Thermal Implications of On-Chip Voltage Regulation: Upcoming Challenges and Possible Solutions

Selçuk Köse
Department of Electrical Engineering
University of South Florida
Tampa, Florida
kose@usf.edu *

ABSTRACT

The primary objective of this paper is to investigate and evaluate the thermal implications of high power density on-chip voltage regulators. This paper is a first attempt to highlight the importance of the number, size, and location of on-chip voltage regulators on the thermal hotspots and thermal gradient. The physical location of on-chip voltage regulators is explored to distribute the hotspot locations and achieve spatial low pass filtering of the hotspots. A new thermal-aware physical design and power management technique are proposed to spatially and temporally distribute the hotspot locations over the cooler areas within an integrated circuit. The proposed technique eliminates the thermal gradient due to on-chip voltage regulators without any performance loss.

Categories and Subject Descriptors

EDA3.1 [**Cross-Layer Power Analysis and Low-Power Design**]: System-level low-power design and thermal analysis, simulation and management

Keywords

On-chip power conversion, power and thermal management

1. INTRODUCTION

Power consumption has become the primary bottleneck for modern integrated circuits (ICs). To satisfy the continuous demand for increased functionality, greater computational power, and larger memory resources under a stringent power budget in modern integrated circuits, aggressive techniques are employed for voltage regulation, distribution, and management [1]. On-chip voltage regulation enables fast load regulation and small voltage scaling time for dynamic voltage/frequency scaling (DVFS), and lowers the power noise due to off-chip parasitic impedances [2].

Distributed on-chip voltage regulation is an emerging research area where multiple voltage regulators are connected in parallel, delivering current to the same power network close to the load circuits [3–9]. Switched capacitor (SC) and low dropout (LDO) voltage regulators are used for distributed voltage regulation due to the small size. Parallel LDO regulators can suffer from device mismatch, offset voltages among parallel regulators, overall system stability, and balanced current sharing. Stability is therefore a significant problem for parallel LDO regulators [5]. Alternatively, multiple phase stages within an SC voltage converter are interleaved to reduce the output voltage ripple, maintain a stable operation, and achieve parallel voltage regulation.

Bulzacchelli *et al.* achieved 500 ps transient response time with a system of eight distributed LDO regulators [3]. Lai *et al.* have provided detailed guidelines to ensure stability of a distributed voltage regulation system composed of LDO regulators [5] based on a hybrid stability theory. Zhou *et al.* have investigated distributed SC voltage regulators and proposed algorithms to determine the optimum number and location of the regulators [6]. Anderson *et al.* have recently proposed an SC regulator with 4.6 W/mm^2 power density and 86% efficiency using deep-trench capacitors [8].

Despite the significant amount of research on on-chip voltage regulation, to the best of author's knowledge, the thermal implications of on-chip voltage regulation have yet been investigated. The trend for integrating smaller on-chip voltage regulators that can provide higher load current leads to high-power density regulators. Due to the finite conversion efficiency, a portion of the power is dissipated as heat, causing thermal gradient and possible hotspots.

Dynamic thermal management is an important technique to reduce the thermal gradients, and spatial and temporal hotspots at runtime in modern ICs [10–12]. Although power management techniques generally reduce the on-chip temperature, power and thermal management can sometimes conflict [12]. Power management is typically used when certain circuit blocks go into a reduced power state, whereas temperature management is almost always required when the chip is fully-utilized. Additionally, since the changes in temperature are relatively slower than those in the power consumption of circuit blocks, power management techniques need to be active for an extended time period to also affect thermal behavior [12]. Furthermore, existing thermal and power management techniques typically focus on the load circuits and provide techniques to reduce the temperature and/or power consumption of only the load circuits. In this paper, on-chip voltage regulators are exploited to reduce the thermal gradient and hotspots.

Our contributions in this paper are as follows:

*This work was supported in part by the National Science Foundation CAREER grant under contract No. CCF-1350451.

DAC '14, June 01 - 05 2014, San Francisco, CA, USA

http://dx.doi.org/10.1145/2593069.2593231.

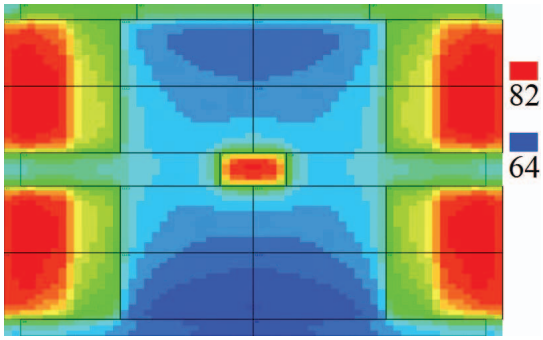

Figure 1: Thermal map of an 8-core Xeon processor when all the cures are active. Voltage regulators are implemented off-chip.

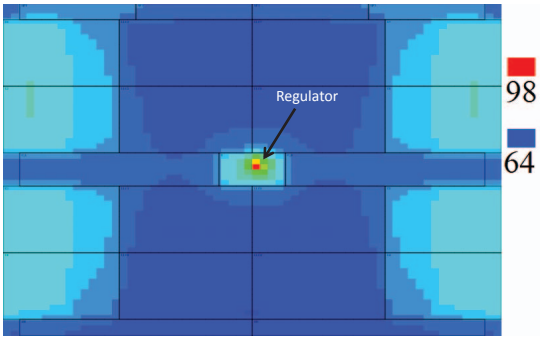

Figure 2: Thermal map of an 8-core Xeon processor. All the cores are active and a single voltage regulator placed at the center of the die provides ∼50% of the current demand of the cores.

- Thermal implications of on-chip voltage regulation are explained with extensive thermal simulations for various scenarios.
- An optimization function is proposed to determine the optimum location of individual parallel regulators for thermal-aware physical design.
- A thermal-aware regulator activity management technique, regulator-gating (ReGa), is proposed to adaptively turn on and off individual voltage regulators. Preliminary simulations are performed to evaluate the speed of ReGa.
- The working principles of the proposed ReGa methodology are explained with a flowchart for noise- and thermal-aware management of on-chip voltage regulators.

The remaining part of the paper is organized as follows. Thermal implications of on-chip voltage regulation and related background are presented in Section 2. In Section 3, the proposed thermal-aware physical design and regulator-gating techniques are discussed. The paper is concluded in Section 4.

2. THERMAL IMPLICATIONS OF ON-CHIP VOLTAGE REGULATION

A quite important and overlooked aspect of on-chip voltage regulation is the heat generated by the on-chip voltage regulators. As a result of the limited power efficiencies, miniaturizing voltage regulators to minimize on-chip area can lead to increased thermal gradient and local hotspots. The power efficiency of a linear voltage regulator is

$$\eta_{ldo} = P_{out}/P_{in} = (I_{out} * V_{out})/(I_{in} * V_{in}). \quad (1)$$

Since the current efficiency (I_{out}/I_{in}) is typically over 98% for LDO regulators, the power efficiency can be approximated as V_{out}/V_{in}. The power loss $P_{loss} = \sim I_{out} * (V_{in} - V_{out})$ is dissipated as heat at the output pass transistor, which typically occupies 30 - 40% of the regulator area. Alternatively, the power efficiency of an SC regulator is

$$\eta_{sc} = P_{out}/P_{in} = P_{load}/(P_{load} + P_{cond} + P_{sw} + P_{par} + P_{cont}), \quad (2)$$

where P_{cond}, P_{sw}, P_{par}, and P_{cont} are, respectively, the conduction, switching, parasitic, and control related power losses. The conduction (P_{cond}) and switching power loss (P_{sw}) are the primary loss mechanisms. With high density trench capacitors and charge recycling techniques, P_{loss} can be mildly mitigated [8]. The conduction loss due to the resistive switches, however, remains as a significant power loss mechanism. The resistive switches typically occupy less than ∼30% of the SC regulator area. Please note that the pass transistors within an LDO and the resistive switches within an SC regulator occupy a fraction of the regulator area where most of the power is dissipated. The power dissipation density (and as a result temperature) on these active parts of a voltage regulator is significantly higher than on the rest of the regulator.

For example, a fully on-chip switched capacitor (SC) regulator with 4.6 A/mm^2 current density and 86% power efficiency occupies 3,444 μm^2 chip area while generating 0.9 volts [8]. Since the power efficiency is 86%, the dissipated power density within the regulator is 4.6*(14/86)=0.75 W/mm^2. 65% of the regulator area is occupied by the trench capacitor (∼2,240 μm) whereas the drivers and switches occupy ∼1,205 μm. Since most of the power is dissipated in the switches and drivers as compared to the power dissipated by the trench capacitor, the power dissipation density in the switches and drivers can be approximated as 2.14 W/mm^2. Assuming a cooling limit of 1 W/mm^2 [13], these regulators will cause potential hotspots. The power-efficiency further reduces from 86% to ∼70% while delivering a higher load current [8]. In this case, the power dissipation and, thus, heat density will increase over 2×. *ISSCC 2013 Trends* predicts that the power density of on-chip voltage regulators will continue to increase [14].

To highlight the importance of this task, the thermal gradient map of an 8-core Xeon processor [15], obtained using HotSpot [16], is illustrated in Fig. 1. The highest on-chip temperature for this base case without any on-chip regulators is ∼82°C. Assuming an on-chip voltage regulator, proposed in [8], is placed physically at the center of the die providing ∼50% of the total current demand of the cores during their full utilization, a severe thermal hotspot is generated around the voltage regulator, as shown in Fig. 2. The maximum local temperature increases to ∼98°C (*i.e.*, ∼16°C higher local temperature as compared to the base case without on-chip regulation). In the next subsection, the effects of spatial and temporal low pass filtering of the thermal gradient are briefly explained.

2.1 Spatial and Temporal Temperature Low Pass Filtering

The location and duration of the power dissipation directly impact the temperature. An analogy has been drawn between the electrical characteristics of a first order RC filter and the heat transfer in the spatial domain in [11]. The thermal spatial capacitance has been described to explain

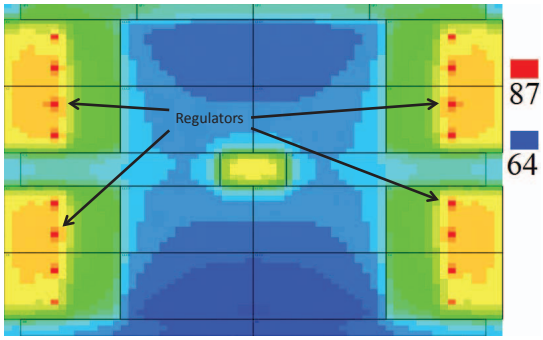

Figure 3: Thermal map of an 8-core Xeon processor with 16 on-chip voltage regulators distributed across the die within the cores.

the spatial temperature low pass filtering effect of distributing thermal hotpots of highly active circuit blocks. In this paper, spatial temperature low pass filtering is used to distribute the hotspots generated by the on-chip regulators rather than the hotspots generated by active load circuits.

The Fourier Law of heat transfer states that

$$Q(x)(t_2 - t_1) = k\frac{dT(x)(t_2 - t_1)}{dx}, \tag{3}$$

where the $Q(x)$, k, and $T(x)$ are, respectively, the heat transfer at location x, thermal conductivity of the medium, and temperature distribution function. The average heat energy transfer between times t_1 and t_2 is proportional to the amplitude of the temperature distribution. The average temperature between t_1 and t_2 can be reduced if the related heat generators (*i.e.* power dissipating circuits) are turned off between t_1 and t_2. By adaptively turning off the voltage regulators, temporal low pass filtering of temperature is achieved, as shown this paper.

3. THERMAL-AWARE DESIGN

The thermal hotspots can be mitigated by being spread in time (temporal filtering) or in space (spatial filtering) [11]. Thermal-aware physical design aims to spread the hotspots spatially whereas architectural level techniques spread the concentrated heat both temporally and spatially. In this section, floorplanning and architectural level thermal-aware design techniques are proposed.

3.1 Thermal-Aware Physical Design

The primary objective of parallel on-chip voltage regulation is to place the regulators much closer to the load circuits to minimize the effect of parasitic impedance of the on-chip power grid, and thus minimize power noise [3, 5, 6]. In this paper, for the first time, parallel voltage regulation is exploited to spread the concentrated heat generated by on-chip voltage regulators. Rather than providing all the current demand with a large stand-alone voltage regulator and, thus, potentially having a thermal hotspot in close proximity to the regulator, multiple smaller voltage regulators provide the required current to the power grid and spread the concentrated heat to multiple locations across the die.

The thermal implications of using multiple smaller regulators instead of a single relatively larger regulator are evaluated for an 8-core Xeon processor. An optimization function is proposed to determine the optimum location of parallel LDO voltage regulators that minimize thermal gradient. The location of the voltage regulators is determined by maximizing the effective resistance i) among different voltage regulators and ii) between the voltage regulators and high activity load circuits.

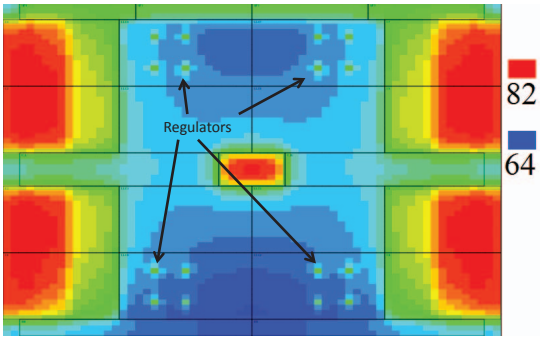

Figure 4: Thermal map of an 8-core Xeon processor with 16 on-chip voltage regulators distributed regularly within cooler areas of the die.

$$\begin{aligned} Minimize \quad & \sum_{j=1}^{m}\sum_{i=1}^{n} C_{P_{ij}}(R_{out}(P_i) + R_{eff}(P_i, L_j)) \\ & - \sum_{k=i}^{n-1}\sum_{i=k+1}^{n} R_{eff}(P_k, P_i) \end{aligned} \tag{4}$$

where n, m, R_{out}, and R_{eff} are, respectively, the number of LDO regulators, the number of high activity load circuits, output resistance of the LDO, and effective resistance between any two nodes within the power grid [17]. In the optimization function, the first term is used to place the regulators farther from the high-activity load circuits which potentially cause hotspots while the second term is used to increase the physical distance among different LDO regulators to further improve the distributed nature of the voltage regulation.

First, to highlight the effect of distributed on-chip voltage regulators that are placed close to the high-activity load circuits, the voltage regulators are only allowed to be placed within the cores. The optimum location of 16 LDO regulators is determined by using (4). When 16 SC voltage regulators [8] are placed in the floorplan that provide ∼50% of the core-current, the maximum local temperature decreases significantly from 98°C (die with a single voltage regulator placed at the center of the die) to 87°C, as shown in Fig. 3. The local temperature is 5°C degrees higher than the base case without any on-chip voltage regulator (87°C as compared to 82°C). This difference can be reduced and a more uniform temperature gradient map can be obtained by increasing the number of voltage regulators and distributing them across the die. Increasing the number of on-chip voltage regulators is, however, not free and comes with overheads. When the number of parallel regulators connected to the same power grid increases, maintaining the stable operation becomes more challenging [5]. Furthermore, the area of a voltage regulator does not scale linearly while increasing the number of regulators. Although the output pass transistor within an LDO can scale quite linearly, most of the remaining portion of the regulator circuit should still be implemented for each smaller regulator. The advantages and possible tradeoffs of increasing the number of on-chip voltage regulators are summarized in [18].

Second, the constraint on placing the regulators within the cores is lifted to allow the optimizer to place the LDOs in cooler areas. The maximum local temperature is reduced by placing the regulators around cooler areas such as the L2 cache, as illustrated in Fig. 4. Please note that the voltage regulators still increase the local temperature as indicated

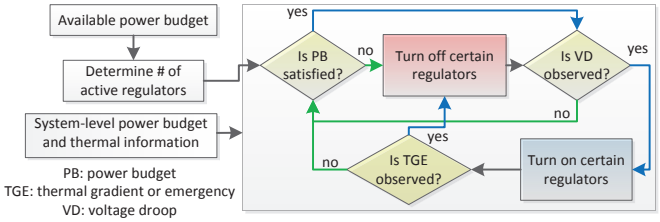

Figure 5: Proposed ReGa management methodology that can be integrated into the mainstream power management flow.

with light blue regions in the dark blue portions of the thermal map. However, since the regulators are placed at some of the lowest temperature regions of the die, the additional heat generated by the voltage regulators does not increase the temperature sufficiently higher to form a local hotspot.

Placing the on-chip regulators farther from the warmer regions of the die in temperature-aware physical design can conflict with the primary motive of parallel on-chip voltage regulation, which is to move the regulators closer to the active circuits to minimize the power noise and improve the response time. This conflict can be easily solved if the regulators are adaptively moved farther from the hotspot locations during thermal emergencies and moved closer to the active load circuits at other times. Since the utilization of different circuit blocks varies at runtime, the current demand of these circuits and the corresponding thermal gradient map of the die change. An adaptive management of voltage regulators, which is discussed in the next section, greatly enhances the noise and thermal behavior of the circuits.

3.2 Thermal-Aware Regulator-Gating

Existing work focuses on mitigating the thermal hotspots generated by the active load circuits by temporally or spatially distributing their functionality. In this paper, a thermal-aware voltage regulator management technique, *regulator-gating* (ReGa) is proposed. Regulator-gating is defined as adaptively turning on and off individual voltage regulators within a system of parallel voltage regulators at runtime. Since multiple voltage regulators provide current to the same power grid, some of the regulators can be turned off when the current demand is reduced.

3.2.1 Regulator-gating methodology

The thermal-aware ReGa is utilized to turn off the regulators close to the thermal hotspot locations during thermal emergencies. This provides both temporal and spatial low-pass filtering of the hotspots by spreading the heat across the die and will reduce the on-chip thermal gradient [11]. Utilizing ReGa will does not completely solve the on-chip thermal problems but significantly mitigates the hotspot creation by the on-chip regulators. This thermal-aware ReGa management considers the thermal emergencies and thermal gradient, and can be integrated into the mainstream power and thermal management flow.

The proposed ReGa management methodology is illustrated in Fig. 5 The number of required active regulators is determined based on the available power budget dictated by the system level power control unit. After the initiation of the power delivery system, the power budget, quality of the supply voltage, and the thermal emergencies are continuously monitored at runtime to control the activity of the individual voltage regulators. The performance and activity counters, thermal sensors, and voltage or current sensors which already exist in high-performance low-power

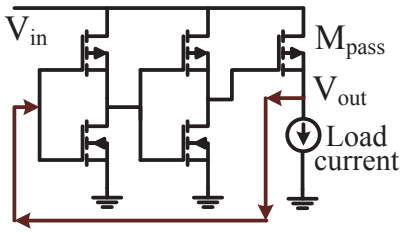

Figure 6: Digital LDO (DLDO) proposed in this paper to support fast load regulation.

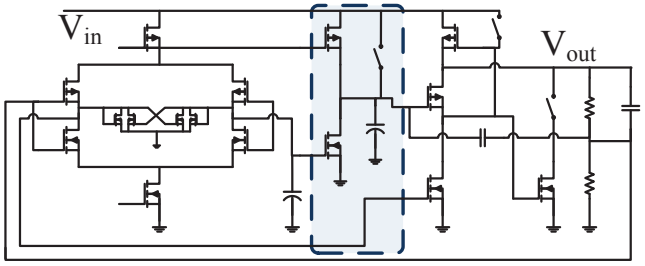

Figure 7: LDO, proposed by Lai *et al.* [20], has been used in the simulations.

integrated circuits are leveraged while implementing ReGa methodology.

3.2.2 Regulator-gating: proof of concept

A distributed power delivery network is constructed with parallel LDO and digital-LDO (DLDO) regulators to provide a *proof of concept* for the thermal-aware ReGa methodology. A simple DLDO regulator has been used, as shown in Fig. 6, with two skewed inverters to sense the changes at the output voltage and generate a transient signal to control the gate voltage of a pass transistor M_{pass}, permitting an instant response to transient changes. A drop at the output voltage V_{out} causes the pass transistor to provide a higher current due to the increased gate voltage. The DLDO regulator is similar to the circuit proposed in [19] with certain fundamental differences. The voltage sense portion is significantly simpler and a single pass transistor is used.

Due to the smaller area of the sense transistors, multiple copies of the proposed DLDO regulator can be distributed across the die in parallel with LDO regulators and can provide a fast response time of ∼400 ps (see Fig. 9). Multiple copies of these DLDO regulators are connected in parallel with the LDO regulator proposed by Lai *et al.* in [20]. The inverting amplifier stage of the LDO regulator, shown in dotted box in Fig. 7, has been modified to enhance the dynamic response while minimizing the quiescent current consumption at this stage.

Seven LDO and three DLDO regulators are connected to a small power network with 400 nodes, as depicted in Fig. 8. The current contribution from individual regulators to the power grid is shown in Fig. 9 when the load current demand increases from 11 mA to 80 mA. While only one LDO regulator (LDO 7) is sufficient to provide a robust 11 mA current to the load, the rest of the LDO regulators turn on and start providing current to the power grid when the load current demand increases to 80 mA. DLDO regulators turn on immediately after sensing a voltage drop at the power grid and provide instant current to the grid while the LDO regulators are turning on, as shown in Fig. 9. The DLDO regulators remain active only for a couple of nanoseconds (∼4 ns) until the LDO regulators turn on. The DLDO reg-

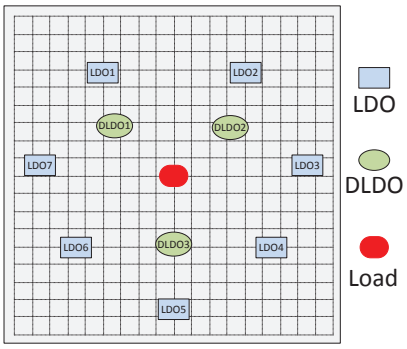

Figure 8: Illustration of a distributed power network with 7 LDO and 3 DLDO regulators connected in parallel.

ulators are self-activated, whereas the LDO regulators are controlled by the system-level (global) controller for power and thermal management.

The current efficiency of the LDO regulators is around 98% (∼300 μA quiescent current while providing 11 mA current). The quiescent current increases while providing a lower output current[1], as discussed in [20], and doubles to ∼600 μA when the output current is lower than 2 mA. In this case, the current efficiency of the LDO regulator becomes 2/2.6=∼77%. In the case study where the load current is 11 mA, a single LDO regulator provides the required current with 98% current efficiency. If all of the seven LDO regulators were active while providing 11 mA load current, each LDO regulator would contribute less than 2 mA current to the load with a current efficiency of less than 77%. Without regulator-gating, the total power dissipated during voltage conversion while providing 11 mA load current would be $V_{in} * I_{in} - V_{out} * I_{out}$ = 1.2 V * (11 mA + 7 * 0.6 mA) - 1 V * 11 mA = 7.24 mW. With regulator-gating, the total power dissipated during voltage conversion is $V_{in} * I_{in} - V_{out} * I_{out}$ = 1.2 V * 11.3 mA - 1 V * 11 mA = 2.56 mW. This preliminary study demonstrates that power delivery system not only can be utilized to reduce local temperatures but also is ∼3× more power-efficient with ReGa when certain regulators are gated during the idle periods of time.

On-chip voltage regulators come with some overheads, such as area and reduced power efficiencies. Below, some of the possible unique overheads of ReGa technique are addressed and justified. The justification is performed assuming that modern ICs already have a basic power management system infrastructure such as performance and activity counters, thermal sensors, and voltage or current sensors.

Speed of ReGa.

With the utilization of DLDO regulators, the turn on and off time is decreased to sub-nanosecond range (400 ps in our example). For most of the applications, this turn on and off time does not degrade system performance.

Area overhead of ReGa.

Assuming that the chip already has on-chip voltage regulation, control circuitry for power/clock gating, and on-chip sensors and performance counters, the area overhead of ReGa will be the additional DLDO regulators, which are already used without ReGa [19]. When a firmware is used, there is no additional area overhead for ReGa power and thermal management.

[1]An increase in the quiescent current is typically observed when an LDO regulator has an AB amplifier type output stage

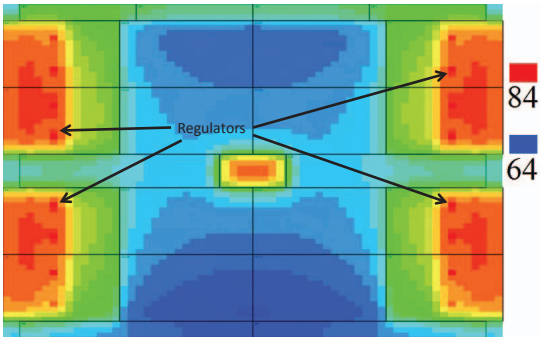

Figure 10: Thermal map of an 8-core Xeon processor with 16 on-chip voltage regulators. ReGa is performed to mitigate the thermal hotspots.

Power overhead of ReGa.

The additional overhead of the proposed ReGa methodology is the power dissipation during turning on and off voltage regulators. The power dissipation to turn on an LDO is less than ∼0.1 mW, and the power dissipated by the DLDO is negligible (*i.e.* ∼0.02 mW) when providing 15 mA output current.

3.2.3 *Thermal-aware regulator-gating*

We assume that the cores are utilized 50% of the time and 16 regulators are placed within the cores in the 8-core Xeon processor. When a core goes into an idle state, the voltage regulators within this core can be temporarily turned off during the idle period. In the analysis, we assume that each core is idle for 1 ms and then becomes active for 1ms. Each consecutive core goes into the idle state with a 125 μs delay (*i.e.*, if core1 goes into idle state at time t μs, core2 and core3 enter idle state at $(t + 125)\mu s$ and $(t + 250)\mu s$, respectively). We also assume that ReGa takes negligible time to turn on and off the individual regulators since the power delivery system can react in 400 ps, as discussed in Section 3.2.2. The corresponding thermal map of the 8-core Xeon processor with 16 voltage regulators, whose activity is controlled by ReGa technique, is shown in Fig. 10. For the same floorplan where the regulators are placed at the same locations within the cores, with ReGa technique, the maximum temperature can be decreased to 84°C from 87°C, which translates into 60% reduction in the additional temperature increase due to the on-chip voltage regulators. While reducing the maximum temperature, ReGa also smoothens the thermal gradients as shown in Fig. 10. By utilizing ReGa, on-chip voltage regulators can be placed closer to the load circuits (*i.e.*, within the cores) while keeping the maximum temperature at a reasonable level.

4. CONCLUSIONS

Thermal implications of on-chip voltage regulation have been investigated in this paper for the first time. The thermal simulations are performed for an 8-core Xeon processor using HotSpot. An optimization technique is developed to determine the optimum location of voltage regulators to minimize thermal gradient. Thermal-aware physical design and regulator activity management techniques are proposed. The pros and cons of these techniques that emerge as a result of mitigating the thermal hotspots and spreading the thermal gradients are explained. Turning on and off local voltage regulators to adaptively control the regulator activity further decreases the thermal hotspots and smoothens the thermal gradients.

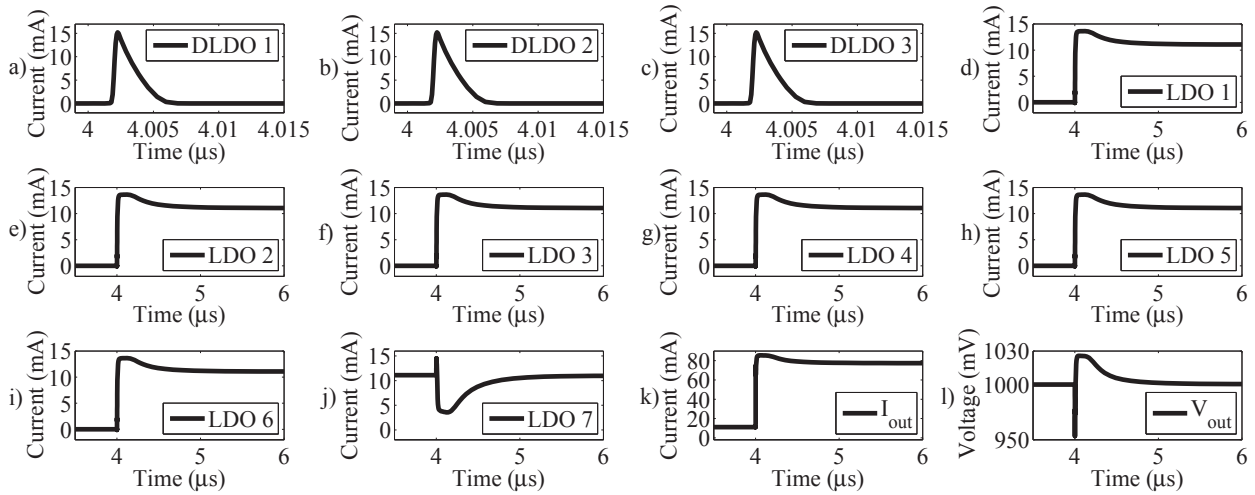

Figure 9: Response time of multiple LDO and DLDO regulators connected in parallel to the same power grid when the load current is increased from 11 mA to 80 mA.

5. REFERENCES

[1] R. Jakushokas *et al.*, *Power Distribution Networks with On-Chip Decoupling Capacitors, Second Edition*, Springer, 2011.

[2] S. Kose *et al.*, "Active Filter Based Hybrid On-Chip DC-DC Converters for Point-of-Load Voltage Regulation," *IEEE Transactions on VLSI Systems*, Vol. 21, No. 4, pp. 680–691, April 2013.

[3] J. F. Bulzacchelli *et al.*, "Dual-Loop System of Distributed Microregulators With High DC Accuracy, Load Response Time Below 500 ps, and 85-mV Dropout Voltage," *IEEE Journal of Solid-State Circuits*, Vol. 47, No. 4, pp. 863–874, April 2012.

[4] S. Kose and E. G. Friedman, "Fast Algorithms for IR Voltage Drop Analysis Exploiting Locality," *IEEE/ACM Design Automation Conference*, pp. 996–1001, June 2011.

[5] S. Lai, B. Yan, and P. Li, "Stability Assurance and Design Optimization of Large Power Delivery Networks with Multiple On-Chip Voltage Regulators," *IEEE/ACM International Conference on Computer-Aided Design*, pp. 247–254, November 2012.

[6] P. Zhou *et al.*, "Optimization of On-Chip Switched-Capacitor DC-DC Converters for High-Performance Applications," *IEEE/ACM International Conference on Computer-Aided Design*, pp. 263–270, November 2012.

[7] S. Kose and E. G. Friedman, "Distributed On-Chip Power Delivery," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, Vol. 2, No. 4, pp. 704–713, December 2012.

[8] T. M. Andersen *et al.*, "A 4.6 W/mm^2 power density 86% efficiency on-chip switched capacitor DC-DC converter in 32 nm SOI CMOS," *IEEE Applied Power Electronics Conference and Exposition*, pp. 692–699, March 2013.

[9] O. A. Uzun and S. Kose, "Converter-Gating: A Power Efficient and Secure On-Chip Power Delivery System," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, August 2014.

[10] A. K. Coskun, T. S. Rosing, and K. Whistnan, "Temperature Aware Task Scheduling in MPSoCs," *Conference on Design, Automation and Test in Europe*, pp. 1659–1664, April 2007.

[11] K. Sankaranarayanan *et al.*, "Architectural Implications of Spatial Thermal Filtering," *Integration, the VLSI Journal*, Vol. 46, No. 1, pp. 44–56, January 2011.

[12] J. Kong, S. W. Chung, and K. Skadron, "Recent Thermal Management Techniques for Microprocessors," *ACM Computing Surveys (CSUR)*, Vol. 44, No. 3, pp. 13:1–13:42, June 2012.

[13] P. Hazucha *et al.*, "A Linear Regulator with Fast Digital Control for Biasing Integrated DC-DC Converters," *Proceedings of the IEEE International Solid-State Circuits Conference*, pp. 2180–2189, February 2006.

[14] "ISSCC 2013 Trends," 2013. [Online]. Available: http://isscc.org/doc/2013/2013_Trends.pdf.

[15] S. Rusu *et al.*, "A 65-nm Dual-Core Multithreaded Xeon® Processor with 16-MB L3 Cache," *IEEE Journal of Solid-State Circuits*, Vol. 42, No. 1, pp. 17–25, January 2007.

[16] W. Huang *et al.*, "HotSpot: A Compact Thermal Modeling Methodology for Early-Stage VLSI Design," *IEEE Transactions on VLSI Systems*, Vol. 14, No. 5, pp. 501–513, May 2006.

[17] S. Kose and E. G. Friedman, "Effective Resistance of a Two Layer Mesh," *IEEE Transactions on Circuits and Systems II: Express Briefs*, Vol. 58, No. 2, pp. 739–743, November 2011.

[18] W. Kim, M. S. Gupta, G.-Y. Wei, and D. Brooks, "System Level Analysis of Fast, Per-Core DVFS Using On-Chip Switching Regulators," *Proceedings of the IEEE International Symposium on High Performance Computer Architecture*, pp. 123–134, February 2008.

[19] Y. Okuma *et al.*, "0.5-V Input Digital Low-Dropout Regulator (LDO) with 98.7% Current Efficiency in 65 nm CMOS," *IEICE Transactions on Electronics*, Vol. 94, No. 6, pp. 938–944, June 2011.

[20] S. Lai and P. Li, "A Fully On-Chip Area-Efficient CMOS Low-Dropout Regulator with Fast Load Regulation," *Analog Integrated Circuits and Signal Processing*, Vol. 72, No. 2, pp. 925–1030, February 2012.