RC25409 (WAT1309-012) September 9, 2013
Electrical Engineering

IBM Research Report

Distributed System of Digitally-Controlled Microregulators Enabling Per-Core DVFS for the POWER8™ Microprocessor

Zeynep Toprak-Deniz[1], Michael Sperling[2], John Bulzacchelli[1], Gregory Still[3], Ryan Kruse[4], Seongwon Kim[1], David Boerstler[4], Tilman Gloekler[5], Raphael Robertazzi[1], Kevin Stawiasz[1], Timothy Diemoz[2], George English[2], David Hui[2], Paul Muench[2], Joshua Friedrich[4]

[1]IBM Research Division
Thomas J. Watson Research Center
P.O. Box 218
Yorktown Heights, NY 10598
USA

[2]IBM Systems and Technology Group
Poughkeepsie, NY USA

[3]IBM Systems and Technology Group
Research Triangle Park, NC USA

[4]IBM Systems and Technology Group
Austin, TX USA

[5]IBM Systems and Technology Group
Boeblingen, Germany

Research Division
Almaden – Austin – Beijing – Cambridge – Dublin - Haifa – India – Melbourne - T.J. Watson – Tokyo - Zurich

Distributed System of Digitally-Controlled Microregulators Enabling Per-Core DVFS for the POWER8™ Microprocessor

Zeynep Toprak-Deniz[1], Michael Sperling[2], John Bulzacchelli[1], Gregory Still[3], Ryan Kruse[4], Seongwon Kim[1], David Boerstler[4], Tilman Gloekler[5], Raphael Robertazzi[1], Kevin Stawiasz[1], Timothy Diemoz[2], George English[2], David Hui[2], Paul Muench[2], Joshua Friedrich[4]

[1]IBM T. J. Watson Research Center, Yorktown Heights, NY

[2]IBM Systems and Technology Group (STG), Poughkeepsie, NY

[3]IBM STG, Research Triangle Park, NC

[4]IBM STG, Austin, TX

[5]IBM STG, Boeblingen, Germany

Integrated voltage regulator modules (iVRMs) [1] provide a cost-effective path to realizing per-core dynamic voltage and frequency scaling (DVFS), which can be used to optimize the performance of a power-constrained multi-core processor. This paper presents an iVRM system developed for the POWER8™ microprocessor which functions as a very fast, accurate low-dropout regulator (LDO) with 90.5% peak power efficiency (only 3.1% worse than an ideal LDO). At low output voltages, efficiency is reduced but still sufficient for beneficial energy savings with DVFS. Each iVRM features a bypass mode so that some of the cores can be operated at maximum performance with no regulator loss. With the iVRM area including the input decoupling capacitance (DCAP) but not the output DCAP inherent to the cores, the iVRMs achieve a power density of 34.5W/mm^2, which exceeds that of inductor-based or SC converters by at least 3.4X [2].

The POWER8™ microprocessor comprises 12 chiplets. Within each chiplet, the power grids for the logic supply (Vdd) and SRAM supply (Vcs) are divided into two regions – one for the main core (Vdd_core, Vcs_core) and one for the L3 cache (Vdd_cache, Vcs_cache); this allows the L3 cache to remain on (for data retention) while the main core

is power gated. The 48 regulated domains (4 per chiplet) are powered from 2 external supplies: Vdd_in for the Vdd_core and Vdd_cache domains, and Vcs_in for the Vcs_core and Vcs_cache domains. The power manager (PM), which programs the iVRMs to the desired voltage levels for DVFS, also controls the voltage levels of the external VRMs to maximize iVRM efficiency.

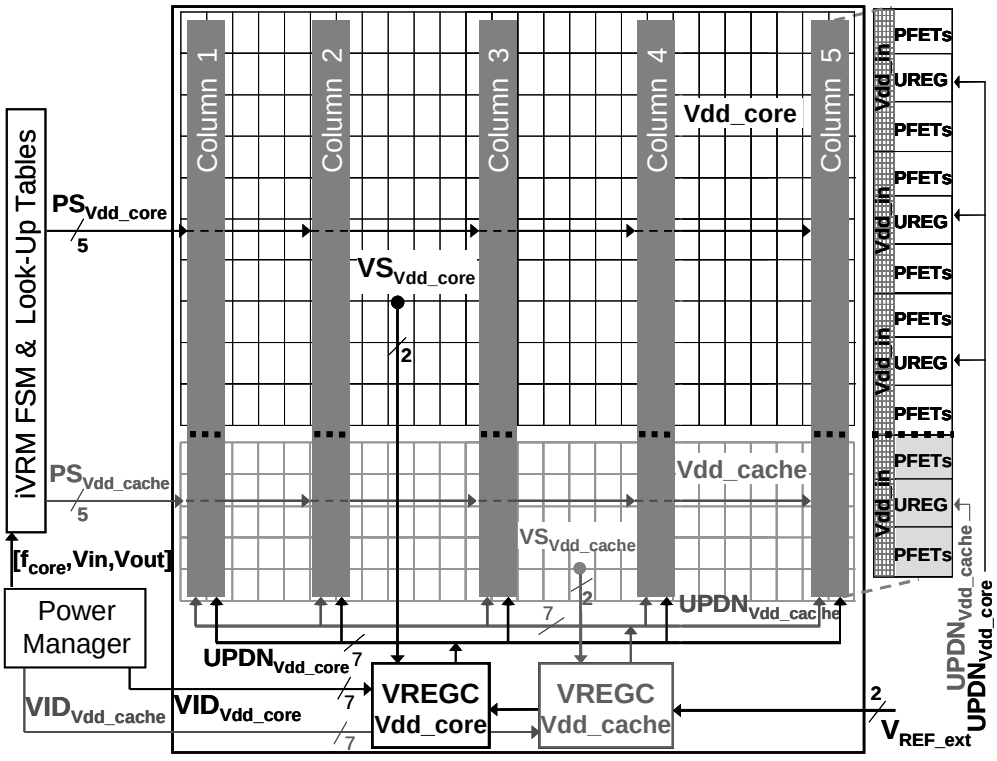

Figure 1. Distributed iVRMs for Vdd_core and Vdd_cache domains of a single chiplet.

Figure 1 shows the iVRM systems for the Vdd_core and Vdd_cache domains of one chiplet. The iVRM of each domain is implemented as a distributed system with a single voltage regulator controller (VREGC) governing the operation of multiple microregulators (UREGs). The input voltage grids, UREGs, and power headers (PFETs) are placed in 5 columns. The UREGs do not receive an accurate DC reference voltage; instead, they receive digital up/down correction signals from VREGC that affect the trip point of each UREG. VREGC compares the regulated voltage (e.g., Vdd_core) at a sense point (VS_{Vdd_core}) on the grid to a programmable voltage derived from a high-precision external reference (V_{REF_ext}) and feeds back a digital code ($UPDN_{Vdd\_core}$) to all the UREGs. This digital

distribution of up/down codes is more suitable for noisy processor environments than the analog distribution of up/down currents used in [3]. To optimize iVRM performance over a wide range of operating conditions, PMOS strength (PS) calibration is used to adjust the active width of the regulator passgate. A FSM employing look-up tables predictively calculates the optimum passgate width as a function of core frequency (f_{core}) and input and output voltages, an approach that is inherently much faster than the analog calibration loop of [3].

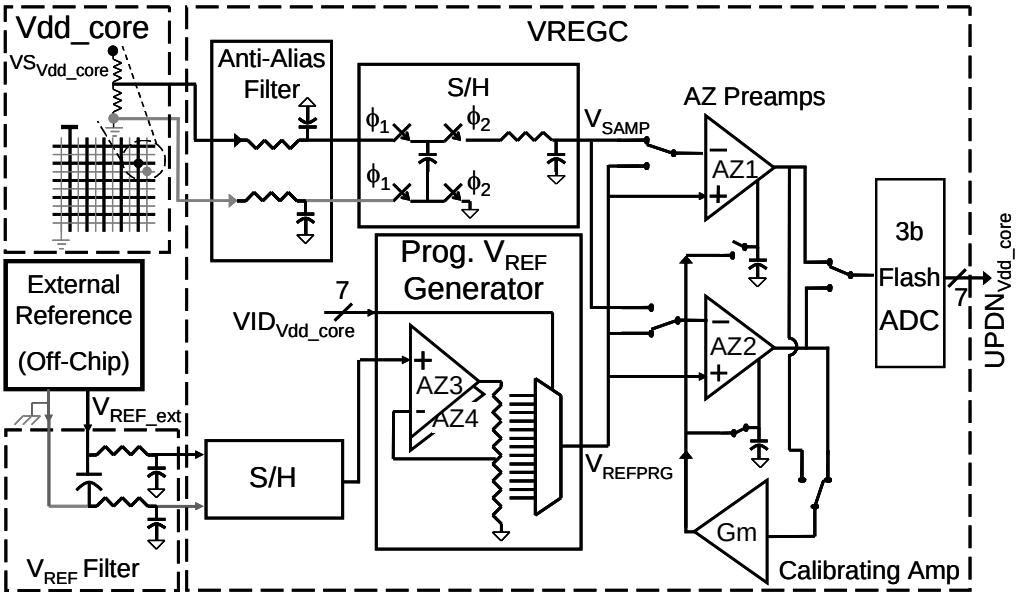

Figure 2. Block diagram of VREGC for Vdd_core domain.

Figure 2 shows the block diagram of VREGC. To avoid errors due to ground drops, the voltage at VS_{Vdd_core} is sampled differentially and converted to a single-ended signal V_{SAMP} (referenced to local ground) with a S/H. An RC filter in front of the S/H ensures that high-frequency ripple on Vdd_core is not aliased to a lower frequency inside the regulator control loop bandwidth. A similar S/H converts V_{REF_ext} to a single-ended signal, from which is generated a programmable reference level (V_{REFPRG}) set by a 7b code (VID_{Vdd_core}) from the PM. Vdd_core can be programmed with 6.25mV nominal resolution. A preamplifier senses the error between V_{SAMP} and $V_{REFPRG,}$ and its output is converted to a 7b thermometer code $UPDN_{Vdd\_core}$ with a 3b flash ADC. With the preamplifier, a +/-4mV error on Vdd_core drives the ADC to full scale. The auto-zeroed (AZ) preamplifier employs a ping-pong architecture in which

one amplifier (e.g., AZ1) is in use while the other (e.g., AZ2) is being offset compensated; the correction voltage is stored on a capacitor. Similar circuitry (not shown in figure) is used for offset compensation of amplifiers AZ3/AZ4.

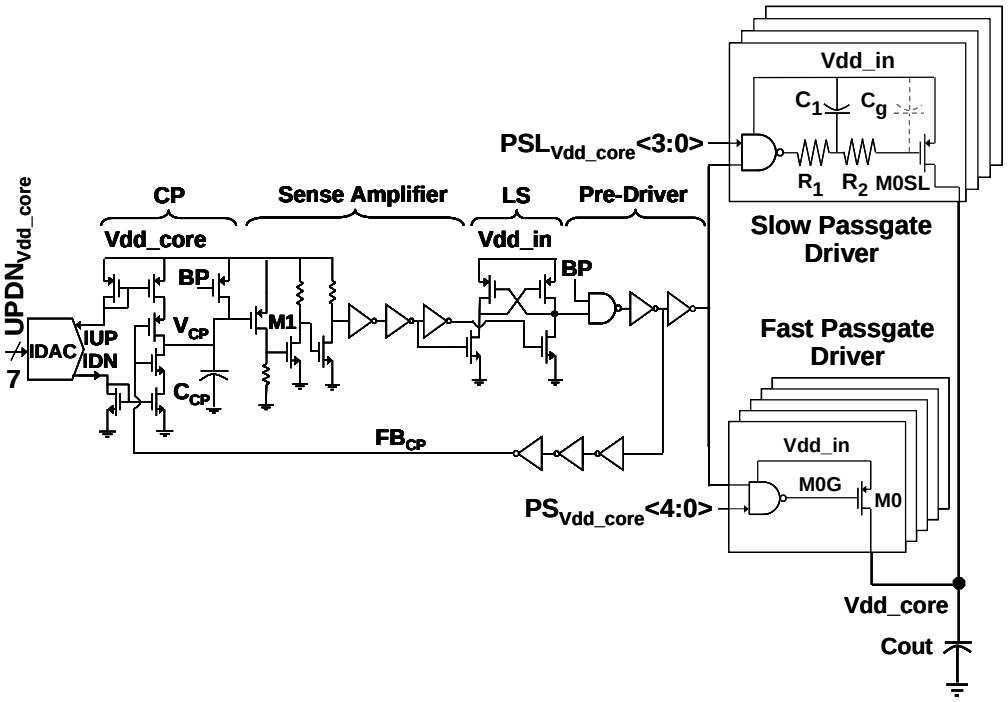

Figure 3. Simplified schematic of UREG.

The UREG (Fig. 3) features a comparator with sub-ns response time [3] which turns a PMOS passgate M0 on and off in a bang-bang fashion. The comparator trip point is tuned for high DC accuracy with a local charge pump (CP), whose output (V_{CP}) serves as a reference voltage for an error amplifier (common-gate stage M1). A current-steering IDAC converts the UPDN code from VREGC to IUP and IDN currents for the CP. If D is the duty cycle of M0 conduction, CP balance is achieved when IUP/IDN=D/(1-D). Since every UREG receives the same UPDN code, the UREG CP voltages are automatically adjusted to ensure equal duty cycles (balanced load sharing) even in the face of comparator offsets. The M1 stage output is amplified to rail-to-rail levels and then level-shifted (LS) to the Vdd_in domain. Driving the M0 gate capacitance with CMOS inverters and gates is power-efficient in modern processes [4]. For a UREG of this power level (≈40X greater than that in [3]), the power overhead of the sensing stages is negligible, which greatly increases current efficiency. Supplementing the fast switching passgate M0 with another

passgate M0SL, whose gate is not fully modulated, improves the tradeoff between self-generated ripple and current handling. The slower M0SL gate signal is generated locally within each UREG using a 2nd-order RC filter instead of being globally distributed as in [3]. PMOS strength calibration by the FSM further reduces self-generated ripple by adjusting the active widths of M0 and M0SL to handle the maximum load current without oversizing at strong corners. A binary-weighted code PS_{Vdd_core}<4:0> and a thermometer code PSL_{Vdd_core}<3:0> set the active widths of M0 and M0SL.

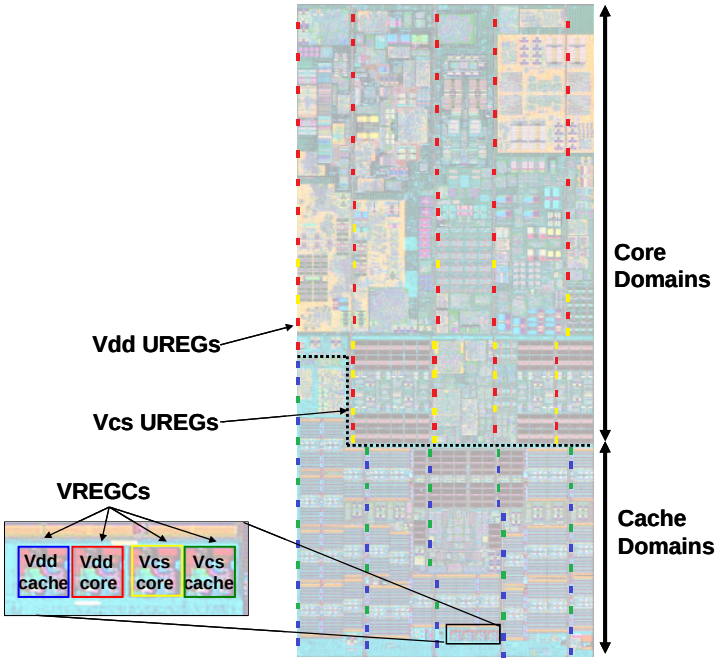

Figure 4. Micrograph of POWER8™ chiplet showing placement of regulator components for four different voltage domains.

The iVRMs were integrated into the POWER8™ chiplets (Fig. 4) and fabricated in a 22nm SOI CMOS process. The highest current iVRM (Vdd_core) uses 64 UREGs and 90nF of deep-trench (DT) input DCAP (shared with the Vdd_cache domain); these components and the Vdd_core VGREGC occupy about 1% of the chiplet area. The output DCAP (also DT) for this domain is 750nF. Figure 5 shows DC measurements of Vdd_core as a function of VID_{Vdd_core} with different loading conditions and values of Vdd_in. High loading is achieved both with custom test code intended to stress the current capacity of the iVRM and by raising f_{core} above its rated operating range. Low loading is achieved

by gating off the clocks of the core. With Vdd_in=1.1V and 0.61V≤Vdd_core≤1.05V, load regulation error is less than 3mV. With adequate headroom (Vdd_in-Vdd_core>50mV), absolute voltage error (Fig. 5(b)) is below 9mV, and the variation with Vdd_in is less than 5mV. Figure 6 shows the measured power efficiency as a function of Vdd_core (with high load). With Vdd_in=1.1V, the iVRM achieves peak power efficiency of 90.5% supplying 11.9A at Vdd_core=1.03V, at a power density of 34.5W/mm^2.

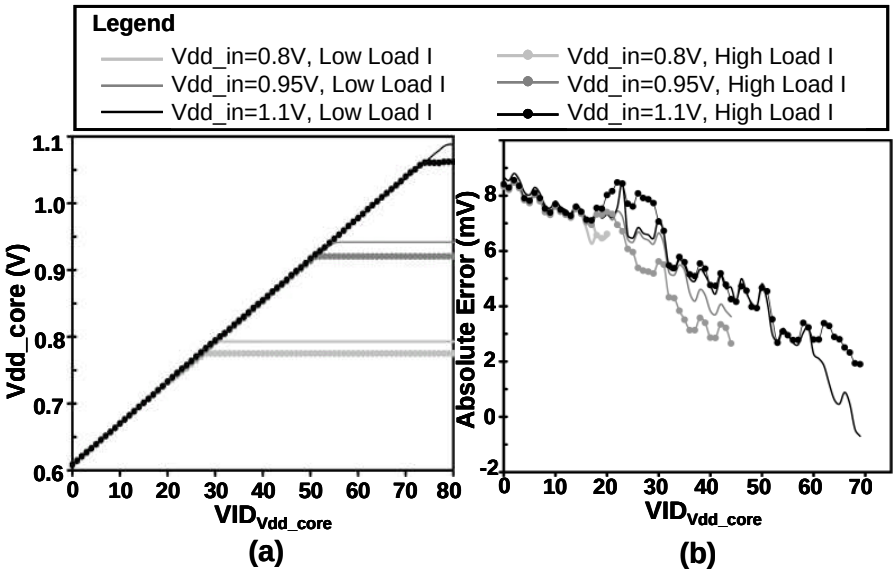

Figure 5. Measured (a) Vdd_core voltage and (b) deviation from its nominal value as function of VID_{Vdd_core}. Deviation from nominal value is plotted only for cases with Vdd_in-Vdd_core>50mV.

Dynamic tracking between Vdd_core and Vdd_cache is important to avoid the overhead of level shifters between domains. Since the output slew rate of each iVRM depends on loading, the reference voltages of the domains are moved in small steps slowly enough to ensure tracking. Figure 7 shows measurements of Vdd_core being moved up and down in 12.5mV steps. When Vdd_core is lowered, the PM decreases f_{core} with a DPLL before updating VID_{Vdd_core}. Because the PM must wait for the DPLL response, the downward movement is slower than the upward one. Finally, measurements of maximum core operating frequency (Fmax) show virtually identical results in bypass and regulated modes for the same values of Vdd_core, indicating that iVRM dynamic performance meets application requirements.

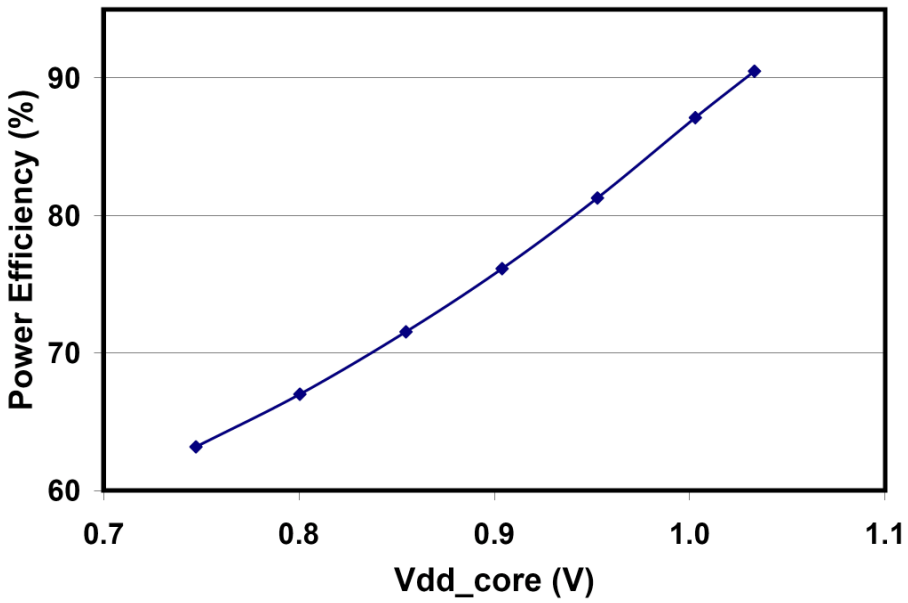

Figure 6. Measured power efficiency as function of regulated output voltage under high load conditions with Vdd_in=1.1V.

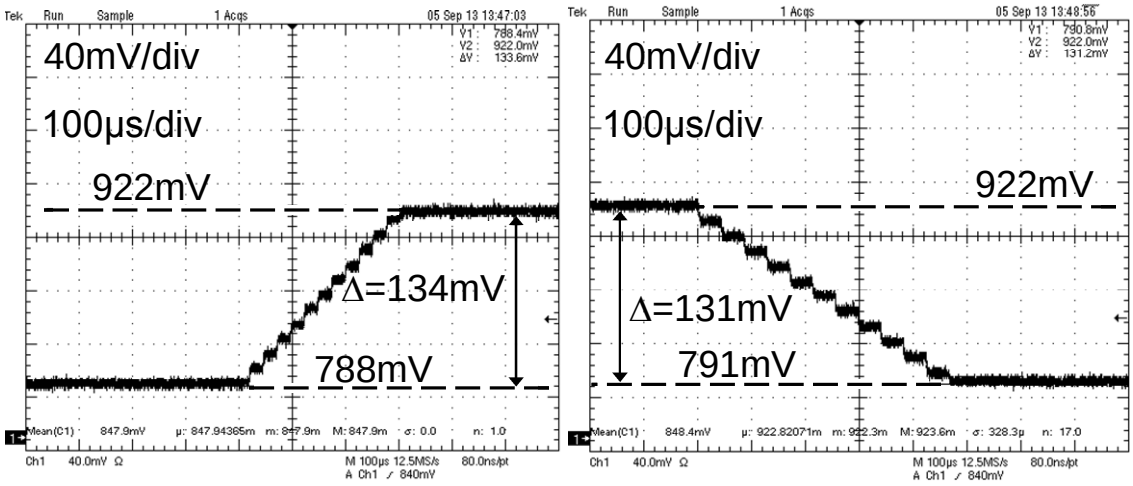

Figure 7. Measured Vdd_core voltage showing 12.5mV steps in a) upward and b) downward directions. At each step, PS_{Vdd_core} is updated by the FSM.

Acknowledgments:

The authors thank L. Acevedo and A. Wu for verification work and the IBM processor design and manufacturing teams for project support.

References:

[1] W. Kim, D. M. Brooks, and G.-Y. Wei, "A Fully-Integrated 3-Level DC/DC Converter for Nanosecond-Scale DVS with Fast Shunt Regulation," *ISSCC Dig. Tech. Papers*, pp. 268-269, Feb. 2011.

[2] S. R. Sanders et al., "The Road to Fully Integrated DC-DC Conversion via the Switched-Capacitor Approach," *IEEE Trans. Power Electron.*, vol. 28, pp. 4146-4155, Sept. 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 J. Solid-State Circuits*, vol. 47, pp. 863-874, Apr. 2012.

[4] P. Hazucha et al., "A Linear Regulator with Fast Digital Control for Biasing Integrated DC-DC Converters," *ISSCC Dig. Tech. Papers*, pp. 536-537, Feb. 2006.