Accurate Fine-Grained Processor Power Proxies

Wei Huang†, Charles Lefurgy*, William Kuk,‡, Alper Buyuktosunoglu*,**
Michael Floyd, Karthick Rajamani*, Malcolm Allen-Ware*, Bishop Brock****
†AMD, *IBM Research, ‡Purdue University, **IBM System and Technology Group
The 45th Annual IEEE/ACM International Symposium on Microarchitecture, 2012

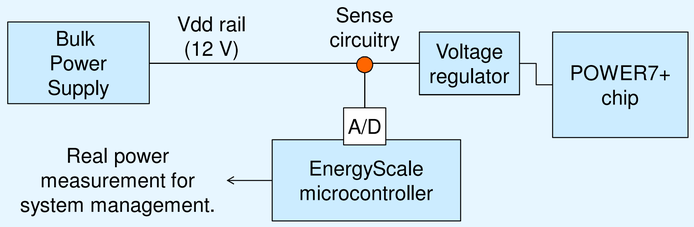
Situation

Practical ways to directly measure power consumption of a core within a microprocessor do not exist.

Servers commonly measure power consumption of chips at the voltage regulator. Power management could be improved by providing accurate, fine-grain power measurements for individual processor cores.

Core power measurement would enable virtual machine energy billing and forecasting the effects of power management actuators.

Conventional power measurement of a chip voltage rail



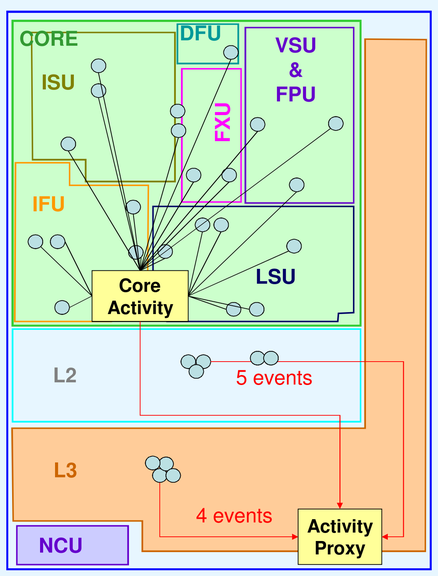
Opportunity

POWER7 includes on-chip hardware to compute per-chiplet *activity proxies* used to estimate active power.

Activity counters and the calculation of *activity proxies* are implemented in hardware logic of each core. Instead of implementing a weighted sum, some weights are applied to groups of activity counters to reduce circuit area.

$$ActivityProxy = \sum\left(W_g \times \sum\left(W_{i_g} \times A_{i_g}\right)\right)$$

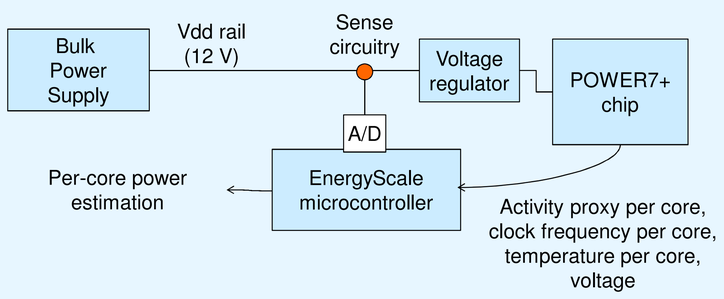
POWER7 Chiplet (core + L2 + L3)



○ = Activity Sense point

The weights to different activity events are programmable by writing to special on-chip registers. The EnergyScale microcontroller receives the activity proxies and adjusts them to account for the effects of leakage, temperature, process variations and voltage to form chip and core *power proxies*.

Estimation of core power



Solution

Determine idle power model.

On idle chip, sweep voltage and frequency (253 measurement points)

↓

Measure power on Vdd rail and measure chip temperature

↓ <Power,Voltage,Frequency, Temp> x 253 x 4 chips

Genetic Algorithm-based Optimization

Find fitting parameters to minimize $P_{measure} - P_{idle}$ for all measurement points

↓ Fitting parameters $S_0, \beta, P_{leak\_nom}, \gamma, m_0$

Idle power model

$$P_{idle} = P_{clock} + P_{leak}$$
$$= \frac{F}{S_0}\left(\frac{V}{V_{nom0}}\right)^{\beta} + P_{leak\_nom}\left(\frac{V}{V_{nom}}\right)^{\gamma}\left(1 + m_0(T - T_0)\right)$$

- Idle power model has accuracy of 3% across voltage and frequency range.
- Fit using 4 chips from distinct process corners.

Determine active power model.

Run training kernel workload (V=Vnom, F=Fnom)

↓

Measure activity counters
Measure power on Vdd rail
$P_{active} = P_{measured} - P_{idle}(V,F,T)$

↓ Counters and active power A_{i_g}, P_{active}

Genetic Algorithm-based Optimization

Find weights to minimize P_{active} - ActivityProxy/R

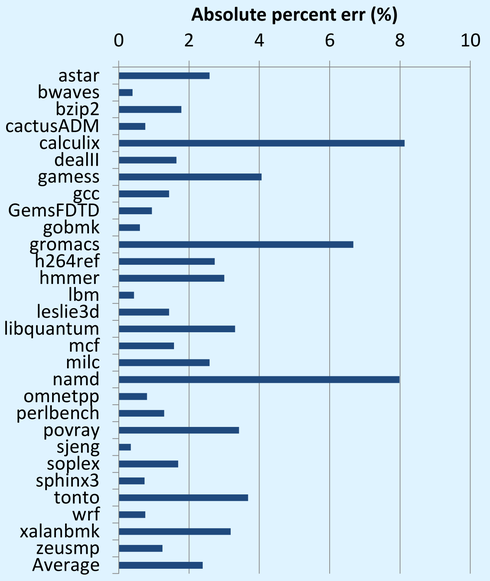
↓ 51 weights W_g, W_{i_g}

Active power model

$$P_{active} = \frac{ActivityProxy}{R_0}\left(\frac{V}{V_{nom0}}\right)^{\alpha}$$
$$ActivityProxy = \sum\left(W_g \times \sum\left(W_{i_g} \times A_{i_g}\right)\right)$$

- Trained with 762 kernels, spanning a range of memory sizes and threading modes.

Validation: Pchip measurement matches Pidle + Pactive models.



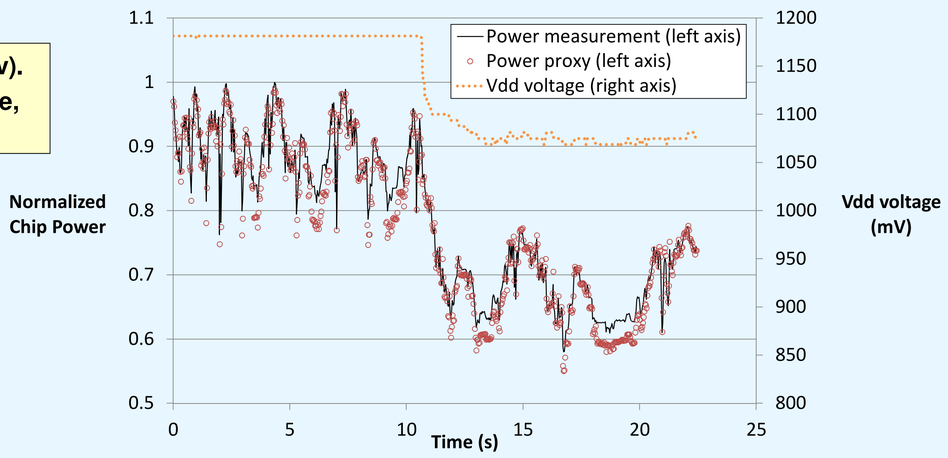
- Unsigned error is 1.8% (2.0% std dev) across all tested workloads (only SPEC CPU2006 shown).

Results

Chip Vdd power proxy has a mean error of 0.2% (2.6% std dev).
Power proxy tracks change in voltage, frequency, temperature, and workload activity.

Experiment
- Calibrate POWER7+ power proxy hardware.
- Run workloads (SPEC CPU2006, SPECpower_ssj2008, etc.).
- Measure power of Vdd voltage rail.

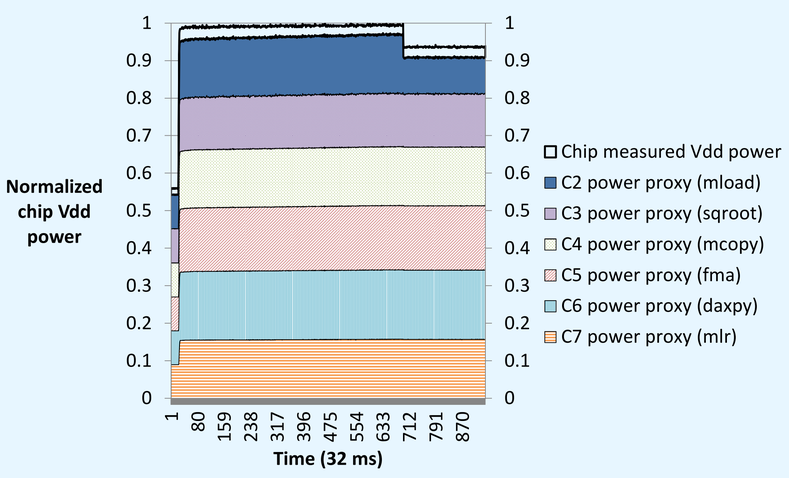
Observations
- Tracks changes in voltage, frequency, temperature, and workload activity
- Power estimation made every 32 ms (30x faster than prior work)
- Power proxy is accurate even when voltage and frequency do not have fixed pairings. Useful for undervolting (with fixed frequency) and overclocking (with fixed voltage) scenarios.
- Power proxy implementation on service processor and system management network does not impact workload performance.



- Fixed frequency run of dealII workload.
- Power proxy continues to track actual power while undervolting up to 112.5 mV.

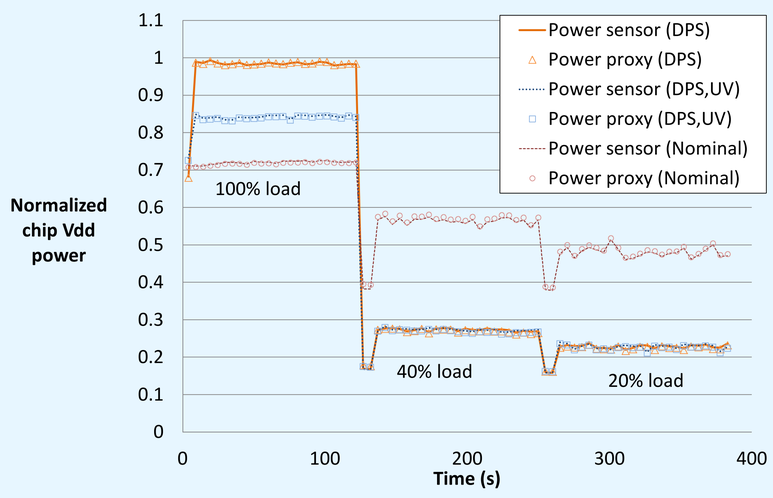
Applications

Enable billing of energy consumption for virtual machines on a per-core basis.



- Run 6 workloads on 6-core chip and compare to real chip Vdd power (3% err).
- Power proxy tracks thermal rise.

Improve power management controllers by forecasting power due to change in voltage, frequency, temperature, and workload.



- SPECPower_ssj2008 run under different conditions
  UV = undervolting, DPS = Dynamic Power Saver (voltage and frequency scaling).