Adaptive Power Delivery System Management for Many-Core Processors with On/Off-Chip Voltage Regulators

Haoran Li, Jiang Xu, Zhe Wang, Peng Yang, Rafael K. V. Maeda, Zhongyuan Tian
Department of Electronic and Computer Engineering, The Hong Kong University of Science and Technology
Email: {hr.li, jiang.xu}@ust.hk

***Abstract*—The power delivery system (PDS) plays a crucial role of guaranteeing the proper functionality of many-core processors. However, as PDS is usually optimized to provide power to the target chip at its best performance level, its energy efficiency can be seriously degraded under highly dynamic workloads, making it a major source of system power losses. On-chip voltage regulators (VR), which are able to achieve fast and fine-grained power control, have been popular choices for PDS implementation and provided design opportunities for improving system energy efficiency. In this paper, we propose the adaptive Quantized Power Management (QPM) scheme to dynamically adjust the PDS with both on-chip and off-chip VRs based on run-time workloads. Experimental results on different applications show that QPM applied on a hybrid PDS with both on/off-chip voltage regulators (VR) achieves 74.1% average overall energy efficiency, 12.3% higher than the conventional PDS with single off-chip VR.**

I. Introduction

Parallel computing capability of multi-core or many-core processors has been well exploited to improve the system performance. However, as the constraint of power consumption and dissipation becomes increasingly critical in the "Dark Silicon" era[1], the maximum performance is seriously throttled by the system energy efficiency.

Power delivery system (PDS), which is responsible of supplying sufficient and reliable power from external power source to all functional units, plays a crucial role of maintaining system functionality. PDS usually includes DC-DC voltage regulators (VR), power delivery networks (PDN) and power management logic units. VRs are critical elements of PDS, converting higher or lower input voltage to the required voltage levels. Compared to simple low-dropout regulators (LDO) and switched-capacitor regulators, buck converters are able to maintain high conversion efficiency in a large range of output load as well as voltage conversion ratio[2]. In conventional solutions of PDS, off-chip VRs are equipped with bulky inductors and capacitors and put on the Printed Circuit Board (PCB). In recent years, on-chip VRs, which composed of integrated power transistors, control circuits and passive components, have been preferred by designers.They have the advantages of: (1) fast and fine-grained voltage control with MHz-level frequency; (2) small area overhead of realizing multi-rail on-chip power delivery paths and multiple voltage domains; and (3) alleviating the I^2R power loss and the need of package power pins by lowering through-package load current.

This work is partially supported by GRF16235616, IRS17EG48, and FSGRF15EG04.

The unpleasant PDS energy efficiency degradation has become a serious design issue of the whole system, and the intrinsic VR conversion losses and PDN delivery losses take large proportion of system total power consumption. To make things worse, power/thermal constraints limit the maximum performance of many-core systems, and power management techniques such as Dynamic Voltage Frequency Scaling (DVFS), Clock Gating (CG) or Power Gating (PG) have been extensively applied to improve system energy efficiency, resulting in highly dynamic workloads in processors. Unfortunately, PDS is usually designed and optimized statically at its best performance level, while significant PDS efficiency degradation will be encountered in light/heavy-workload scenarios. Traditionally, VRs at the same stage will supply power evenly to the next stage or target chip regardless workload change, while the implementation of multiple on/off-chip VRs provides chances of applying novel PDS management schemes. Research has shown the potential of dynamically adjusting VRs or the whole PDS to improve system energy efficiency[3][4][5][6]. Most of the previous works mainly focus on circuit-level optimization or dynamic management of LDO networks, while wide design space is observed with system-level management of PDS constructed by buck converters. In this paper, we propose an adaptive power management scheme to dynamically manage the PDS with both on/off-chip VRs based on the run-time workloads, which can effectively improve the PDS overall energy efficiency for many-core processors.

II. Modeling and Static Optimization of Power Delivery Systems

In this paper, buck converters are adopted for DC-DC voltage regulation. Fig.1 gives a schematic view of a multi-phase buck converter with Pulse-Width-Modulation (PWM) feedback control. It is usually composed of MOS power transistors, inductors, output capacitors and feedback control circuits. There are five main sources of power losses in a buck converter: the switching loss of the power transistors and

978-3-9815370-8-6/17/$31.00

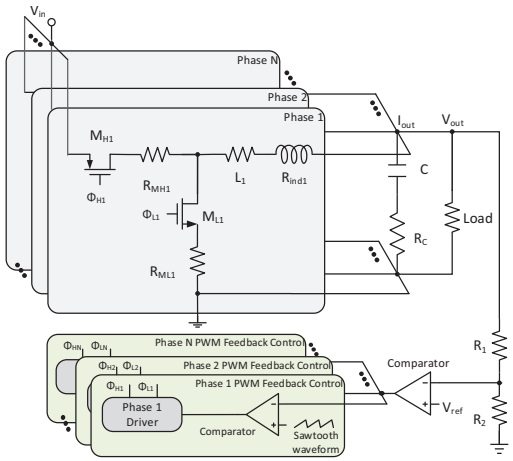


Fig. 1. Schematic diagram of a multi-phase buck converter with PWM feedback control

driver circuits P_{sw}, the on-state resistive loss of the transistors P_{Ron}, the inductor resistive loss P_{ind}, the power consumed by the feedback control circuits P_{ctrl}, and the static power P_{stat} as well as the transistors short-circuit loss P_{sc}. Thus the energy efficiency η of a buck converter can be described as the following equation, where N is the number of phases:

$$\eta = \frac{P_{out}}{P_{out} + N(P_{sw} + P_{Ron} + P_{Rind} + P_{ctrl} + P_{stat} + P_{sc})}. \quad (1)$$

Some of the power losses for one phase are modeled as follows:

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

$$P_{Ron} = (DR_{on,H} + (1-D)R_{on,L} + R_{ind})(I_{ind}^2 + \frac{\Delta I_{ind}^2}{12}), \quad (3)$$

$$P_{Rind} = R_{ind}(I_{ind}^2 + \frac{\Delta I_{ind}^2}{12}). \quad (4)$$

$$P_{ctrl} = I_{ctrl}V_{driver}. \quad (5)$$

C_{MOS} and C_{driver} are the effective capacitance of power transistors and their driver circuits. V_{driver} is the supply voltage of drivers and control circuits. f_{sw} is the switching frequency. $R_{on,H}$ and $R_{on,L}$ are the on-resistance for the high/low-side power transistors. R_{ind} is the resistance of inductors. I_{ind} and ΔI_{ind} are, respectively, the average and peak-to-peak value of inductor current in each phase. D is the duty cycle of the gate signal.

Voltage scaling $T_{scaling}$ is an important performance indicator in this respect, which can be derived as

$$T_{scaling} = \beta\sqrt{\frac{2L_{ind}(C_{out} + C_{Load})\Delta D}{D_{avg}(1 - D_{avg})(1 + 0.5\Delta D)}}. \quad (6)$$

D_{avg} is the average duty cycle of the initial and final voltage regulation states, and ΔD is the difference between both duty cycles. β is the fitting empirical factor.

In order to capture accurate characteristics of the power delivery system, PDN parasitics such as pins and bumps resistance should be modeled in detail, as they would degrade system efficiency and performance, especially at a high load

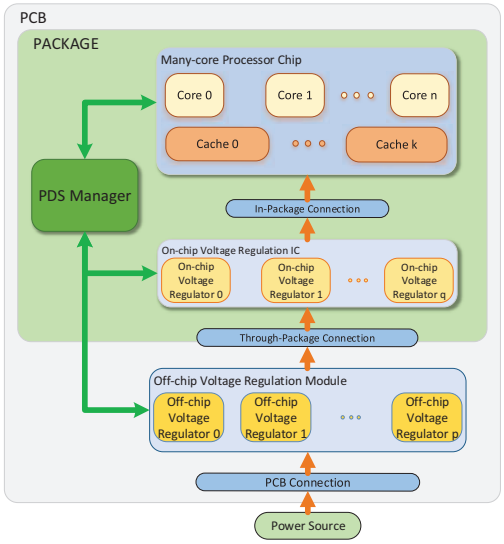


Fig. 2. Control diagram of the hybrid power delivery system

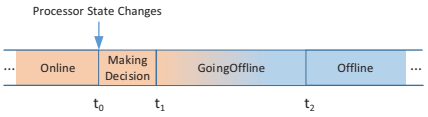
current. The PDN parasitics includes the PCB wires connecting power source and off-chip VRs $R_{pcb,src}$ and $R_{pcb,off}$, package pins/wires/vias R_{pkg}, C4-bumps between on/off-chip VRs and processors $R_{bump,off}$ and $R_{bump,on}$, on-chip processor power grid R_{grid}, etc. For example, the power loss along the PDN for a two-stage hybrid PDS can be modeled as

$$P_{PDN} = 2R_{pcb,src}I_{pcb}^2 + 2(R_{pcb,off} + R_{pkg} + R_{bump,off})I_{pkg}^2 + 2(R_{bump,on} + R_{grid})I_{proc}^2. \quad (7)$$

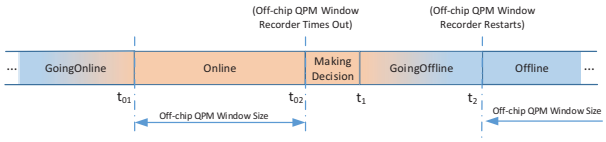
In this paper, we utilize PowerSoC[7], a modeling and analysis platform for the entire PDS, to find optimal design parameters such as MOS power transistors channel width, switching frequency and passive components. Geometric Programming(GP)-based convex optimization is applied to reach maximum PDS overall efficiency under given constraints, and different PDS paradigms such as conventional PDS with only off-chip VRs and hybrid PDS with both on-chip and off-chip VRs are constructed and statically optimized at the assumed average workloads.

III. Quantized Power Management Scheme

For many-core processors, more than half of the cores are predicted to be in low-power modes or even powered off due to power and thermal limitations[1], resulting in highly dynamic run-time chip workloads, and significant energy efficiency degradation occurs as the PDS is usually designed and optimized to work at the average workload and suffer from low conversion efficiency at light load. Traditionally, bulky off-chip VRs are placed on PCB to provide power to the target processor. The introduction of multiple on-chip VRs makes it possible to establish a hybrid PDS with both on/off-chip VR stages, which achieves better trade-off between energy efficiency and fast/fine-grained power management[7]. Conventionally, homogeneous VRs at the same stage provide



(a) Timing process of on-chip QPM going offline



(b) Timing process of off-chip QPM going offline with Moving-Window strategy

Fig. 3. Timing process of on/off-chip QPM going offline

the output load evenly to the next stage or the target functional units, which in this paper is labeled as Load-Balance Power Management (LBPM). Unfortunately, although LBPM performs well when supporting workloads around the optimal point, PDS suffers much lower energy efficiency at light load with all VRs actively working (called "Online" in this paper, and similarly "offline" for shutdown VRs). The primary reason is that P_{sw} and P_{ctrl} are unavoidable and dominant at light workload[8]. Intuitively, reducing the number of online VRs is promising to improve PDS efficiency. In this paper, we propose the Quantized Power Management (QPM) scheme to dynamically manage online VRs according to run-time processor workloads such that the remaining online VRs are working near their optimal load points.

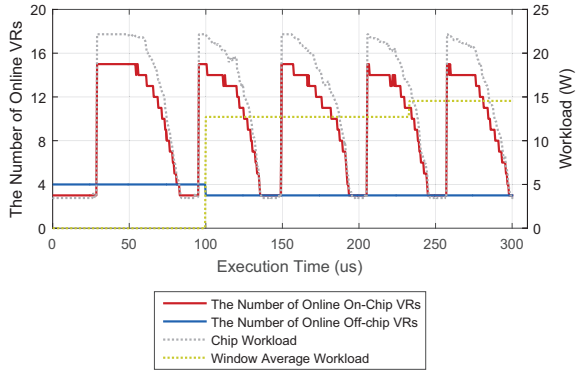
As shown in Fig.2, a PDS manager is responsible of receiving information from the processor chip and managing both stages of on/off-chip VRs. In this paper, VRs at the same stage are assumed identical, and the number of optimal online VR in the same stage can be simply estimated based on:

$$N_{opt\_online\_VR} = \begin{cases} \lceil \frac{P_{cur}}{P_{opt}} \cdot N_{VR} \rceil & (if\ P_{cur} \leq P_{opt}), \\ N_{VR} & (if\ P_{cur} > P_{opt}) \end{cases} \quad (8)$$

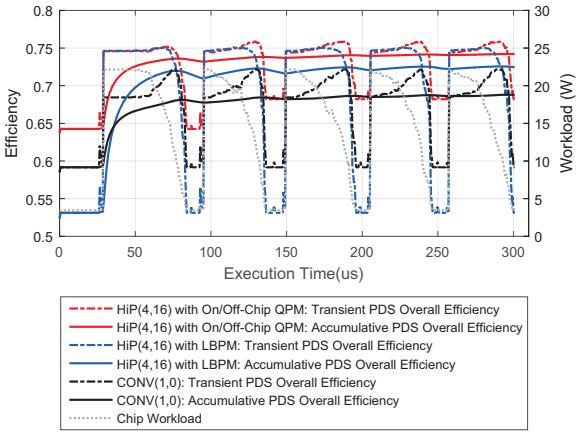
where P_{cur} is the current chip workload, and P_{opt} is the average workload that PDS is designed and optimized at. N_{VR} is the number of VRs at each stage.

Fig.3(a) illustrates the detailed timing process of an on-chip VR going offline. When the processor states change at t_0, for example several cores go idle, power-gated or clock-gated, the chip workload varies accordingly, invoking the PDS manager to manage online VRs. If a going-offline command is sent to this VR, the VR will be disconnected from the PDS and go offline at t_1. However, the going-offline procedure should be gradual, including powering off the control/driver circuits and discharge the output capacitor etc. to avoid spike and possible damage. For the going-online procedure, soft-start is also required to protect the circuit from overshoot voltage. With much smaller passive components and much higher frequency, on-chip VRs can achieve much shorter going-online/offline time at the level of dozens of ns to hundreds of ns[9]. At t_2 the on-chip VR is fully offline and consumes no extra power.

QPM can be applied to off-chip VRs to further improve system efficiency. However, as off-chip VRs suffer much



(a) Snapshot of online VR transition with on/off-chip QPM scheme



(b) Snapshot of power delivery system overall efficiency with different power management schemes

Fig. 4. Transient evaluation of on/off-chip QPM, running RSd-32x28x8 for 5 iterations

slower response time due to their kHz-level frequencies and bulky passive components, a myopic management decision based on a short term of workload deviation may lead to a long term of efficiency deterioration before another readjustment. Therefore, a moving-window strategy for off-chip QPM is adopted. A recorder will sample the average workload within a timing window, and the off-chip QPM decision will be made when the window times out. Fig.3(b) shows a off-chip VR going offline with moving-window QPM. When the VR is back online from the last transition procedure at t_{01}, the workload recorder begins to sample average workload within the window size. At t_{02} the workload recorder times out and invoke PDS manager making decision based on the average workload. If the VR is scheduled to go offline, at t_1 the off-chip VR will start going-offline transition until t_2, and the window will be reset and restarted. The going-online timing processes for both on-chip and off-chip QPM are similar.

IV. Experimental Analysis and Evaluation

A. Simulation Setup

The key modeling parameters for PDS are estimated as follows. The off-chip VRs are implemented with Fairchild power trench MOSFET[10]. The capacitors and inductor

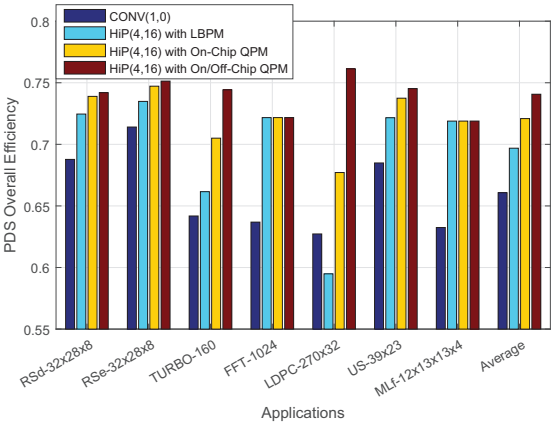


Fig. 5. Power Delivery system overall efficiency comparison with different power management schemes and applications for 5 iterations

parameters are from [11][12]. For the on-chip VRs, they are implemented with $130nm$ CMOS power transistors[13] on a separated die in the same package with processors. Their air-core inductors and integrated capacitors are estimated from [14]. The parameters of PDN parasitics, including PCB wires, package wires/vias/pins, C4-bumps etc., are adopted from [15][16]. Based on the above parameters, we design and optimize different PDS models using PowerSoC[7]. The conventional PDS with only one off-chip VR is labeled as *CONV(1,0)*, and the hybrid PDS is constructed with 4 off-chip VRs and 16 in-package on-chip VRs for evaluation, which is labeled as *HiP(4,16)*. The input voltage for off-chip and on-chip VRs is 12V and 1.2V. The going-online/offline period for the optimized on-chip and off-chip VRs are estimated as $68ns$ and $33\mu s$ respectively based on Eq(6). The off-chip QPM window size is set to be $100\mu s$. The simulation is based on JADE[17], a heterogeneous multiprocessor system platform for fast initial architecture explorations. Realistic applications are adopted for evaluation. The target processor models are 256 ARMv7-A cores at 7nm FinFET technology with mesh 16x16 network-on-chip. The core clock rate is 1GHz and the chip supply voltage is 0.45V. The average chip workload is assumed as 25W.

B. Transition Analysis and Energy Efficiency Evaluation

Fig.4(a) gives a demonstration of on/off-chip QPM running RSd-32x28x8 for 5 iterations. The number of online on-chip VRs is tightly adapted to the chip workloads changes, while the off-chip QPM decisions are made based on the sampled average workload within the window. At 100us one off-chip VR goes offline based on the average workload, and the number of online off-chip VRs is reduced from 4 to 3.

In Fig.4(b), the corresponding efficiency snapshots of different PDS and power management schemes are also presented. With fewer online VRs, *HiP(4,16)* with on/off-chip QPM can effectively improve the PDS overall efficiency at light workloads compared to LBPM and *CONV*. And benefited from reduced the through-package current, *HiP(4,16)* with both LBPM and QPM can significantly reduce PDS power loss at heavy workloads compared to *CONV(1,0)*.

In Fig.5, we evaluate the performance of the proposed on/off-chip QPM on *HiP(4,16)* over several applications, all of which are simulated for 5 iterations. For some applications with long terms of workload deviation such as LDPC and TURBO, *HiP(4,16)* with on/off-chip QPM can achieve significant improvement over *CONV(1,0)* and LBPM. With the help of adaptive on/off-chip QPM based on run-time workload, *HiP(4,16)* achieves 25.0% PDS power loss reduction compared to *CONV(1,0)*, and the average PDS overall efficiency is improved from 66.1% to 74.1%, which is 12.3% and 6.8% higher than *CONV(1,0)* and LBPM, respectively.

V. Conclusion

In this paper, we propose the adaptive Quantized Power Management(QPM) scheme that can dynamically manage the on/off-chip voltage regulators(VR) in hybrid power delivery systems(PDS) based on the run-time workload. Experimental results on different applications show its potential of improving system light-workload energy efficiency for large-scale many-core processors. When applied on a hybrid PDS with both on/off-chip voltage regulators(VR), QPM achieves 74.1% average overall efficiency, which is 12.3% higher than the conventional PDS with single off-chip VR.

References

[1] H. Esmaeilzadeh *et al.*, "Dark silicon and the end of multicore scaling," in *ISCA'11*. IEEE, 2011, pp. 365–376.
[2] G. Patounakis *et al.*, "A fully integrated on-chip dc-dc conversion and power management system," *IEEE Journal of Solid-State Circuits*, vol. 39, no. 3, pp. 443–451, 2004.
[3] P. Zumel *et al.*, "Efficiency improvement in multiphase converter by changing dynamically the number of phases," in *Power Electronics Specialists Conference*, 2006, pp. 1–6.
[4] G. Yan *et al.*, "Agileregulator: A hybrid voltage regulator scheme redeeming dark silicon for power efficiency in a multicore architecture," in *IEEE HPCA'12*. IEEE, 2012, pp. 1–12.
[5] W. Lee, Y. Wang, D. Shin, N. Chang, and M. Pedram, "Optimizing the power delivery network in a smartphone platform," *IEEE TCAD*, 2014.
[6] I. Vaisband and E. G. Friedman, "Energy efficient adaptive clustering of on-chip power delivery systems," *INTEGRATION, the VLSI journal*, 2015.
[7] X. Wang *et al.*, "An analytical study of power delivery systems for many-core processors using on-chip and off-chip voltage regulators," *IEEE TCAD*, Sept 2015.
[8] H. Li *et al.*, "Energy-efficient power delivery system paradigms for many-core processors," *IEEE TCAD*, vol. PP, no. 99, pp. 1–1, 2016.
[9] E. Burton *et al.*, "FIVR - Fully integrated voltage regulators on 4th generation Intel Core SoCs," *IEEE APEC'14*, pp. 432–439, 2014.
[10] Fairchild semiconductor. [Online]. Available: https:// www.fairchildsemi.com
[11] Bourns. [Online]. Available: https:// www.bourns.com
[12] Murata. [Online]. Available: https:// www.Murata.com
[13] Mosis integrated circuit fabrication service, usc information sciences institute. [Online]. Available: https:// www.mosis.com
[14] P. Hazucha *et al.*, "A 233-MHz 80%-87% efficient four-phase DC-DC converter utilizing air-core inductors on package," *IEEE Journal of Solid-State Circuits*, vol. 40, no. 4, pp. 838–844, 2005.
[15] M. S. Gupta *et al.*, "Understanding voltage variations in chip multiprocessors using a distributed power-delivery network," *DATE'07*, vol. 1, pp. 624–629, 2007.
[16] R. Zhang *et al.*, "Architecture implications of pads as a scarce resource," *ISCA'14*, pp. 373–384, 2014.
[17] R. K. V. Maeda *et al.*, "JADE : a Heterogeneous Multiprocessor System Simulation Platform Using Recorded and Statistical Application Models." HiPEAC Workshop on Advanced Interconnect Solutions and Technologies for Emerging Computing Systems, 2016.