

VR-Scale: Runtime Dynamic Phase Scaling of Processor VRs for Improving Power Efficiency

**Hadi Asghari-Moghaddam, Hamid Reza Ghasemi,
Abhishek Arvind Sinkar, Indrani Paul, Nam Sung Kim**



ILLINOIS
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

DAC-53
July 9th, 2015

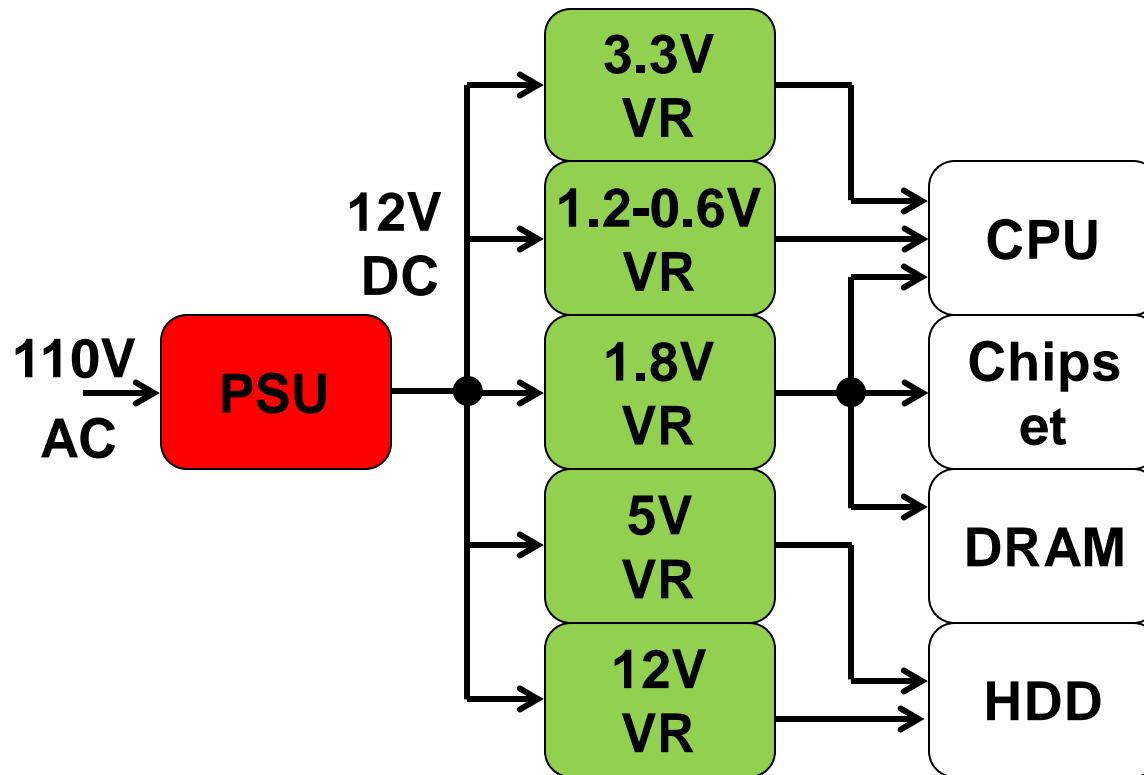
-
- Background
 - P-state and C-state
 - System power delivery
 - Multi-phase VRs
 - VR efficiency
 - Benchmarks behavior analysis
 - P-state
 - C-state
 - Current
 - VR-Scale
 - Conclusion

- C state: low-power sleep state
 - ⊙ Deeper C states consume lower standby power at higher wake-up penalty
 - ⊙ C0 (idle), C1 (halt), C3 (sleep), and C6 (off) states

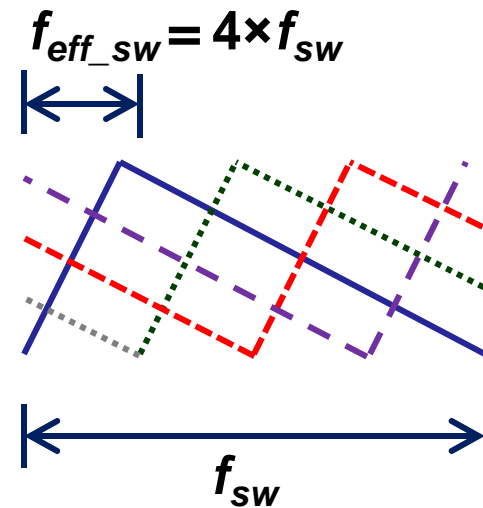
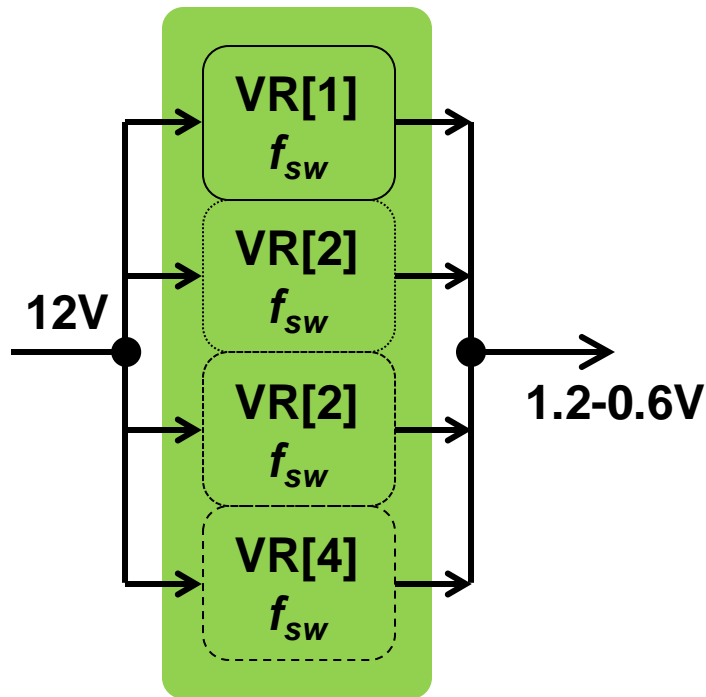
- P state: dynamic voltage/frequency (V/F) state
 - ⊙ Deeper P states operate cores at lower V/F and thus consume lower active power at longer runtime
 - P0 (3.5GHz) – P19 (1.6GHz) for our evaluation system.

Platform Power Delivery Architecture 3

- Hierarchical power delivery to maximize efficiency
 - 110V AC to 12V DC
 - 12V DC to various levels of DC using multiple VRs



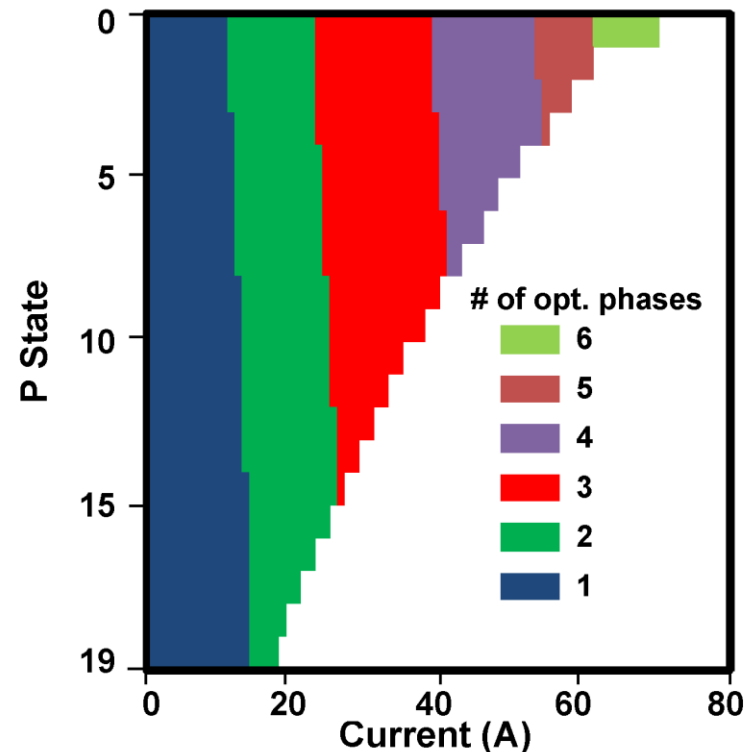
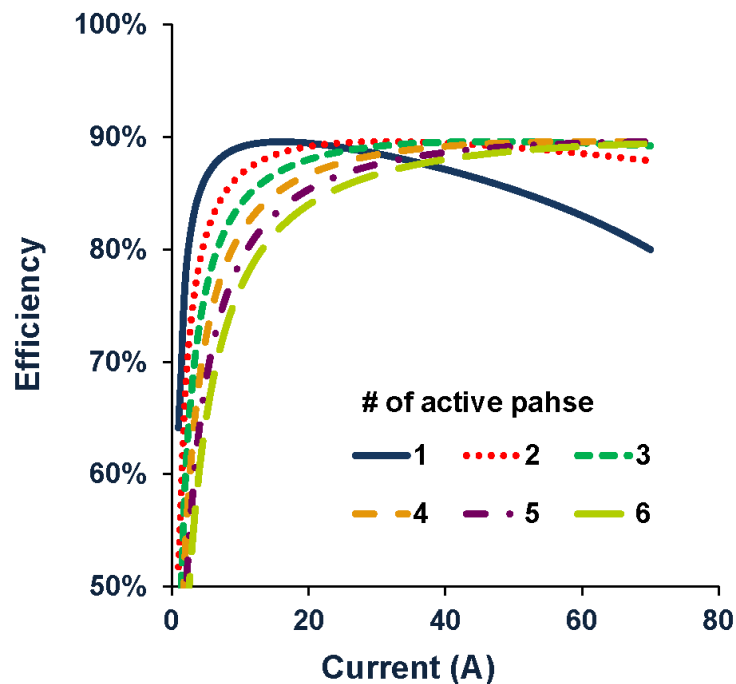
- Comprised of N phases
 - Each of N phases is turned on at $1/f_{sw}$ interval.
 - Lowers switching loss
- High-performance processors use as many as 6-8 phases.



VR Efficiency vs Load Current

5

- Function of number of active phases, current, and voltage
 - 5A load current 64% w/ 6 phases vs 86% w/ 1 phase at 1.2V
- Optimal # of active phases
 - Vary as a function of voltage (P state) and current
 - Improve power efficiency 30-50% at low load current



- System configuration

CPU	Intel i7-3770K	Technology	22nm
Microarchitecture	Ivy Bridge	# of cores/threads	4/8
Turbo freq	3.9GHz~3.6GHz	P-state freq	3.5GHz~1.6GHz
Max TDP	77W	Memory	DDR3-1600/8GB

- Workload

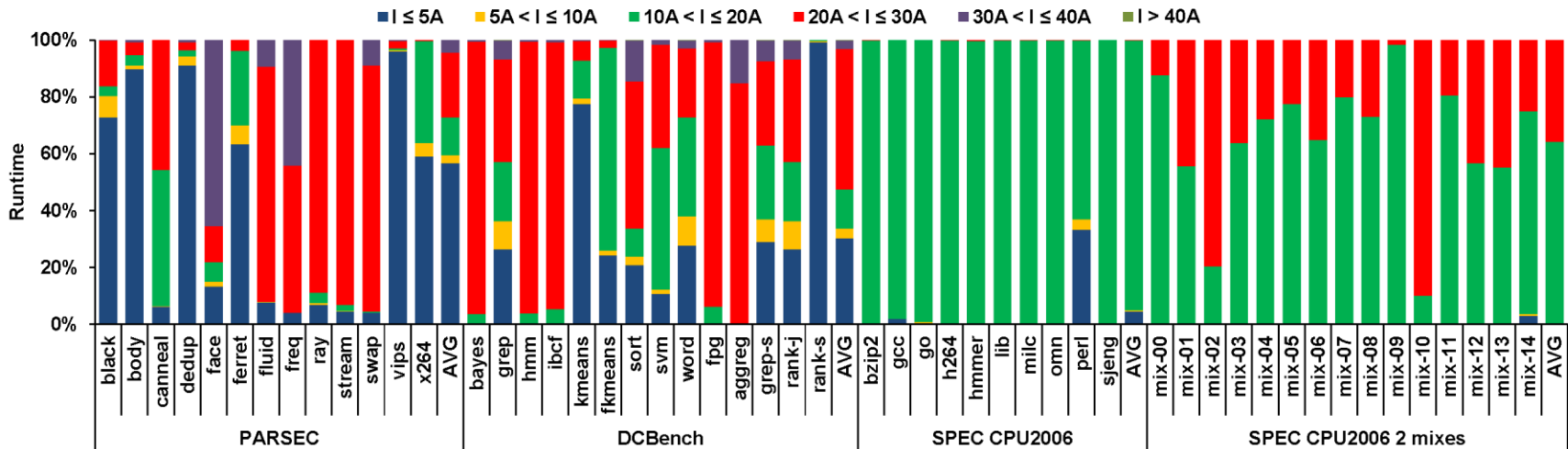
- PARSEC w/ native input set
- SPEC CPU2006 (one, two, and four mixes)
- DCBench for classification algorithms processing big data

- Energy, frequency, and VID

- Obtained through processor MSR's
- 1ms sampling interval

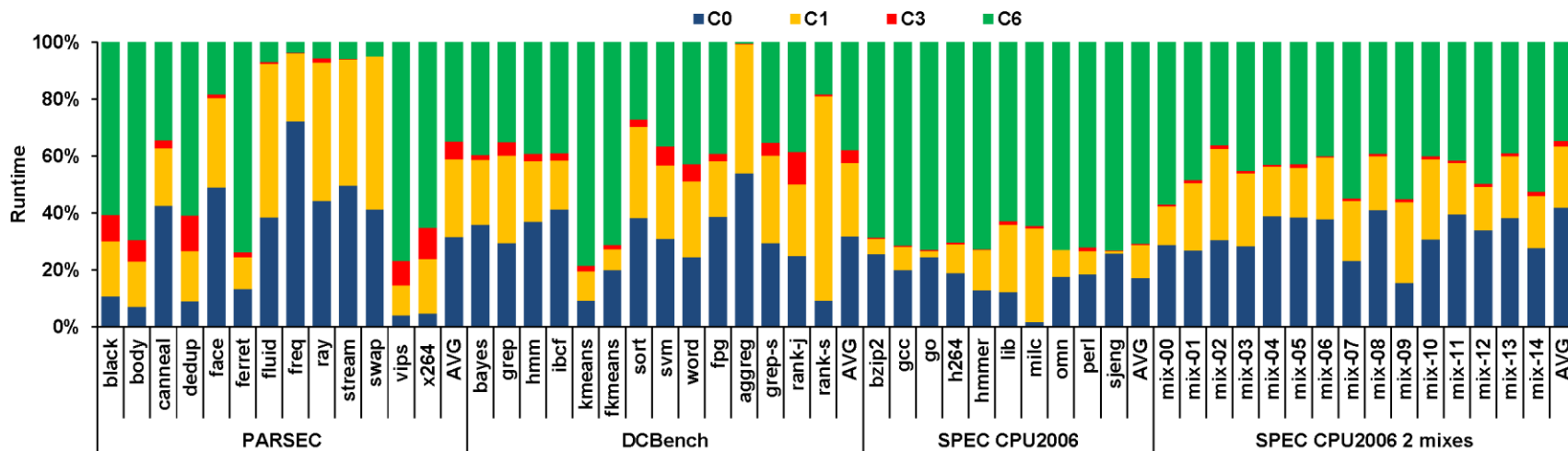
Load Current Distribution

- PARSEC
 - Less than 5A (30A) for more than 55% (95%) runtime
- DCBench
 - Less than 5A (30A) for more than 30% (95%) runtime
- SPEC(1)
 - Less than 15A for more than 90% runtime



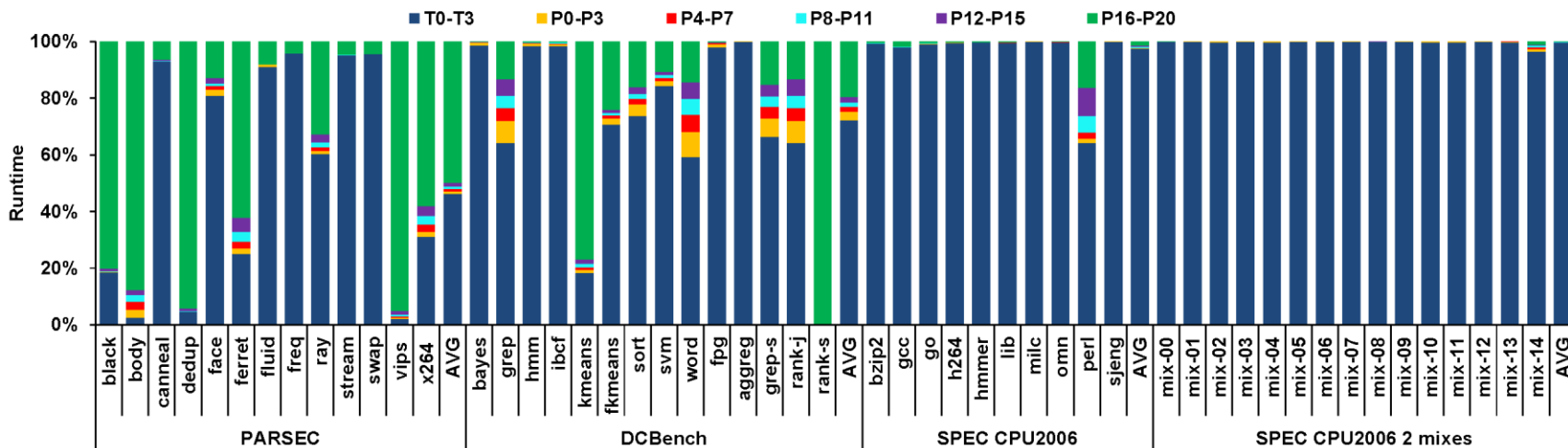
C-State Residency

- PARSEC
 - More than 68% runtime in C1/C3/C6 sleep states
- DCBench
 - More than 65% runtime in C1/C3/C6 sleep states
- SPEC
 - More than 71% runtime in C6 states

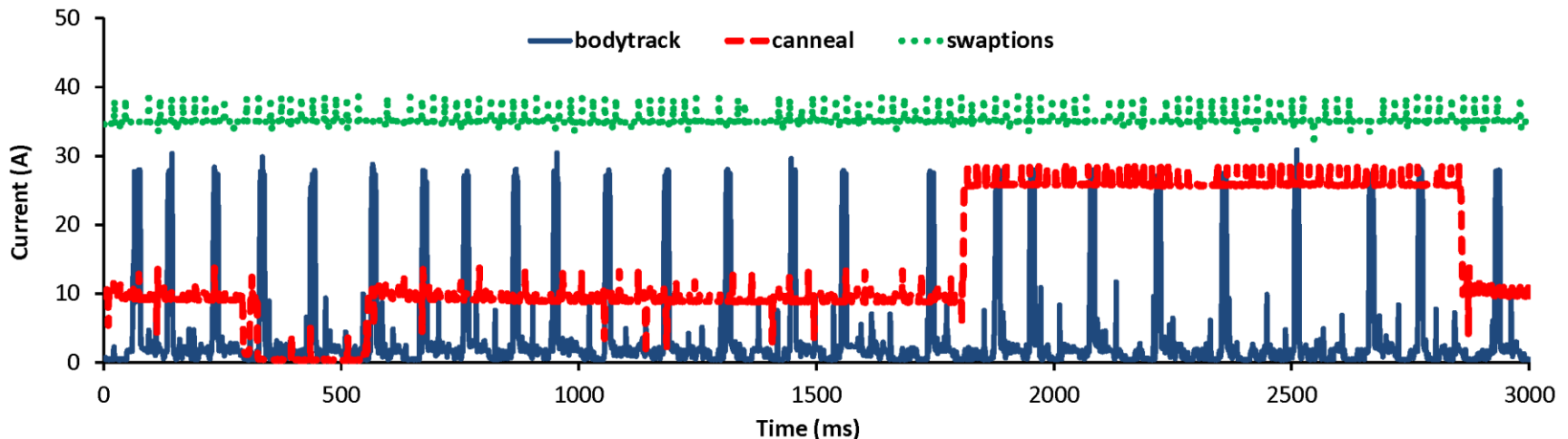


P-State Residency

- PARSEC
 - 50% runtime in P16-P20 (lowest V/F) P states
- DCBench
 - 20% runtime in P16-P20 (lowest V/F) P states
- SPEC
 - 95% runtime in T0-T3 (highest V/F) P states

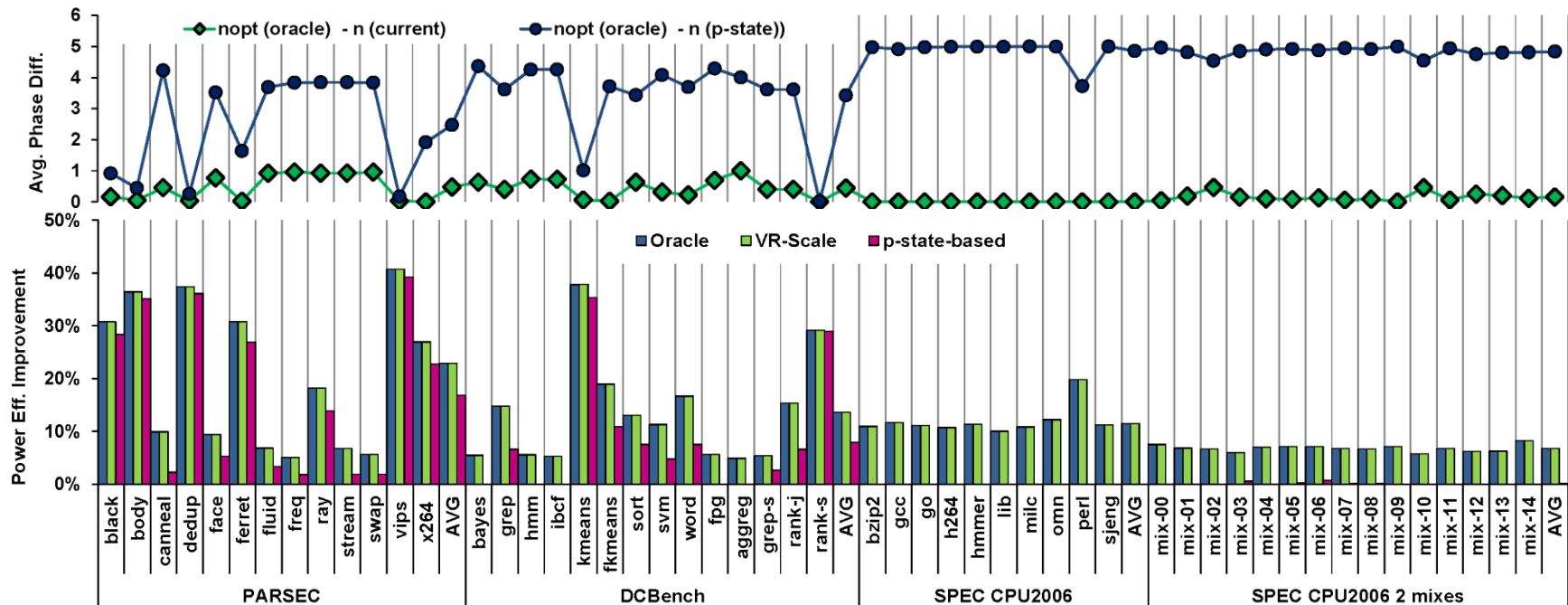


- Load current can significantly change over time but it changes slowly.
 - VR-Scale predicts the load current for the next interval.
- The algorithm determines the optimum number of active phases based on the predicted current.
 - 3.3 μ s processor halt for changing number of active phases



Power Efficiency Improvement

- PARSEC
 - 23% power efficiency improvement w/ 91% pred accuracy
- DCBench
 - 13% power efficiency improvement w/ 80% pred accuracy
- SPEC
 - 11% power efficiency improvement w/ 99% pred accuracy



-
- Voltage regulator (VR) efficiency is strongly dependent on input current, # of VR active phases, and output voltage
 - Parallel benchmarks consume consistently low current for many runtime intervals
 - ⊙ VR efficiency at low current is low in traditional VRs activating all the phases.
 - VR-Scale predicts current consumption and adjusts # of active phases accordingly
 - ⊙ Improves VR efficiency by 19% w/ negligible performance impact

Questions?