Zhou, and F. Lee, "Analytical loss model of power mosfet," *Power Electronics, IEEE Transactions on*, vol. 21, no. 2, pp. 310–319, March 2006.
- [25] J. Sun, J.-Q. Lu, D. Giuliano, T. P. Chow, and R. J. Gutmann, "3d power delivery for microprocessors and high-performance asics," in *Applied Power Electronics Conference, APEC 2007 - Twenty Second Annual IEEE, 25 2007-march 1 2007*, pp. 127–133.
- [26] P. Zumel, C. Fernandez, A. de Castro, and O. Garcia, "Efficiency improvement in multiphase converter by changing dynamically the number of phases," in *Power Electronics Specialists Conference, 2006. PESC '06. 37th IEEE, 2006*, pp. 1–6.
- [27] S. Park, J. Park, D. Shin, Y. Wang, Q. Xie, M. Pedram, and N. Chang, "Accurate modeling of the delay and energy overhead of dynamic voltage and frequency scaling in modern microprocessors," *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, vol. 32, no. 5, pp. 695–708, 2013.
- [28] R. Redl, B. Erisman, and Z. Zansky, "Optimizing the load transient response of the buck converter," in *Applied Power Electronics Conference and Exposition, 1998. APEC '98. Conference Proceedings 1998., Thirteenth Annual*, vol. 1, 1998, pp. 170–176 vol.1.
- [29] A. Soto, A. De Castro, P. Alou, J. Cobos, J. Uceda, and A. Lotfi, "Analysis of the buck converter for scaling the supply voltage of digital circuits," *Power Electronics, IEEE Transactions on*, vol. 22, no. 6, pp. 2432–2443, 2007.
- [30] "Coilcraft. [online]." <http://www.coilcraft.com>.
- [31] Z. Xu, X. Gu, M. Scheuermann, K. Rose, B. Webb, J. Knickerbocker, and J.-Q. Lu, "Modeling of power delivery into 3d chips on silicon interposer," in *Electronic Components and Technology Conference (ECTC), 2012 IEEE 62nd*, May 2012, pp. 683–689.
- [32] M. S. Gupta, J. L. Oatley, R. Joseph, G.-Y. Wei, and D. M. Brooks, "Understanding voltage variations in chip multiprocessors using a distributed power-delivery network," in *Proceedings of the conference on Design, automation and test in Europe*, ser. DATE '07. San Jose, CA, USA: EDA Consortium, 2007, pp. 624–629. [Online]. Available: <http://dl.acm.org/citation.cfm?id=1266366.1266498>
- [33] M. Hershenson, S. Boyd, and T. Lee, "Optimal design of a cmos op-amp via geometric programming," *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, vol. 20, no. 1, 2001.
- [34] D. Grant and J. Gowar, *Power MOSFETS: theory and applications*, ser. Wiley Interscience publication. Wiley, 1989.
- [35] "Mosis integrated circuit fabrication service, use information sciences insitute. [online]." <http://www.mosis.com>.
- [36] "Fairchild semiconductor. [online]." <https://www.fairchildsemi.com>.
- [37] "Bourns. [online]." <http://www.bourns.com>.
- [38] "Murata manufacturing. [online]." <http://www.murata.com>.
- [39] P. Wu and P. Mok, "A monolithic buck converter with near-optimum reference tracking response using adaptive-output-feedback," *Solid-State Circuits, IEEE Journal of*, vol. 42, no. 11, pp. 2441–2450, 2007.
- [40] J. T. DiBene, P. Morrow, C.-M. Park, H. W. Koertzen, P. Zou, F. Thenus, X. Li, S. W. Montgomery, E. Stanford, R. Fite, and P. Fischer, "A 400 amp fully integrated silicon voltage regulator with in-die magnetically coupled embedded inductors," in *IEEE Applied Power Electronics Conference and Exposition, Special Presentation*, Feb 2010.
- [41] A. Sinker, H. Wang, and N. S. Kim, "Workload-aware voltage regulator optimization for power efficient multi-core processors," in *Design, Automation Test in Europe Conference Exhibition (DATE), 2012*, March 2012, pp. 1134–1137.
- [42] R. Kumar, K. Farkas, N. Jouppi, P. Ranganathan, and D. Tullsen, "Single-isa heterogeneous multi-core architectures: the potential for processor power reduction," in *Microarchitecture, 2003. MICRO-36. Proceedings. 36th Annual IEEE/ACM International Symposium on*, Dec 2003, pp. 81–92.



Xuan Wang (S'12) received the B.S. degree in electrical engineering from Shanghai Jiaotong University, Shanghai, China in 2009. From 2009, he has been a Ph.D. student in the department of Electronic and Computer Engineering in Hong Kong University of Science and Technology (HKUST). His research interests include embedded system, multiprocessor system, network-on-chip and fault tolerant design and reliability issues in very deep submicron technologies.



Jiang Xu (S'02-M'07) received his Ph.D. degree from Princeton University in 2007. From 2001 to 2002, he worked at Bell Labs, NJ, as a Research Associate. He was a Research Associate at NEC Laboratories America, NJ, from 2003 to 2005. He joined a startup company, Sandbridge Technologies, NY, from 2005 to 2007 and developed and implemented two generations of NoC-based ultra-low power multiprocessor systems-on-chip for mobile platforms. Dr. Xu is an Associate Professor at Hong Kong University of Science and Technology. He is

the founding director of Xilinx-HKUST Joint Lab and establishes Mobile Computing System Lab. He currently serves as the Area Editor of NoC, SoC, and GPU for ACM Transactions on Embedded Computing Systems and Associate Editor for IEEE Transaction on Very Large Scale Integration Systems. He is an ACM Distinguished Speaker and IEEE Distinguished Lecturer. He served on the steering committees, organizing committees and technical program committees of many international conferences. Dr. Xu authored and coauthored more than 80 book chapters and papers in peer-reviewed journals and international conferences. His research areas include network-on-chip, multiprocessor system-on-chip, optical interconnects, embedded system, computer architecture, low-power VLSI design, and HW/SW codesign.

Zhe Wang (S'14) received his B.S. degree in Electronic Engineering from Shanghai Jiao Tong University, Shanghai, China, in 2011. He is currently a Ph.D. candidate in Electronic and Computer Engineering at Hong Kong University of Science and Technology, Hong Kong. His research interests include embedded systems, network-on-chip and design space exploration techniques.

Kevin J. Chen (M'96-SM'06-F'14) received the B.S. degree from Peking University, Beijing, China, and the Ph.D. degree from the University of Maryland, College Park, MD, USA. He is currently a Professor with the Department of Electronic and Computer Engineering, The Hong Kong University of Science and Technology, Hong Kong. Prof. Chen is an Editor for IEEE TRANSACTIONS ON ELECTRON DEVICES.

Xiaowen Wu (S'12) received BSc degree in computer science from the Harbin Institute of Technology, China in 2008. He is currently working towards the Ph.D. degree in Electronic and Computer Engineering at the Hong Kong University of Science and Technology. His research interests include embedded systems, multiprocessor systems, and network-on-chip.

Zhehui Wang (S'11) received B.S. degree in electrical engineering from Fudan University, China in 2010. He is currently working towards the Ph.D. degree at the department of Electronic and Computer Engineering in Hong Kong University of Science and Technology. His research interests include embedded system, multiprocessor systems, network-on-chip, and floorplan design for network-on-chip.

Peng Yang received the B.S. degree in Electronic of Science and Technology from Wuhan University, Hubei, China, in 2013. He is now a Ph.D. student in Electronic and Computer Engineering at The Hong Kong University of Science and Technology, Hong Kong. His research interests include network-on-chip, multiprocessor system-on-chip, and embedded system.

Luan H.K. Duong (S'14) received the B.S. degree in Computer Science from The Hong Kong University of Science and Technology, Hong Kong, in 2012. He is currently a Ph.D. candidate in Electronic and Computer Engineering at the Hong Kong University of Science and Technology, Hong Kong. His research interests include embedded system, system-on-chip and network-on-chip.