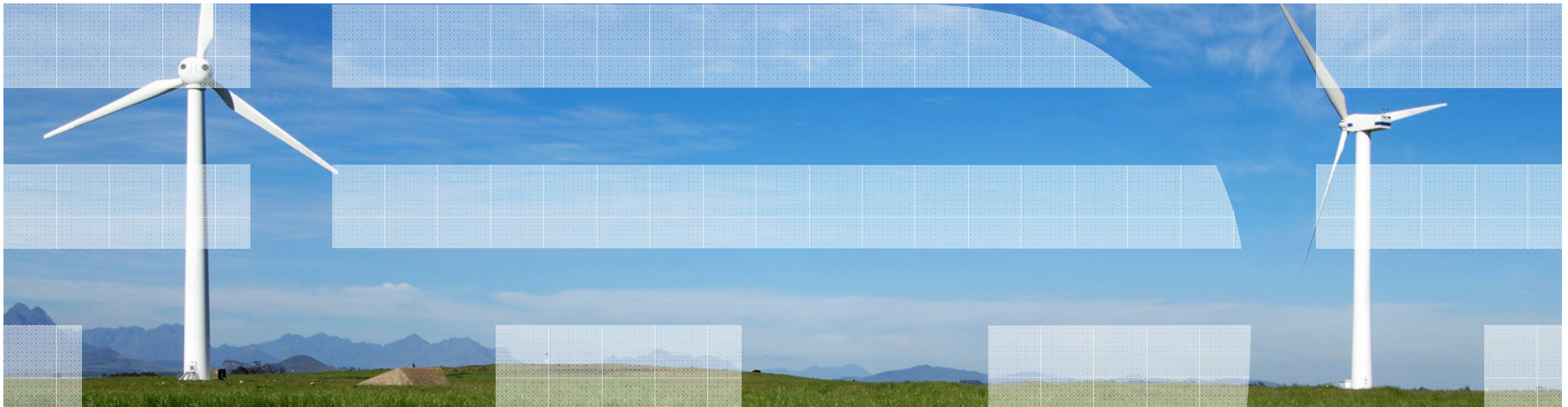


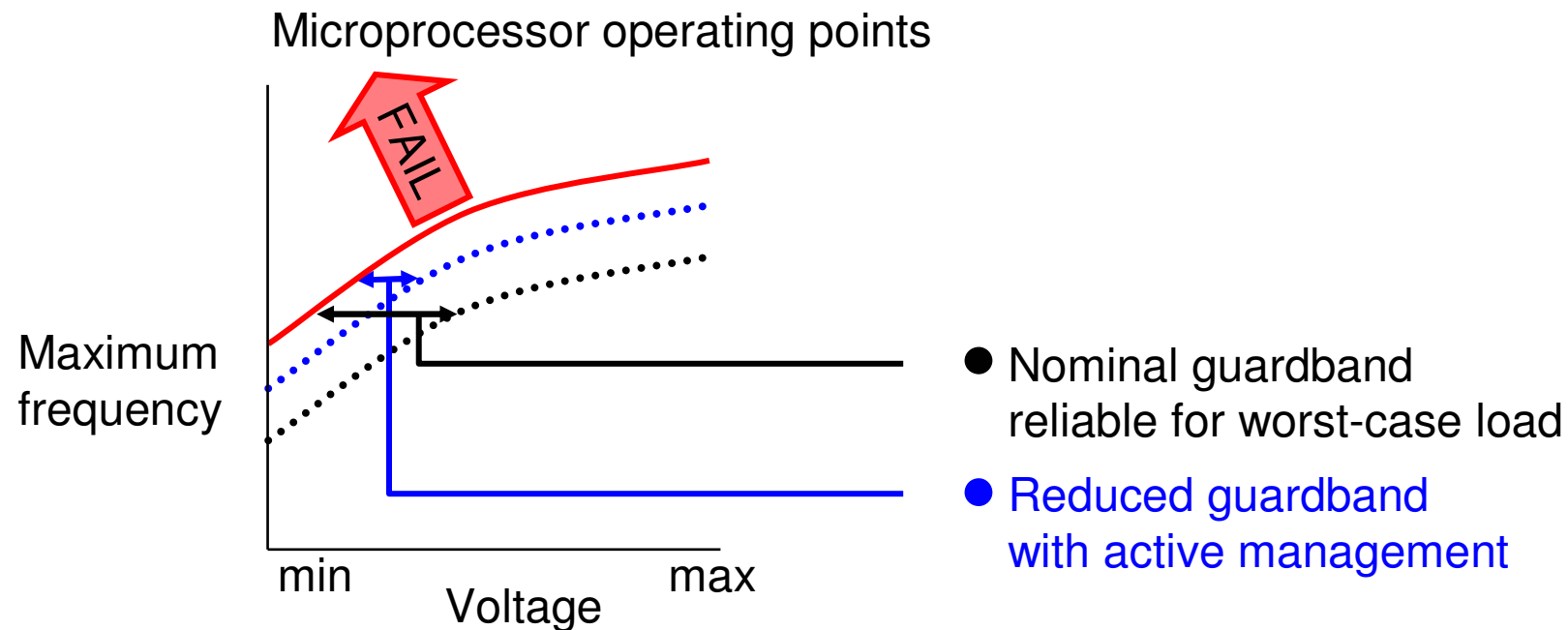
Active Management of Timing Guardband to Save Energy in POWER7

Charles Lefurgy, Alan Drake, Michael Floyd, Malcolm Allen-Ware,
Bishop Brock, Jose Tierno, and John Carter



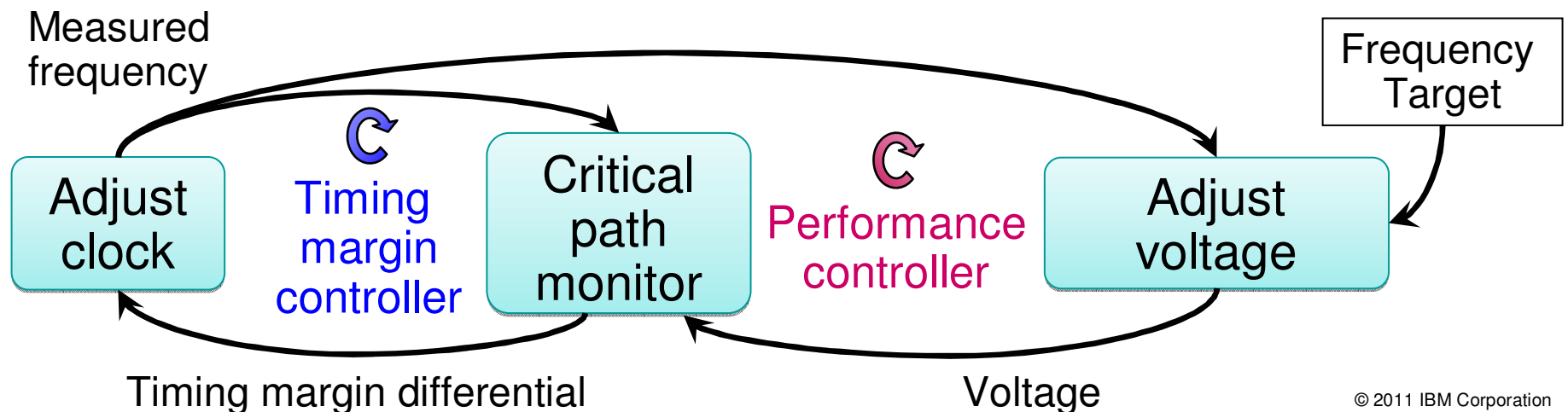
Excess guardband

- The voltage used on a microprocessor is conservative to provide a safe timing margin under worst-case conditions
 - workload-induced voltage droops (dl/dt or load line)
 - high temperature
- **Concern:** Energy-efficiency is reduced to guarantee reliability.
- **Opportunity:** Worst-case conditions rarely occur. Can actual timing margin be controlled?



Our solution

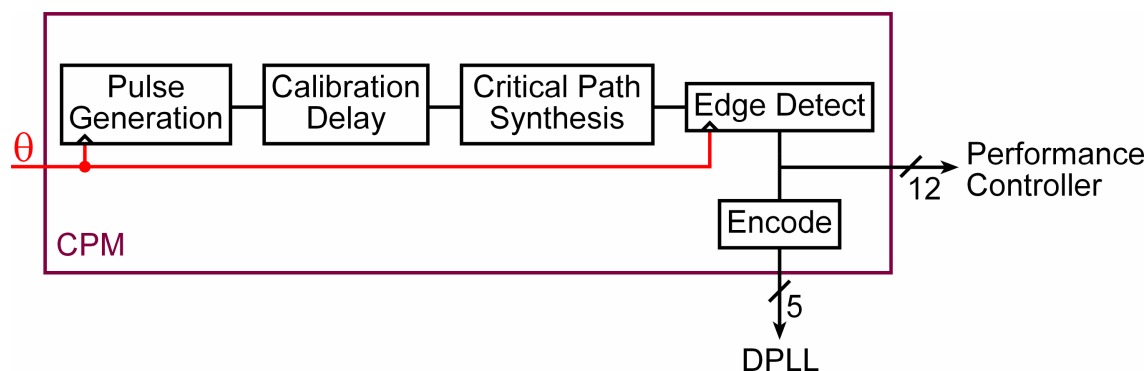
- New capability to keep timing margin nearly constant
 - Convert excess timing margin into a voltage reduction
 - Reduce traditional voltage margin when conditions are not worst-case (Some voltage margin is retained for aging, calibration inaccuracy, etc.)
- 1. **Measure** excess operational margin with timing margin sensor
 - Difference from a calibrated reference point
- 2. **Protect** timing margin against voltage droop by adjusting frequency
 - Hardware-based **timing margin controller**
- 3. **Save energy** by converting excess timing margin into voltage reduction
 - Software-based **performance controller**



Measure timing margin

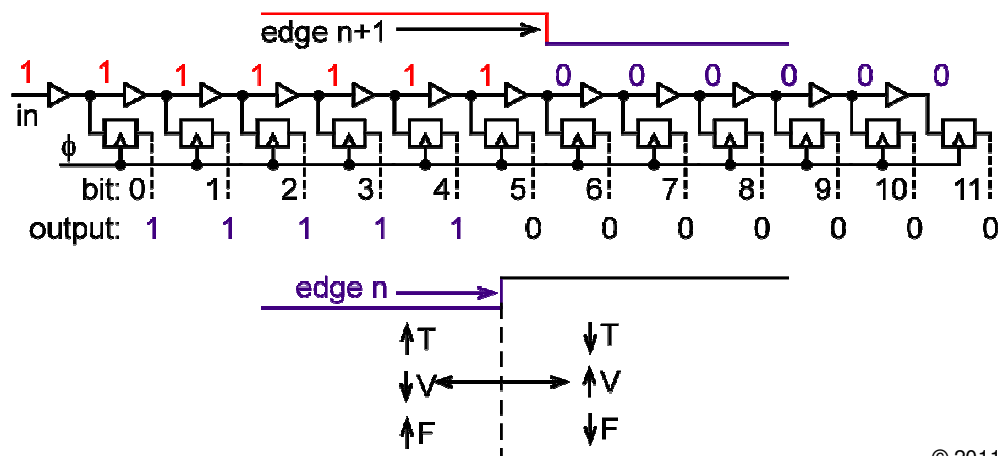
- Use **Critical Path Monitor** (CPM) circuit. Mimics behavior of real critical path.
- Each cycle: generate pulse, traverse synthesized critical path and calibrated delay, capture in edge detector

Critical Path Monitor



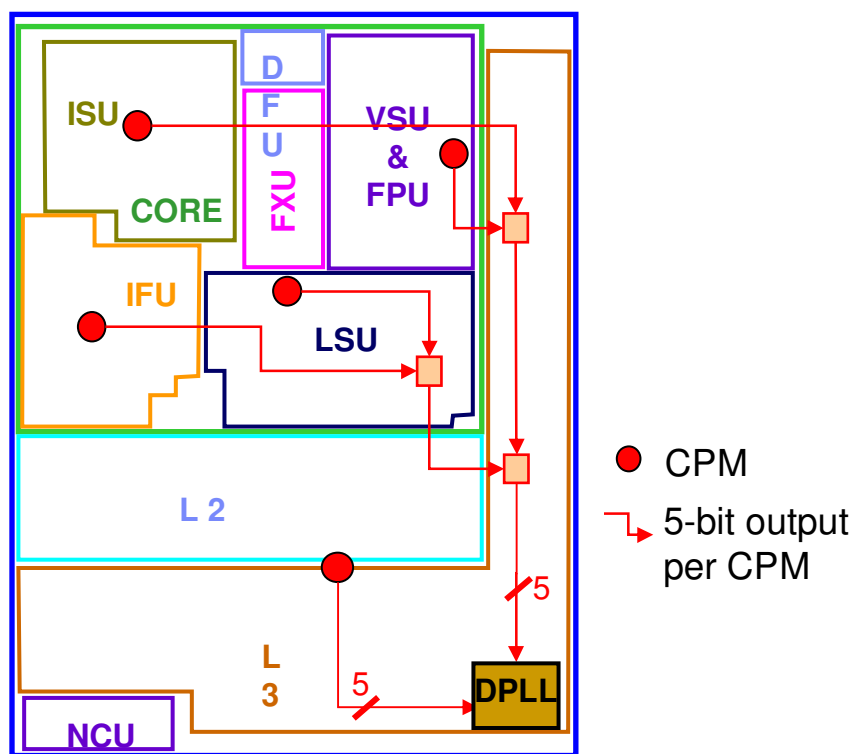
- Edge detector 12-bit output: (bit 0 = less margin, bit 11 = more margin)

Edge Detector



Critical path monitor

- 5 Critical Path Monitors per core in POWER7 (8 core chip)
- Middle bits of edge detector are forwarded to DPLL



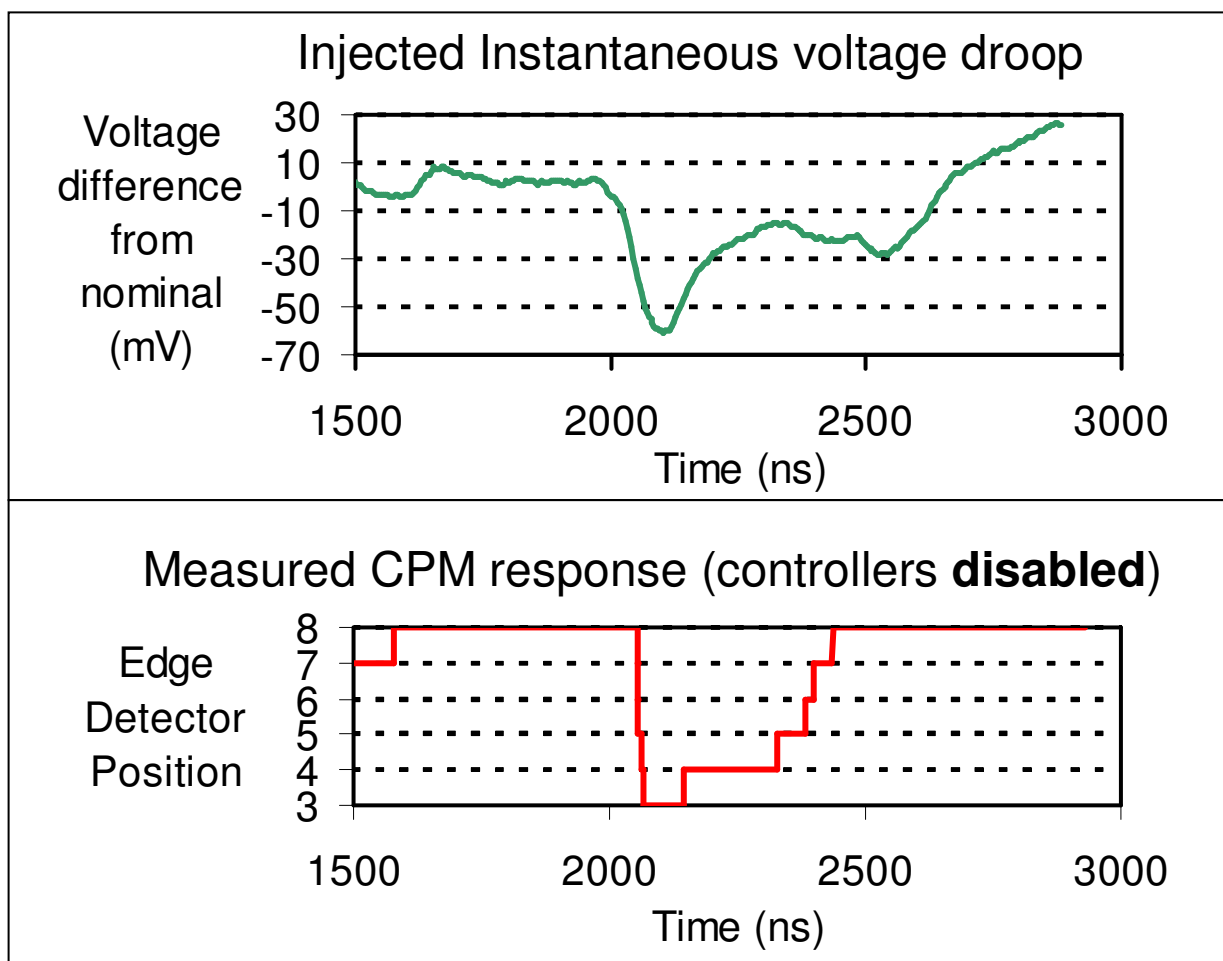
POWER7 core chiplet

CPM output

“11111” = large margin
 “11110” = some margin
 “11100” = ideal margin
 “11000” = margin too small
 “10000” = not enough margin

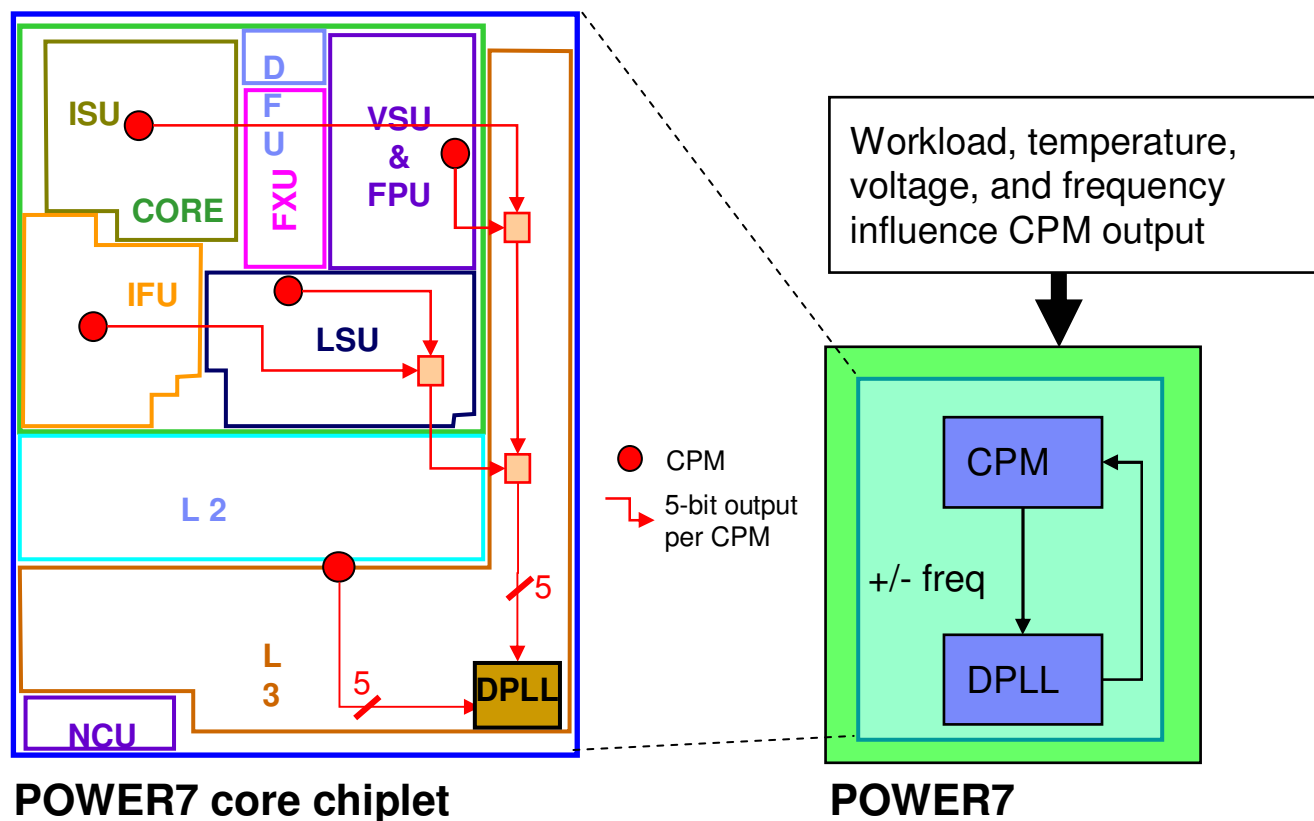
Example of critical path monitor output

- Inject 60 mV droop into Power 755 Express Server (with no load-line)
 - Instruction fetch throttling
- Critical path sensor follows on-chip voltage reduction



Protect timing margin

- **Timing margin controller** responds to changing operating conditions by adjusting frequency to maintain timing margin target.
 - Implemented in hardware of POWER7.
 - Can reduce frequency by -7% in about 5 ns to handle fast voltage droop.



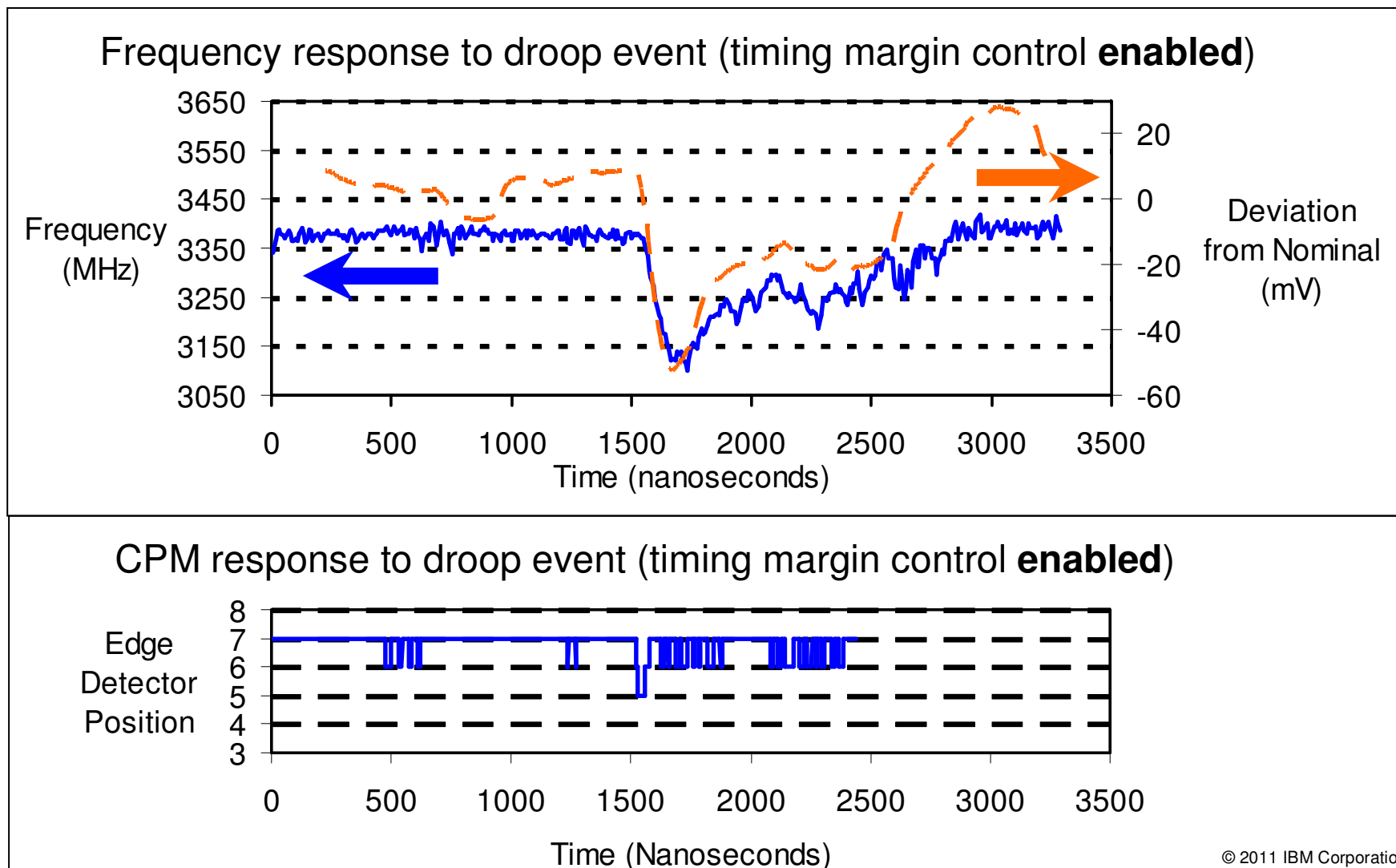
Calibration of critical path

- Teach the chip the desired timing margin to use during field operation
- Done once during manufacturing of chip
- Run chip at desired timing margin
 - Set voltage, frequency, and temperature
 - Run stressful workload
- Find delay setting that places timing edge on position 6 in edge detector
 - Position 6 is the setpoint for the timing margin controller

- Calibrations used in our study
 - “100% guardband” – original product
 - “50% guardband” – removes roughly 50% of the guardband
 - “0% guardband” – removes nearly all guardband

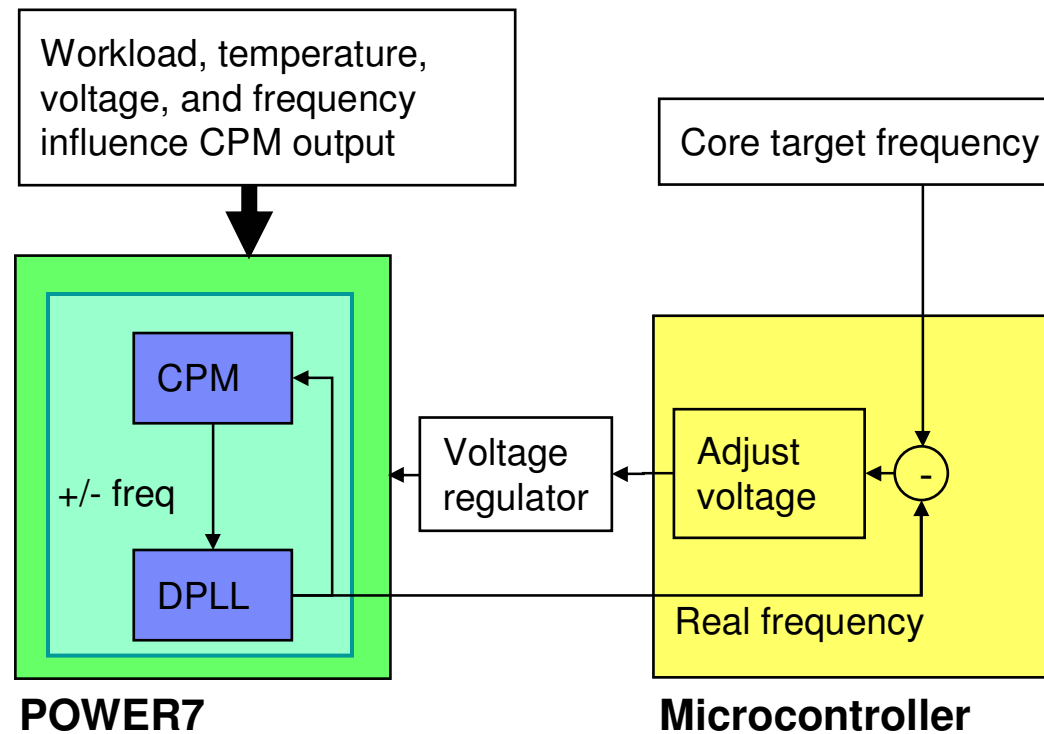
Timing margin controller response time

- Quick enough to follow voltage droops

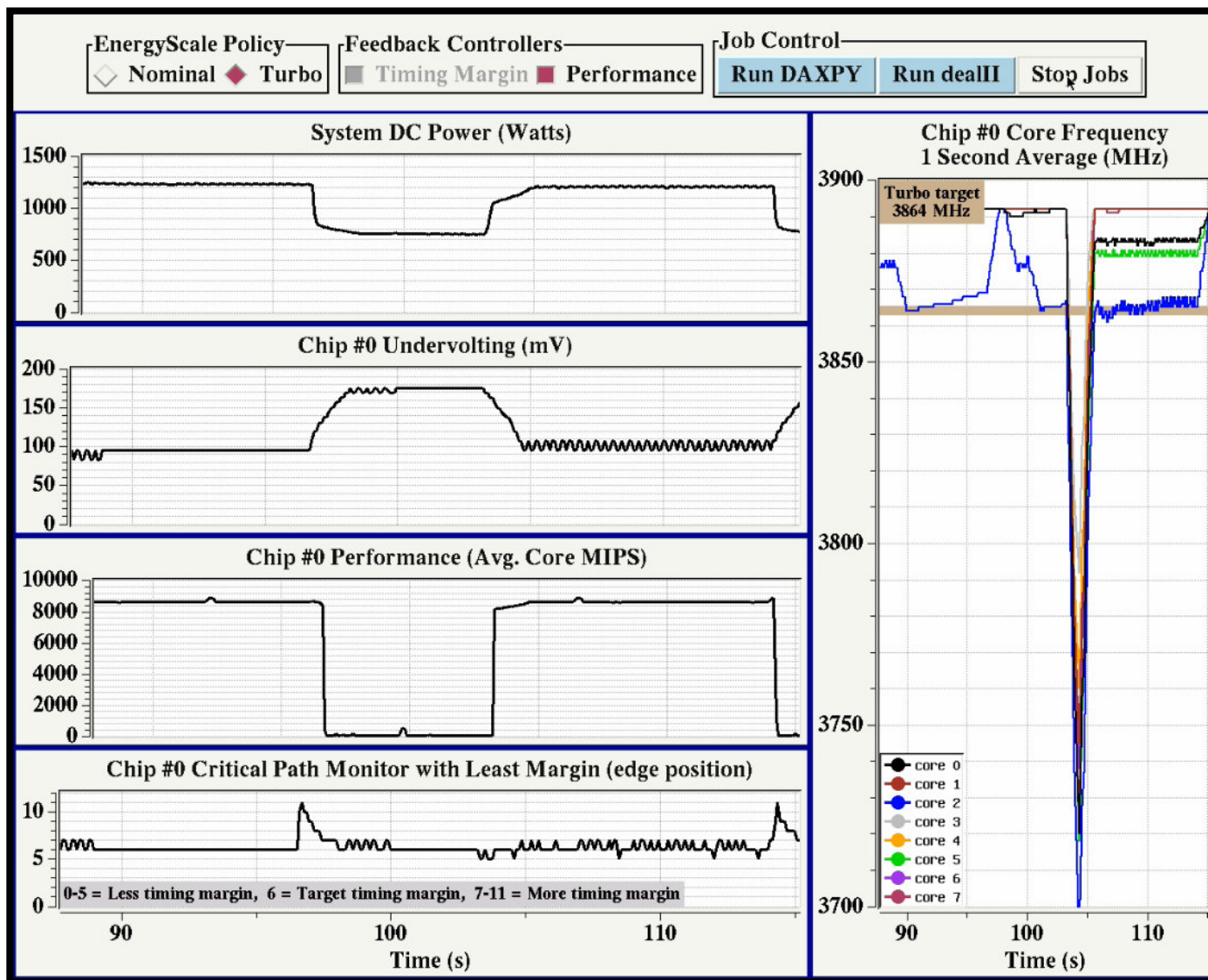


Save energy

- **Performance controller** adjusts voltage to meet desired clock frequency target.
 - Implemented in firmware of on-board microcontroller
 - Frequency is capped at target + 28 MHz (clock resolution)
 - Prevent energy waste
 - Allow for detection of excess timing margin for voltage reduction



Demonstration



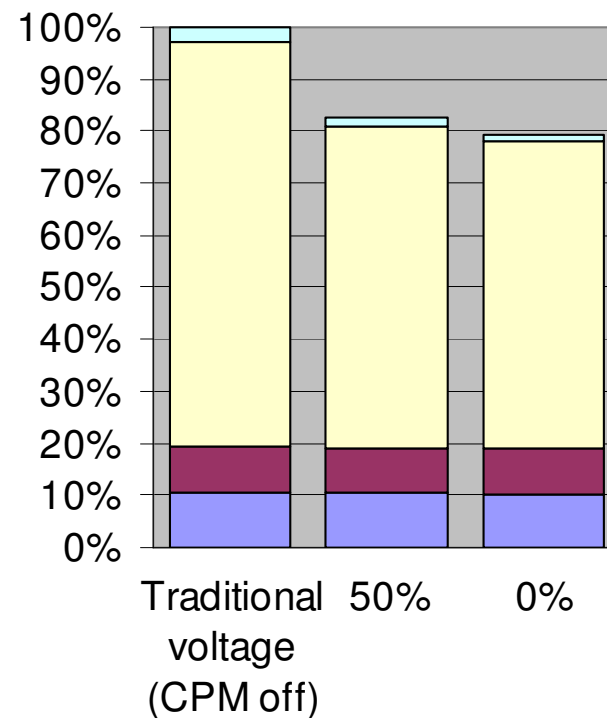
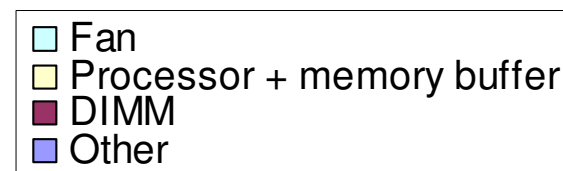
Results

- Implement in Power 750 Express Server (4 POWER7 processors, 64 GB)
 - Run SPEC CPU 2006 workloads
 - Frequency target = 3864 MHz (Turbo)

- Guardband reduced to 50%
 - 20% chip power reduction
 - Voltage reduced 113 mV – 140 mV
 - 18% system power reduction
 - 50% fan power reduction
 - No change in performance

- Upper bound on power reduction (0% guardband)
 - 24% chip power reduction
 - 21% system power reduction

Normalized System DC Power



Guardband

Related work

- Razor [Ernst et al., MICRO-36, 2003]
 - Instrument actual critical paths
 - Compute both speculative and known-good values
 - Rollback and replay instructions when timing error detected
 - Allows safe removal of all timing margin

	Razor	Active Guardband Management in POWER7
Area	Less than 3% of chip [Das, 2009]	0.2% of core
Power reduction	33-50% [Das, 2009]	Up to 24%
Performance	1-3% overhead from roll-back recovery [Ernst, MICRO 2003]	No change
Design and Verification Effort	More (pervasive)	Less (localized)

Conclusions

- Demonstration of a new capability to keep timing margin nearly constant
- Architecture combines two feedback controllers
 - Hardware-based timing margin controller (safety)
 - Software-based performance controller (undervolting)
- Reduced average chip power by 20% for SPEC CPU2006
- Working prototype in POWER7 server
- Commercially viable

Acknowledgements

- Philip Restle
Alexander Rylyakov
Daniel Friedman
Daniel Beece
(IBM Watson Research Lab)

Groundwork for the CPM-clock feedback control loop and extensive modeling to validate the CPM-DPLL feedback implemented in the POWER7 chip.

- Richard Willaman (IBM Austin POWER Systems Lab)

Collected the instantaneous droop scope data.

- This work was supported in part by the Defense Advanced Research Projects Agency under contract #HR0011-07-9-0002.

