

# Enabling Realistic Fine-Grain Voltage Scaling with Reconfigurable Power Distribution Networks

Waclaw Godycki, Christopher Torng, Ivan Bukreyev  
Alyssa Apsel, Christopher Batten  
School of Electrical and Computer Engineering  
Cornell University



## 1 Abstract

Recent work has shown that monolithic integration of voltage regulators will be feasible in the near future, enabling reduced system cost and the potential for **fine-grain voltage scaling** (FGVS). In this project, we use architecture-level modeling to explore a new dynamic voltage/frequency scaling controller called the fine-grain synchronization controller (FG-SYNC+). **FG-SYNC+ enables improved performance and energy efficiency at similar average power for multithreaded applications with activity imbalance.** We then use circuit-level modeling to explore various approaches to organizing on-chip voltage regulation, including a new approach called reconfigurable power distribution networks (RPDNs). **RPDNs allow one regulator to "borrow" energy storage from regulators associated with underutilized cores resulting in improved area/energy efficiency and faster response times.** We evaluate FG-SYNC+ and RPDN using a vertically integrated research methodology, and our results demonstrate a 10–50% performance and 10–70% energy-efficiency improvement on the majority of the applications studied compared to no FGVS, yet RPDN uses 40% less area compared to a more traditional per-core regulation scheme.

## 2 Motivation

Monolithic integration using a standard CMOS process provides a tremendous cost incentive for integrating closed-loop voltage regulators on the die. Recent technology trends suggest that it is now becoming feasible to integrate switching regulators on-chip (e.g., Intel Haswell), enabling **reduced system cost** as well as the potential for **fine-grain voltage scaling** (FGVS) to exploit fine-grain activity imbalance in multi-threaded applications for performance and energy efficiency benefits.

### Integrated Voltage Regulation in Intel Haswell Processors

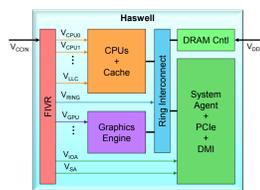
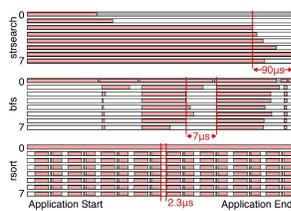


Figure from D. Kanter, MPR'13

### Fine-Grain Activity Imbalance in Multi-Threaded Applications



## 3 Target System

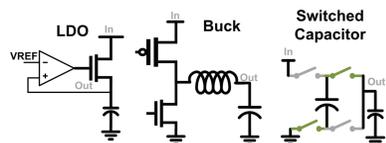
Our target system is an embedded processor composed of: eight in-order, single-issue, five-stage, RISC cores; private, coherent 16 KB instruction and data L1 caches; and a shared 512 KB unified L2 cache. We implemented the core and L1 memory system for this design in RTL and used a commercial standard-cell-based CAD toolflow targeting a TSMC 65 nm process to generate layout. Each core can run at 333 MHz at 1 V and the full eight-core system is approximately 6 mm<sup>2</sup>.



Target System Layout

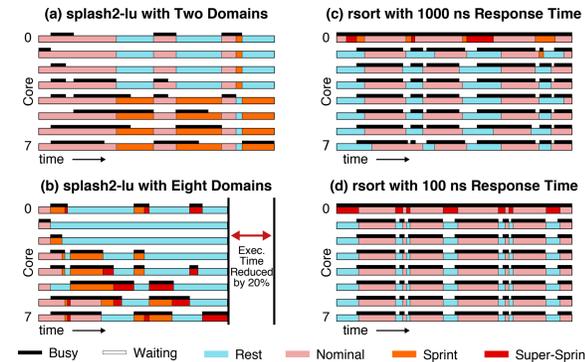
## 4 Types of Integrated Voltage Regulation

The three primary types of step-down voltage regulators are linear regulators (e.g., LDOs), inductor-based switching regulators (e.g., buck), and capacitor-based switching regulators (e.g., switched-capacitor).



## 5 FGVS Architecture Design: FG-SYNC+

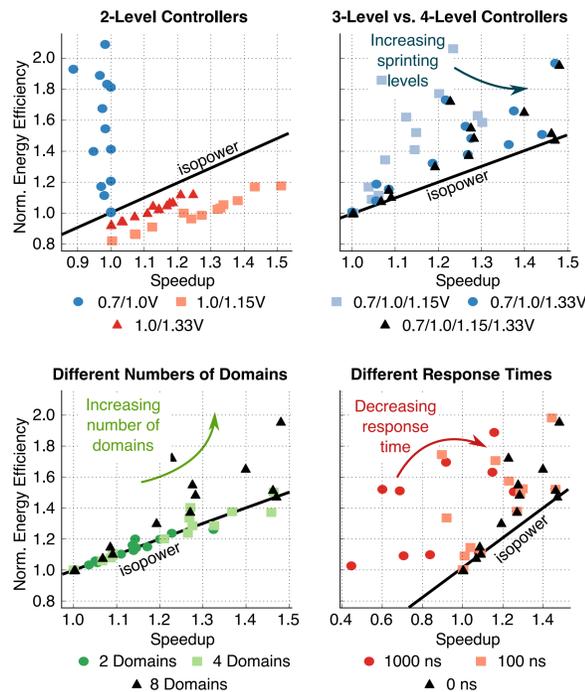
We explore a new FGVS controller called the fine-grain synchronization controller (FG-SYNC+) that exploits fine-grain scaling in **level** (i.e., many voltage levels), **space** (i.e., per-core regulation), and **time** (i.e., fast transition times between levels) to improve performance and energy efficiency while maintaining similar average power. FG-SYNC+ uses a thread library instrumented with hint instructions to inform the hardware about which cores are doing useful work vs. useless work (e.g., waiting for a task, waiting at a barrier).



In these application activity plots illustrating FG-SYNC+, rows show controller decisions per-core and black strips above cores show when that core is active. We compare SPLASH-2 LU factorization with two vs. four voltage domains (a,b). We also illustrate the impact of slow voltage-settling response times over a small excerpt from radix sort (c,d).

### Fine-Grain Voltage Scaling with FG-SYNC+

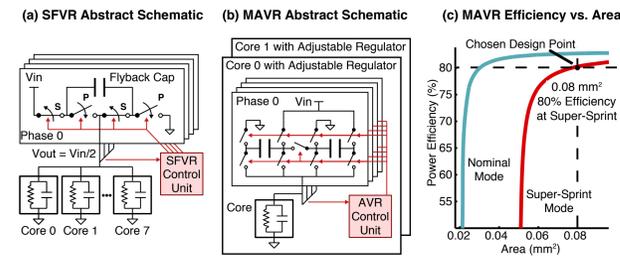
We use three sensitivity studies to understand the implication of varying: (1) the number of voltage levels, (2) the number of voltage domains, (3) and voltage-settling response times.



To exploit fine-grain activity imbalance, (1) at least three levels are required and four levels helps further; (2) more domains results in improved performance and energy efficiency; (3) response times of 100 ns or faster are required.

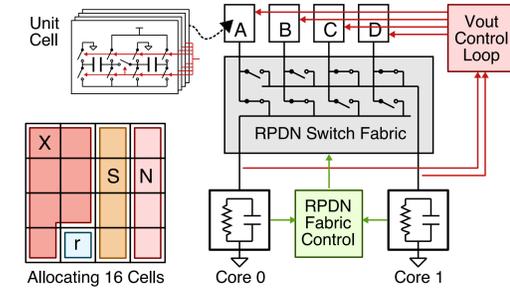
## 6 FGVS Circuit Design: RPDN

Shown in (a), we use a **single fixed-voltage regulator (SFVR)** as a baseline to compare against more sophisticated regulation schemes. We choose a configuration that can provide 80% efficiency at 1 V with an area of 0.26 mm<sup>2</sup> (4% of the core/L1 area). Shown in (b), **multiple adjustable voltage regulators (MAVR)** enable fine-grain voltage scaling in space and level. The power efficiency vs. area plot in (c) shows how we choose a per-core regulator area of 0.08 mm<sup>2</sup> to allow efficient voltage regulation for super-sprint. Note that designing for super-sprint significantly over-provisions for rest, nominal, and sprint modes; also, only one or two cores will ever be using the super-sprint mode at any given time.

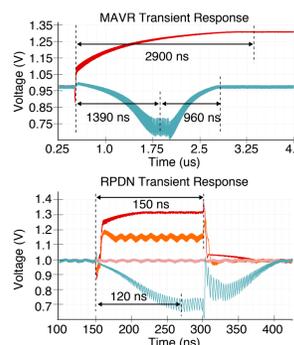


### Reconfigurable Power Distribution Networks

We propose a new approach called **reconfigurable power distribution networks** (RPDNs). As shown below, RPDNs include many small "unit cells" that each contain the flyback capacitance and regulator switches required for a SC regulator. These cells can be flexibly reconfigured through a switch fabric and combined with per-core control circuitry to effectively create multiple differently-sized SC regulators "on-demand" for cores. The inset shows how 16 unit cells can be allocated to four cores operating in four different modes.



The efficiency vs. power plot shown for a single core illustrates how lower-voltage modes require less area for high efficiency. A resting core only requires a single cell; RPDN would allow a super-sprinting core to "borrow" three cells from a resting core while maintaining high efficiency.

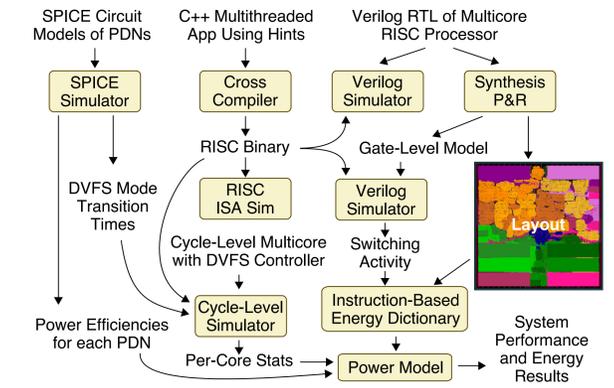


The RPDN architecture provides area savings of 40% over MAVR when supporting per-core supply regulation across the same number of cores. In addition to reducing area overhead, RPDN significantly reduces the voltage-settling response time. The waveforms shown here illustrate the difference in the transient responses for RPDN and MAVR when transitioning between modes. For RPDN, the response time for the nominal to super-sprint transition takes 150 ns while the same transition takes 2.9 μs with MAVR.

## 7 Evaluation

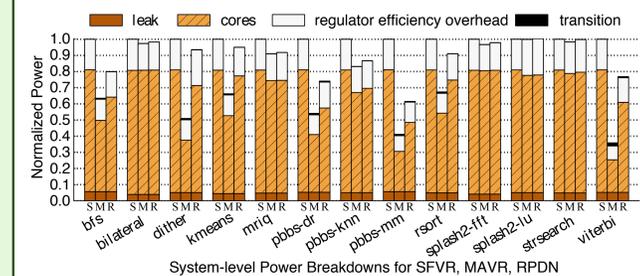
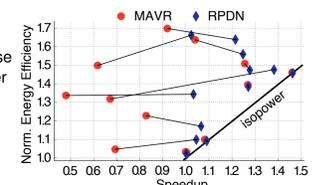
### Evaluation Methodology

We use a vertically integrated evaluation methodology that uses a mix of circuit-, gate-, register-transfer-, and architectural-level modeling.



### Evaluation Results

We evaluate the performance, energy efficiency, and power of applications on our target system with each PDN. We use a 4-level, 8-domain FG-SYNC+ controller and account for realistic voltage-settling response times and regulator power efficiencies in each DVFS mode for varying load currents.



Our promising results suggest that there is an important opportunity for architecture and circuit co-design of integrated voltage regulation in future systems.

## 8 Acknowledgments

This work was supported in part by NSF CAREER Award #1149464, a Spork Fellowship, and donations from Intel Corporation and Synopsys, Inc. The authors acknowledge and thank Ji Kim for his help in developing the instruction-based energy model, and Derek Lockhart, Shreesha Srinath, and Pol Rosello for their help in writing multicore application kernels.