

# Load Balanced On-Chip Power Delivery for Average Current Demand

Divya Pathak  
Drexel University  
divya.pathak@drexel.edu

Mohammad Hossein  
Hajkazemi  
George Mason University  
mhajkaze@gmu.edu

Mohammad Khavari  
Tavana  
George Mason University  
mkhavari@gmu.edu

Houman Homayoun  
George Mason University  
hhomayoun@gmu.edu

Ioannis Savidis  
Drexel University  
isavidis@coe.drexel.edu

## ABSTRACT

A dynamic power management system for homogeneous chip multi-processors (CMP) is proposed. Each core of the CMP includes on chip DC-DC switching buck converters that are interconnected through a switch network. The peak current rating of the buck converter is selected to meet only the average current demand of the load circuit. A real-time load balancing algorithm is developed which reconfigures the power delivery network by combining the output of multiple buck converters when the workload demand exceeds the peak current rating. Simulation results for the proposed power delivery method indicate up to a 44% reduction in the energy consumption of the CMP system. In addition, the on-chip footprint of the power delivery network, including the on-chip voltage regulators and the switching network, is reduced by at least 23%.

## 1. INTRODUCTION

With the paradigm shift in computing systems from performance oriented design to energy efficiency, considerable research effort has focused on optimizing the core configuration by reducing the over-provisioning of the core resources. Little attention, however, is given to the reduction in the over-provisioning of the circuits delivering power to the cores. Conventionally, the voltage regulator and power conditioning circuits are off-chip. The power consumption and the footprint of the voltage regulators and the conditioning circuits is therefore not a concern while optimizing the power delivery to the core(s). The introduction of chip multi-processors (CMPs) resulted in new challenges in the delivery of power to the multiple cores. Providing low latency, per-core dynamic voltage and frequency scaling (DVFS) is challenging with off-chip voltage regulators [1]. The power supply voltage regulation is also reduced due to longer on-chip interconnects connecting the off-chip VR to the multiple load circuits. On chip voltage regulators

(OCVRs) have been extensively researched and successfully introduced in commercial multi-core systems in Intel 4<sup>th</sup> generation processors [2] as well as IBM POWER8 servers [3]. The choice of OCVR topology is dependant on several factors including system level parameters such as the optimal power conversion efficiency and maximum load current consumption as well as the physical design of the passive components. System level tools such as a power virus [4] or McPAT [5] are used to determine a first order estimate of the peak power consumption of the cores, which is typically overestimated. As a result, the OCVR and the power delivery network are over-provisioned to support a peak load current larger than what is consumed by the cores.

In this paper, an interconnected on-chip power distribution network is modeled. Rather than a static configuration designed for the worst case power consumption of the cores, a work load aware reconfigurable power delivery system is developed. A detailed statistical analysis of the cycle accurate power consumption profile of workloads executed on a CMP system is performed. Each OCVR is designed to support a peak current rating equal to the average load current  $I_{avg}$  consumed across all workloads. SPICE simulations indicate that by reducing the peak current rating of the OCVRs to support  $I_{avg}$ , the energy efficiency of the CMP improves and the on-chip area occupied by the OCVRs is reduced. A load balancing algorithm is developed for dynamic power management. The algorithm is executed on the on-chip power management unit (PMU), and is capable of reconfiguring the power delivery network (PDN) to combine the outputs of multiple OCVRs to support load currents in excess of  $I_{avg}$ .

The rest of the paper is organized as follows: Prior work exploring reconfigurable power delivery networks (RPDN) are discussed in Section 2. The system level simulation and the power consumption profile of multi-application workloads are described in Section 3. The proposed power delivery methodology is discussed in Section 4. Simulated results indicating an improvement in the energy efficiency of the CMP system are also included in Section 4. Concluding remarks are provided in Section 5.

## 2. RELATED WORK

Recent work has attempted to improve the energy efficiency of multi-core and many-core systems by reconfiguring the power delivery network according to the power demand of the work load. An RPDN using switched capacitor volt-

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

GLSVLSI '16, May 18-20, 2016, Boston, MA, USA

© 2016 ACM. ISBN 978-1-4503-4274-2/16/05...\$15.00

DOI: <http://dx.doi.org/10.1145/2902961.2903030>

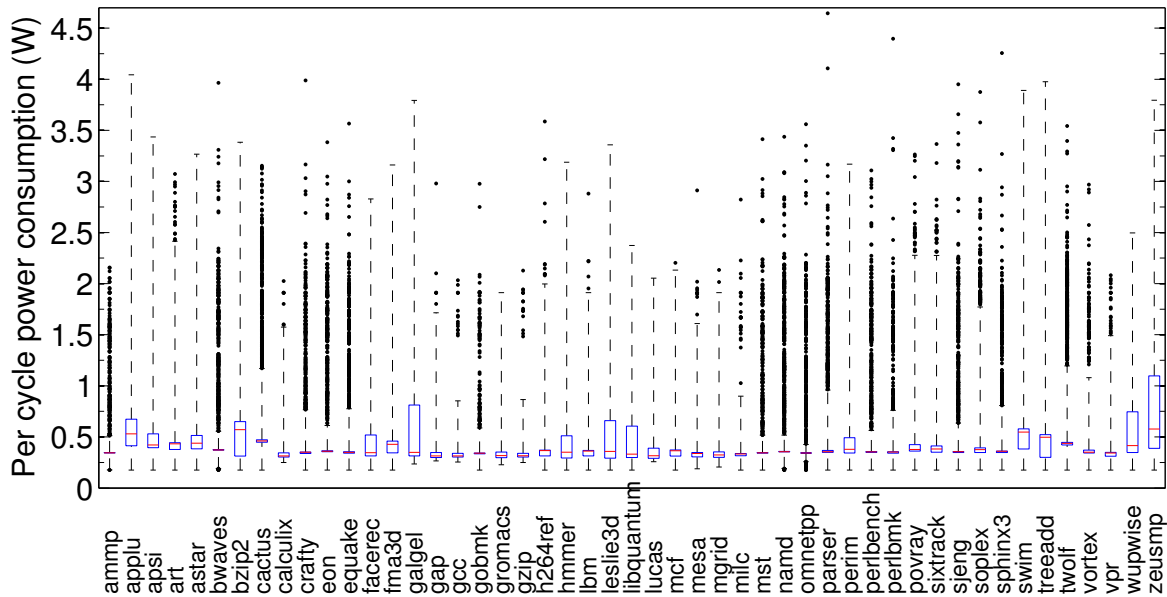


Figure 1: Statistical analysis of the per cycle power consumption of the SPEC CPU 2000 and SPEC CPU 2006 benchmarks.

Table 1: Architectural parameters of the core.

Parameter	Value
Core clock frequency	2.4 GHz
Power supply voltage ( $V_{dd}$ )	1 V
Issue, Commit width	2
INT and FP Instruction Queue	16 entries
Load and Store Queue	16 entries
INT and FP Physical Register File	48 entries
ROB size	48
L1 cache	32 KB, 4-way
L2 cache	256 KB

age regulators (SCVRs) and cross bar switches to serve 8 cores is proposed in [6]. The RPDN consists of 32 cells, where each cell is an SCVR capable of supporting two voltage step down conversions (2:1 and 3:2). The simulation results indicate that the reconfigurable power delivery network offers 40% energy savings as compared to a configuration with per core voltage regulators. The switched capacitor voltage regulators offer 80% power conversion efficiency. The work does not address the inferior voltage regulation offered by the SCVRs.

A run time reconfigurable voltage regulator network of buck converters is described in [7]. The lowest energy consumption across various DVFS levels is determined by solving an integer linear programming (ILP) problem. The timing penalty to set the switching network is not quantified and the ILP is solved for discrete DVFS timing penalties ranging from 5% to 15%. An off-chip buck converter (LTC3816) SPICE model is used instead of an OCVR model, although the OCVR offers an order of magnitude faster voltage response time under DVFS [1].

An RPDN for 3-D many-core systems is proposed in [8]. Single input multiple output buck converters supply power to the many-core system. The energy optimization problem across DVFS operating points (voltage and frequency pairs) is solved through ILP formulation.

Clustering of voltage regulators to boost the energy efficiency of the system is proposed in [9], but the work ignores

the variation in the power conversion efficiency (PCE) of the voltage regulators due to dynamic voltage and frequency scaling. As shown in [10], ignoring the PCE variation of the voltage regulators leads to suboptimal workload mapping and therefore a large penalty on the energy savings possible with DVFS.

Recent work on RPDNs does not offer an analysis of the deterioration in the response time of the power delivery system due to the search and decision time needed to flip the requisite number of switches and reconfigure the connections between the cores [6, 7, 8]. In addition, prior work consider over-provisioned voltage regulators designed for worst case power consumption. The RPDN proposed in this paper is novel as the power network is designed to supply the average power demand of the cores and is adaptable to support peak power demands. The RPDN configuration is managed dynamically without the overhead of solving an off-line or on-line linear or nonlinear programming optimization problem.

### 3. POWER DISSIPATION BEHAVIOR OF APPLICATIONS ON A CMP SYSTEM

The cycle dependant power dissipation of the workloads executing on a CMP system provide insight to optimize the design of the PDN. Workloads which are computation intensive consume higher power during CPU bound phases of the application, but the maximum power consumed is considerably lower than the peak power consumption of a well developed power virus. A detailed power trace analysis of the SPEC2000 and SPEC2006 benchmark suite is performed to obtain realistic power consumption statistics of the workloads. A 16-core CMP in 45 nm technology is modeled using a processor architectural simulator [11]. McPAT [5] is integrated in the simulator to analyze the power consumption of the core. Each core has a 2-way issue and out-of-order execution unit. The micro-architectural parameters of the core used in the simulations are summarized in Table 1.

Each of the 49 benchmarks from the SPEC benchmark suite are simulated at four different timing intervals to cover multiple execution phases. The simulations are run for 10K cycles per time interval. The dynamic and static power consumption is sampled cycle by cycle through McPAT. The statistical dispersion of the per cycle power consumption of different SPEC benchmarks with single phase forwarding is shown through a box plot in Fig. 1. The interquartile range for all the studied benchmarks falls approximately an order of magnitude below the peak power ( $P_{peak}$ ) consumption of 5.73 W reported through McPAT simulations. The number of outliers beyond  $5\sigma$  coverage for each benchmark is an insignificant fraction of the sample size. The combined power dissipation characteristics of all the 49 benchmarks are as follows: The minimum, average, and maximum power consumption is, respectively, 0.175 W, 0.555 W, and 4.755 W. The power dissipation of the applications is between 0.3 W and 0.5 W for approximately 65% of the execution time. The minimum, average, and maximum power dissipation variation per clock cycle is, respectively, 0 W, 0.195 W, and 4.333 W. The power variation is less than 0.1 W for 90% of the time. The studied benchmarks spend 78% of the run-time consuming less than the average power. The maximum power consumption of 4.75 W across all benchmarks is consumed for a very small percentage of the run-time ( $7.5 \times 10^{-5}\%$ ).

## 4. PROPOSED DESIGN OF THE POWER DELIVERY NETWORK

Significant work has been done to optimize the core configuration, work load mapping, and dynamic/static clustering of the cores in a CMP system, but the energy and area loss incurred due to the integration of over-provisioned VRs has been overlooked. In the proposed power delivery system, the loadline of the OCVR serving each core or cluster of cores is designed for the average current ( $I_{avg}$ ) consumed. Designing the OCVR loadline around  $I_{avg}$  reduces the peak voltage demand on the OCVR and significantly decreases the maximum supported load current. The block representation of the power delivery network is shown in Fig. 2. For a CMP system consisting of  $N$  cores or core clusters,  $N$  OCVRs provide the regulated power. In the proposed PDN, if the OCVR serves a cluster of  $n$  identical cores, the power rating of the OCVR is  $n \cdot I_{avg}$ . The output of each OCVR is connected to the inputs of a high-speed switching (HSS) fabric. The  $N$  outputs of the HSS fabric are connected to the local PDN grid of the  $N$  cores or core clusters. The HSS fabric is controlled by the power management unit (PMU). The interconnected power delivery network shown in Fig. 2 provides increased service reliability as compared to conventional on-chip power distribution with a single OCVR serving a single core. In addition, the interconnected network provides opportunity to balance load currents through reconfiguration of the switches.

### 4.1 Load balancing through run-time OCVR clustering

A technique to deliver currents higher than  $I_{avg}$  is described in this section. The current sensors placed in each core are constantly monitored by the PMU. When the sum of the currents sensed from all cores within a cluster ( $I_{sense}$ ) reaches a threshold ( $\Delta I$ ) below  $I_{avg}$ , the PMU configures

**Algorithm 1** Load balanced power delivery with run-time OCVR clustering to support higher than average load current consumption.

---

**Inputs:**  
Current consumption sensed from the core:  $I_{sense,x}$   
Current threshold:  $\Delta I$   
Voltage level applied to the core:  $V_x \in [V_{dd,1}, V_{dd,2}, \dots, V_{dd,m}]$  where  $x \in [1, \dots, n]$ ,  $m = \text{number of DVFS levels}$ ,  $n = \text{number of OCVRs/Cores}$   
Switch matrix:  $Switch_{n,n-1}$   
**Constraints:**  
 $t_{switch} + t_{PMU} < t_{core}$   
 $\sum_{x=1}^n V_x \cdot I_{sense,x} < n \cdot V_{dd,m} \cdot I_{avg}$

- 1: Append array *CORE\_RED* with core id  $x$  where  $I_{sense,x} \geq I_{avg} - \Delta I$
- 2: Append array *CORE\_GREEN* with core id  $y$  where  $I_{sense,y} < I_{avg} - \Delta I$
- 3: **if**  $\text{length}(\text{CORE\_RED}) > 0$  **then**
- 4:   call *OCVR\_CLUSTER*  $\triangleright$  Reconfigure the PDN by clustering the output of OCVRs
- 5: **else**
- 6:   call *OCVR\_DECLUSTER*  $\triangleright$  Reconfigure the PDN by de-clustering the output of OCVRs
- 7: **end if**
- 8: **procedure** *OCVR\_CLUSTER*
- 9:   **for** each  $i$  in  $\min(\text{length}(\text{CORE\_RED}), \text{length}(\text{CORE\_GREEN}))$  **do**
- 10:      $\text{Demand}(i) \leftarrow \text{CORE\_RED}(I_{sense,i}) + \text{CORE\_GREEN}(I_{sense,i})$
- 11:     **if**  $\text{Demand}(i) \leq 2 \cdot I_{avg}$  **then**
- 12:        $V_{\text{CORE\_RED}}(i) \leftarrow V_{\text{CORE\_GREEN}}(i)$   $\triangleright$  Align the  $V_{dd}$  levels of the two cores whose OCVR outputs are being combined
- 13:        $Switch_{\text{CORE\_RED}(i), \text{CORE\_GREEN}(i)} \leftarrow 1$   $\triangleright$  Close the switch so that *CORE\_RED*( $i$ ) is served by the OCVR connected to *CORE\_GREEN*( $i$ )
- 14:       Delete *CORE\_RED*( $i$ ) and *CORE\_GREEN*( $i$ ) from the respective arrays
- 15:     **end if**
- 16:   **end for**
- 17: **end procedure**
- 18: **procedure** *OCVR\_DECLUSTER*
- 19:   **for** each nonzero element in sparse matrix *Switch* **do**
- 20:     **if**  $I_{sense,i} + I_{sense,j} < 2 \cdot I_{avg} - \Delta I$  **then**
- 21:        $Switch_{i,j} \leftarrow 0$   $\triangleright$  Open the switch connecting core  $i$  with OCVR  $j$
- 22:     **end if**
- 23:   **end for**
- 24: **end procedure**

---

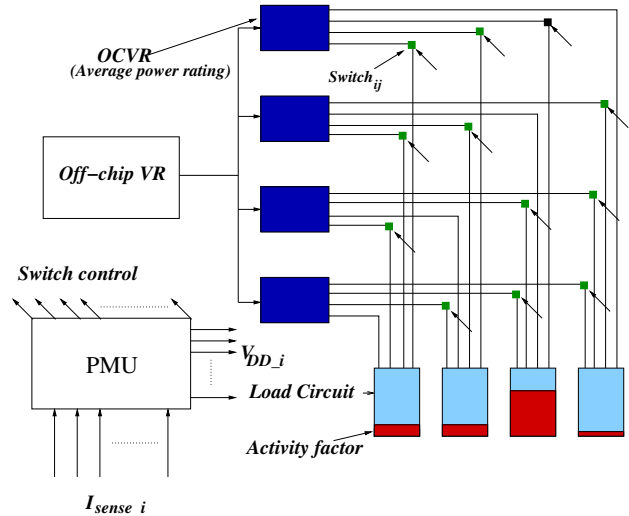


Figure 2: Proposed interconnected on-chip power delivery network with run-time voltage regulator clustering through a switching fabric.

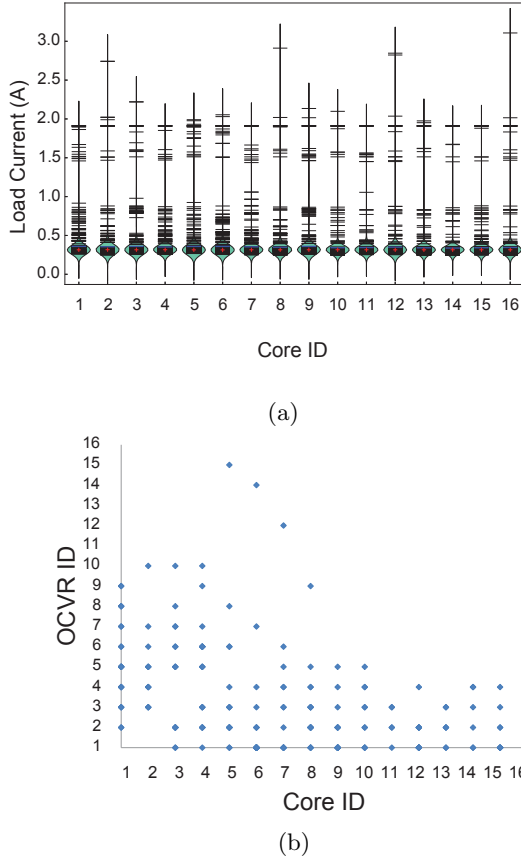


Figure 3: Simulated results of the implementation of Algorithm 1 for a 16-core CMP system with 16 OCVRs each with a peak rating of 0.6 A. (a) Load current variation per core and (b) cumulative configuration of the switches for 1000 CPU cycles.

the HSS to source additional current from the OCVRs which are located nearest to the cluster and are sourcing current less than  $I_{avg}$ . The logic controlling the HSS fabric within the PMU operates on two system parameters; the  $V_{dd}$  levels and the total load current sensed from each core cluster. The analysis of the power consumption provided in Section 3 indicates that the probability of the load current demand exceeding  $I_{avg}$  is 22%. As a result, there is always more than one core operating at or below  $I_{avg}$ . The PMU is provisioned to add at least one additional OCVR to serve a core requiring current higher than  $I_{avg}$ . The sum of the decision time of the PMU and the time to reconfigure the switches must be less than or equal to the load current transient response time (current slew-rate) of an OCVR with a current rating of  $I_{peak}$  to ensure an uninterrupted power supply to the core cluster.

The switching control of the HSS fabric is described by Algorithm 1. Algorithm 1 is implemented in the Python programming language and is analyzed with the parameters summarized in Table 2. A stochastic model of the current consumption of the cores in a CMP system is developed based on the statistical parameters captured from the per cycle power consumption analysis of the SPEC benchmarks. The load current obtained from the stochastic model for

Table 2: Simulation parameters for Algorithm 1.

Parameter	Value
Number of Cores	8, 16, 32, 64, 128
DVS levels	0.7 V, 0.8 V, 0.9 V, 1 V
OCVR current rating ( $I_{avg}$ )	0.6 A
Current threshold ( $\Delta I$ )	0.1 A
Load current variation	SPEC benchmark power consumption statistics
Execution duration	10 million CPU cycles

1000 CPU cycles across 16 cores is shown in Fig. 3(a). The peak current rating of each OCVR is set to 0.6 A ( $I_{peak}$ ), which is one order of magnitude less than the  $I_{peak}$  obtained through McPAT. The active switches required to support the run-time load current variation on each core for 1000 CPU cycles are shown in Fig. 3(b). A generic statistical load current model is also analyzed through Monte Carlo simulations with a maximum possible value of  $I_{peak}$ . Four DVS levels, listed in Table 2, are selected corresponding to the core configuration provided in Table 1.

## 4.2 Energy efficiency of CMP system

The switching DC-DC buck converter offers superior power supply voltage regulation and is therefore an optimum choice to power the cores. The circuit implementation of a buck converter consists of a switching network and a passive low pass filter. The inductor in the low pass filter acts as a low-loss energy transfer device which improves the power conversion efficiency. In this section, the PCE of the buck converter is analyzed for changes with the peak load current rating. The goal is to analyze the impact on the energy efficiency of the CMP when designing the OCVRs to support only the average load current demand of the cores.

### 4.2.1 Power conversion efficiency

The power consumed by the buck converter ( $P_{buck}$ ) is given by (1) [12]. The  $P_{mos}$ ,  $P_{ind}$ ,  $P_{cap}$ , and  $P_{pwm}$  are the power loss in, respectively, the MOS power transistors and the cascaded buffers driving them, the inductor and capacitor of the filter circuit, and the pulse width modulator circuit. The detailed mathematical formulae of each of the components which contribute to  $P_{buck}$  are given in [12].

$$P_{buck} = P_{mos} + P_{ind} + P_{cap} + P_{pwm} \quad (1)$$

The power consumed by the filter circuit, power transistors, and the buffers driving them increases with the maximum supported output current of the buck converter. Alternatively, multiple phases are used to drive higher output currents. The circuit schematic of a buck converter with multiple phases of the filter circuit, MOS power transistors, and cascaded buffers is shown in Fig. 4. Two custom buck converters with maximum output current ratings of 6 A and 0.6 A are implemented [13]. The two converters represent voltage regulators that support the  $I_{peak}$  and  $I_{avg}$  currents of a CMP with core parameters as listed in Table 1. The circuit characteristics and power consumption of the components of the buck converters are listed in Table 3. The buck converter with a maximum output current rating of 0.6 A consumes 22.65 mW, which is 3.3% of the power consumed by the over-provisioned buck converter. The on-chip implementation of the two buck converters yields similar ratios between the power consumed by each passive component, although at a higher switching frequency to reduce the size of the filter inductor and capacitor.

Table 3: Circuit parameters and power consumption for two peak load currents of DC-DC switching buck converters [13].

Parameter	Maximum output current of 6 A	Maximum output current of 0.6 A
Peak to peak inductor ripple current	1.65 A	182.5 mA
Switching frequency	368.25 kHz	100.02 kHz
Duty cycle	43.4%	40.16%
Peak-to-peak output ripple voltage	1.924 mV	2.281 mV
Inductor power dissipation	260.83 mW	675 $\mu$ W
Output capacitor power dissipation	368.91 $\mu$ W	13.56 $\mu$ W
Total power dissipation	688.99 mW	22.65 mW
Footprint	209 $mm^2$	161 $mm^2$

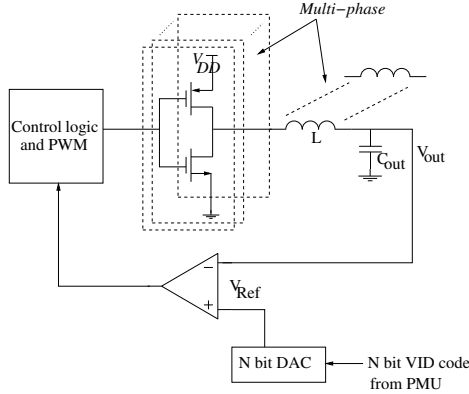


Figure 4: Schematic of a multi-phase DC-DC switching buck converter.

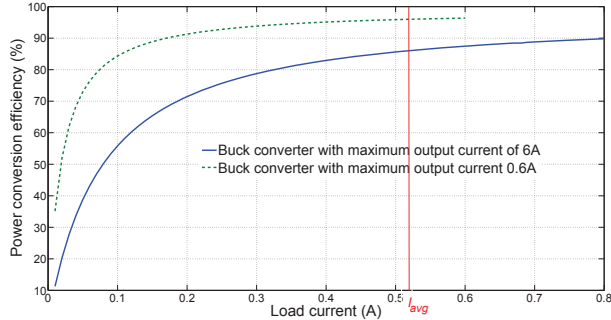


Figure 5: Power conversion efficiency with load current for two DC-DC buck converters with a maximum load current rating of 6 A and 0.6 A.

The large reduction in the power dissipation of the buck converter achieved by reducing the peak load current rating results in an improvement in the PCE. Although the over-provisioned buck converter offers a peak PCE of 89.7% at an output current of 6 A, the reduction in the PCE with decreasing output current is significant. The variation in the PCE with the output current for the two buck converters is shown in Fig. 5. The typical workloads executed on the CMP (with the core configuration listed in Table 1) consume currents in the range of  $I_{avg}$  and therefore the buck converter with an output rating of 0.6 A offers a higher average PCE for a majority of the run-time of the workloads.

#### 4.2.2 Improvement in energy efficiency

The total energy consumption of a CMP system implemented with a conventional PDN for a given execution time

$T_{epoch}$  with  $N$  cores and  $N$  OCVRs is given by (2). The cores are served by over-provisioned OCVRs identical to the buck converter with a maximum output current of 6 A. The dynamic and static power consumed by the cores in the presence of DVFS are given by  $P_{dynamic}$  and  $P_{static}$ , respectively.  $PCE_1$  represents the power conversion efficiency of the over-provisioned OCVR. At low load currents close to  $I_{avg}$ , the  $PCE_1$  offered by the over-provisioned buck converter is 87%. Alternatively, if the power delivery system is designed with each core supported by a buck converter that supplies a maximum output current of 0.6 A, the achieved  $PCE_2$  at  $I_{avg}$  is 96.36%. In addition, the static power consumed by idle cores or core clusters is close to zero as power gating through the HSS fabric is performed. The HSS fabric, however, imposes an additional switching loss  $P_{switch}$ , which is the dynamic power consumed by the PMOS transistors while switching, and a conduction loss  $P_{conduction}$  while in the ON state and passing the average current  $I_{avg}$ .

$$E_{CMP,conventional} = \left\{ \sum_{i=1}^N \frac{(P_{dynamic,i} + P_{static})}{PCE_1} \right\} \cdot T_{epoch} \quad (2)$$

The total energy consumed by the CMP with  $N$  OCVRs, where each OCVR is designed for an  $I_{avg}$  rating, and  $N_x(N-1)$  PMOS switches is given by (3). The parameters  $j$ ,  $k$ , and  $l$  are, respectively, the number of active cores consuming current below  $I_{avg}$ , the number of active core(s) consuming current above  $I_{avg}$ , and the number of idle core(s) power gated through the HSS network. In the case of idle cores, the power consumed by the OCVRs ( $P_{OCVR,leakage}$ ) is the only component contributing to the system energy. As described in Section 3, applications consume current less than  $I_{avg}$  for about 78% of the time. The  $P_{switch}$  loss is therefore incurred for 22% of the execution time of the workloads when the load current demand exceeds  $I_{avg}$ .

$$E_{CMP,proposed} = \sum_{t=1}^{T_{epoch}} \left\{ \sum_{i=1}^j \frac{(P_{dynamic,i} + P_{static})}{PCE_2} + \sum_{i=1}^k \frac{(P_{dynamic,i} + P_{static} + P_{switch} + P_{conduction})}{PCE_2} + \sum_{i=1}^l P_{OCVR,leakage} \right\}; \quad (3)$$

$$j + k + l = N$$

Circuit simulations of a PDN designed to support the  $I_{avg}$  for each of the 16 cores are performed to determine the energy consumption as given by (3). The 16 cores are simulated as piecewise constant current sinks. The current variation with time for the 16 current sinks is shown in Fig.

Table 4: Parameters of the PDN determined through SPICE simulation.

Parameter	Value
PMOS Width	800 $\mu\text{m}$
PMOS switching time	160 ps
Area occupied by 16x15 switching network	9600 $\mu\text{m}^2$
Per core maximum switching capacitance	2.4 nF
Piecewise constant load current variation	Stochastic model
$P_{switch}$	0.1 W
$P_{conduction}$	0.06 W
$P_{static}$ (obtained through McPAT)	0.175 W

3(a). A 16x15 PMOS switching network is implemented in a 45 nm technology. The gates of the PMOS switches are controlled through time varying voltage signals. The  $P_{switch}$  and  $P_{conduction}$  for the PMOS switch with an output capacitance provided by a single core is determined through SPICE simulations. The  $P_{static}$  and  $P_{dynamic}$  of a core is measured through McPAT. The  $P_{dynamic}$  for each core is overestimated as the power consumption per clock cycle is captured at the highest supported DVS level of 1 V. The parameter values of the switching network and the PDN are summarized in Table 4.

The additional switching and conduction losses due to the PMOS switches are an insignificant fraction of the total power consumed by the CMP as both are only consumed when the PDN is reconfigured to combine the outputs of the OCVRs. The energy consumption of the proposed power delivery system is up to 44% less than the energy consumed by the CMP with over-provisioned OCVRs and PDN. On average, there is a 15% reduction in the energy consumption as shown through simulations of the proposed PDN with a stochastic load current modeled on the SPEC benchmark power traces shown in Fig.1. By reducing the maximum rating of the OCVR (buck converter), the percentage reduction in the energy consumed for a core sinking current less than  $I_{avg}$  is 36%. The reduction in energy is due to the optimal PCE offered by the buck converter at the maximum output current supported. The energy efficiency of the CMP is therefore improved significantly by designing the power delivery system to support the average current demand of the cores. In addition, if a system level workload mapping technique is applied to distribute identical workloads on a cluster of cores, serving the core cluster with one OCVR that implements DVFS is advantageous to further reduce the energy consumption [6, 10].

### 4.3 Technique to prevent system failure

The power consumption analysis of the different workloads provided in Section 3 and the construction of the PDN in Section 4 ensure that there is always an OCVR available in the CMP system to support core(s) demanding higher than the average current(s). In the unlikely event that an OCVR is not located to support the higher than average current requirement of a core, the core is stalled by the PMU until an OCVR becomes available for clustering. The probability of not finding an OCVR for clustering is low when an efficient workload mapping technique is implemented. The performance penalty due to stalling the core is therefore negligible.

## 5. CONCLUSIONS

A load balanced circuit technique to deliver average power through on-chip voltage regulators (OCVRs) is developed.

The current rating of each OCVR is reduced to support only the average current demands of typical workloads executed on the CMP system. The reduction in the maximum output current of the OCVRs improves the power conversion efficiency, reduces the footprint of the PDN (voltage regulator and switch fabric) by at least 23%, and improves the energy efficiency of the CMP system by at most 44%. The simulated results indicate that the optimum OCVR configuration for a CMP system depends on the average load current requirement per core. The proposed inter-connected power delivery system is applicable to any OCVR circuit topology and offers higher reliability through a run-time clustering technique that prevents system failure when the current demand of the core exceeds the maximum output current supported by a single OCVR.

## 6. REFERENCES

- [1] W. Kim et al, "System Level Analysis of Fast, Per-Core DVFS Using On-chip Switching Regulators," *Proceedings of the International Symposium on High Performance Computer Architecture*, pp. 123–134, February 2008.
- [2] E.A. Burton et al, "FIVR—Fully Integrated Voltage Regulators on 4th Generation Intel Core SoCs," *IEEE Applied Power Electronics Conference and Exposition*, pp. 432–439, March 2014.
- [3] E.J. Fluhr et al, "The 12-Core POWER Processor With 7.6 Tb/s IO Bandwidth, Integrated Voltage Regulation, and Resonant Clocking," *IEEE Journal of Solid-State Circuits*, Vol. 50, No. 1, pp. 10–23, January 2015.
- [4] mersenne.org, "Great Internet Mersenne Prime Search," <http://www.mersenne.org/download/#stresstest>.
- [5] S. Li et al, "McPAT: an integrated power, area, and timing modeling framework for multicore and manycore architectures," *IEEE/ACM International Symposium on Microarchitecture*, pp. 469–480, 2009.
- [6] W. Godycki et al, "Enabling Realistic Fine-Grain Voltage Scaling with Reconfigurable Power Distribution Networks," *Proceedings of the IEEE/ACM International Symposium on Microarchitecture (MICRO)*, pp. 381–393, December 2014.
- [7] W. Lee et al, "Optimizing a Reconfigurable Power Distribution Network in a Multicore Platform," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 34, No. 7, pp. 1110–1123, July 2015.
- [8] M. Sai et al, "3D Many-core Microprocessor Power Management by Space-Time Multiplexing based Demand-supply Matching," *IEEE Transaction on Computers*, Vol. 64, No. 11, pp. 3022–3036, November 2015.
- [9] I. Vaisband and E.G. Friedman, "Energy efficient adaptive clustering of on-chip power delivery systems," *INTEGRATION, The VLSI Journal*, Vol. 48, pp. 1–9, January 2015.
- [10] M. Tavana et al, "Realizing Complexity-Effective On-Chip Power Delivery for Many-Core Platforms by Exploiting Optimized Mapping," *Proceedings of the International Conference on Computer Design*, pp. 581–588, October 2015.
- [11] D.M. Tullsen, "Simulation and Modeling of a Simultaneous Multithreading Processor," *Proceedings of the International Conference for the Resource Management & Performance Evaluation of Enterprise Computing Systems, CMG. Part 2(of 2)*, pp. 819–828, 1996.
- [12] E. Salman and E.G. Friedman, *High Performance Integrated Circuit Design*, McGraw Hill, 2012.
- [13] Texas Instruments, "WEBENCH Design Center," <http://webench.ti.com>.