IEEE MASCOTS 2016

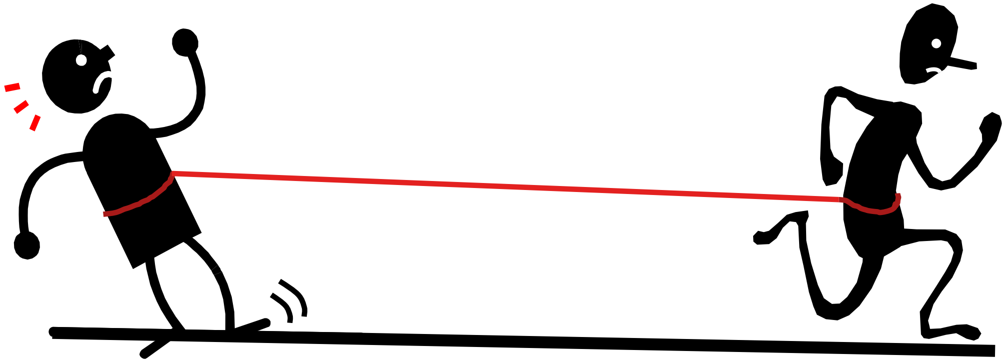
Experimental Calibration and Validation of a Speed Scaling Simulator

Arsham Skrenes
Carey Williamson
Department of Computer Science
University of Calgary

Speed Scaling: Inherent Tradeoffs

Dynamic Speed Scaling: adapt service rate to the current state of the system to balance energy consumption and performance.

Run slower: less ***energy***



Run faster: less ***delay***

- Minimize power consumption ***P***
  - Minimize energy cost $\boldsymbol{\varepsilon}$
  - Minimize heat, wear, etc.
- Minimize response time ***T***
  - Minimize delay
- Maximize job throughput

Background: Theory and Systems

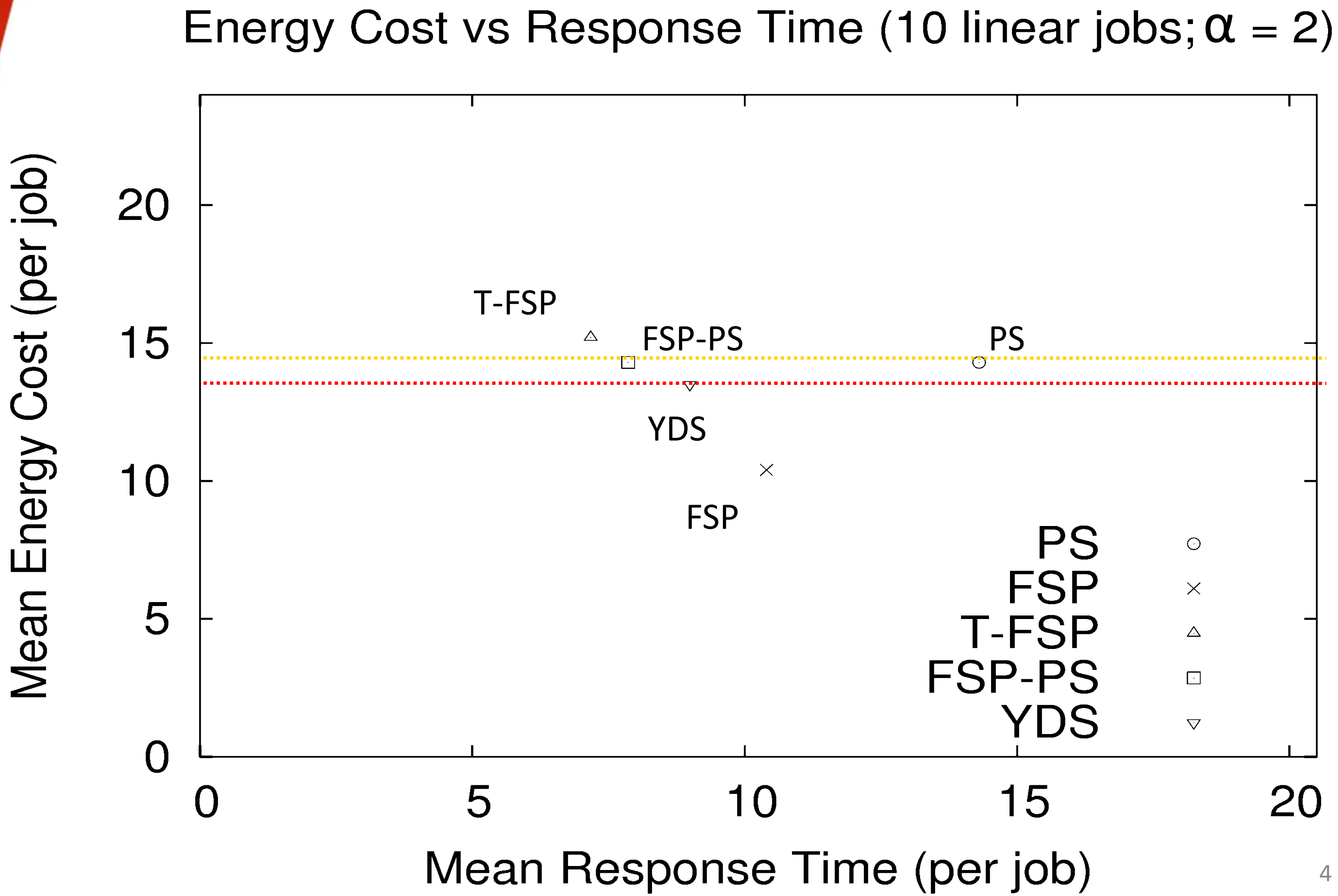
Theoretical Research

- Goal: optimality
- Domains: CPU, parallel systems
- Methods: proofs, complexity, competitive analysis, queueing theory, Markov chains, worst case, asymptotics, simulation
- Metrics: E[T], E[ε], combo, slowdown, competitive ratio
- Power: $P = s^{\alpha}$ $(2 \leq \alpha \leq 3)$
- Schedulers: PS, SRPT, FSP, YDS
- Speed scalers: job-count-based, continuous and unbounded speeds
- Venues: SIGMETRICS, PEVA, Performance, INFOCOM, OR

Systems Research

- Goal: practicality
- Domains: CPU, disk, network
- Methods: DVFS, power meter, measurement, benchmarking, simulation, power gating, over-clocking, simulation
- Metrics: response time, energy, heat, utilization
- Power: $P = a\ C_{eff}\ V^2\ f$
- Schedulers: FCFS, RR, FB
- Speed scalers: threshold-based, discrete and finite speeds
- Venues: SIGMETRICS, SOSP, OSDI, ISCA, MASCOTS, TOCS

Example Simulation Results: IEEE MASCOTS 2014



Simulation Results

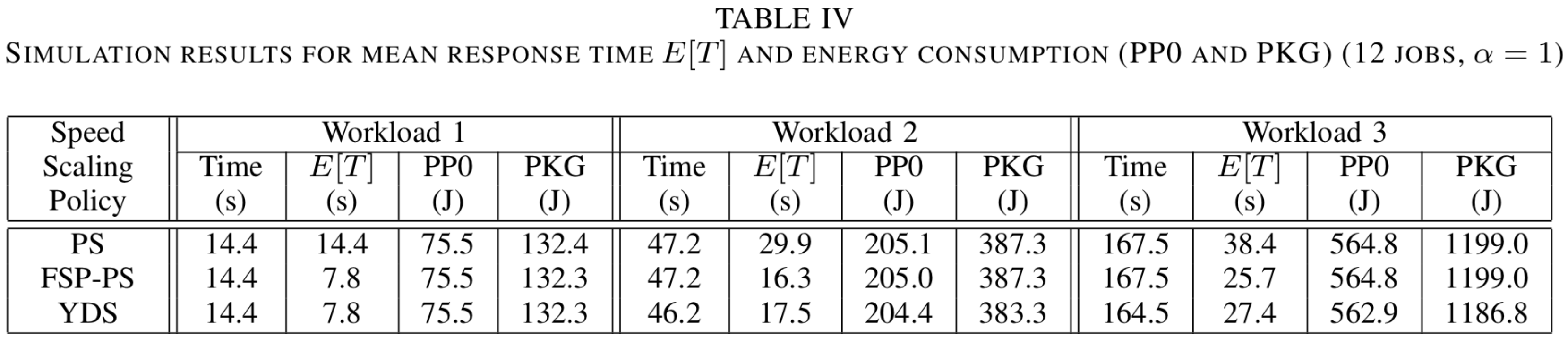


TABLE IV

SIMULATION RESULTS FOR MEAN RESPONSE TIME $E[T]$ AND ENERGY CONSUMPTION (PP0 AND PKG) (12 JOBS, $\alpha = 1$)

Speed Scaling Policy	Workload 1				Workload 2				Workload 3			
	Time (s)	$E[T]$ (s)	PP0 (J)	PKG (J)	Time (s)	$E[T]$ (s)	PP0 (J)	PKG (J)	Time (s)	$E[T]$ (s)	PP0 (J)	PKG (J)
PS	14.4	14.4	75.5	132.4	47.2	29.9	205.1	387.3	167.5	38.4	564.8	1199.0
FSP-PS	14.4	7.8	75.5	132.3	47.2	16.3	205.0	387.3	167.5	25.7	564.8	1199.0
YDS	14.4	7.8	75.5	132.3	46.2	17.5	204.4	383.3	164.5	27.4	562.9	1186.8

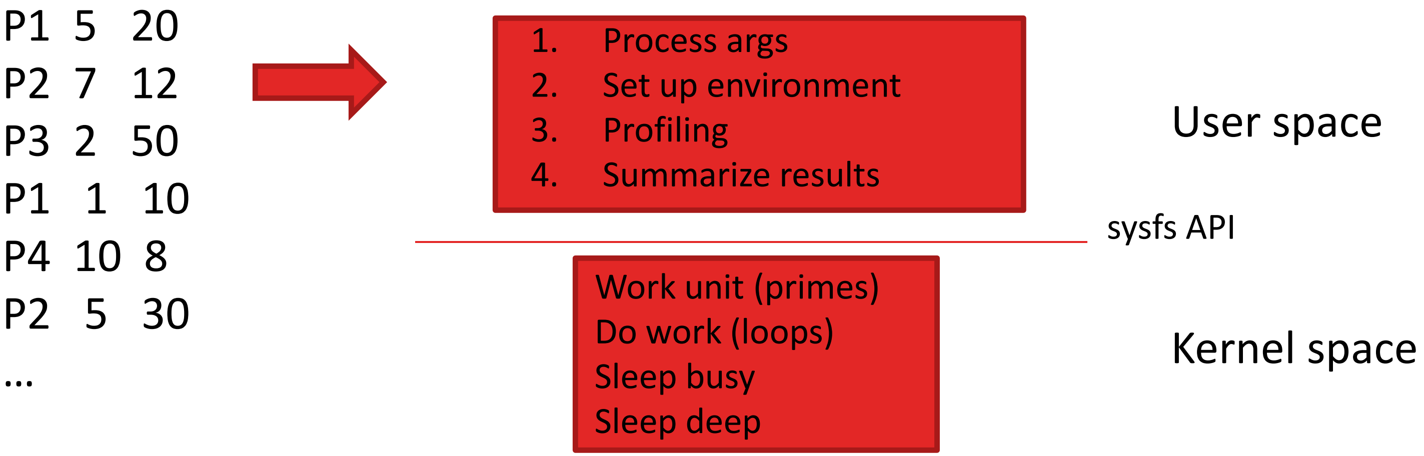
Typical Modeling Assumptions

- Single-server queue for CPU service
- Single batch of n jobs arrive at time 0
- Job sizes known in advance
- Dynamic speed scaling with $s = f(n)$
- Power consumption $P = s^{\alpha}$ where $1 \leq \alpha \leq 3$
- Maximum system speed is unbounded
- System speeds are continuous (not discrete)
- Context switches are free (i.e., zero cost)
- Speed changes are free (i.e., zero cost)

Question: How would they perform on <u>real</u> systems?

Profilo Design [Skrenes 2016]

- Flexible framework for the experimental evaluation of arbitrary scheduling and speed scaling policies
- Hybrid user-mode and kernel-mode implementation
- User space: CSV file input to specify workload
- Kernel space: carefully-controlled job execution, timing, and energy measurement using RAPL MSRs



P1 5 20
P2 7 12
P3 2 50
P1 1 10
P4 10 8
P2 5 30
…

1. Process args
2. Set up environment
3. Profiling
4. Summarize results

User space

sysfs API

Work unit (primes)
Do work (loops)
Sleep busy
Sleep deep

Kernel space

Running Average Power Limit (RAPL)

- Non-architectural model-specific registers (MSRs)
- Four domains (but only three for any given CPU):
  - → PP0: Power Plane 0 for the CPU cores
  - PP1: Power Plane 1 for GPU (consumer machines only)
  - DRAM: Memory energy (server-class machines only)
  - → PKG: Energy usage by rest of the CPU chip package
- Highly accurate power meters for each domain (matches well with external power measurements)
- Experiments conducted on Macbook Pro Retina laptop (2012): 2.3 GHz quad-core Intel i7-3615 QM Ivy Bridge processor; Ubuntu Linux 14.04 LTS; compute-intensive workload with no I/O, memory, or networking involved

Measurement Results

Frequency (MHz)	PP0 (W)	PKG (W)	Context Switch (us)	Speed Switch (us)	Mode Switch (ns)
2301 (3300)	11.5	15.3	1.140	0.76	44.8
2300	5.4	9.2	1.634	1.09	64.2
2200	5.0	8.9	1.708	1.14	67.0
2100	4.8	8.6	1.808	1.20	70.2
2000	4.6	8.4	1.898	1.26	73.7
1900	4.5	8.3	1.999	1.32	78.3
1800	4.3	8.0	2.118	1.38	81.9
1700	4.1	7.9	2.213	1.47	86.7
1600	3.9	7.6	2.369	1.56	92.1
1500	3.7	7.5	2.526	1.67	98.6
1400	3.5	7.3	2.709	1.81	105.3
1300	3.3	7.1	2.886	1.93	113.4
1200	3.1	6.9	3.167	2.09	123.1

Experimental Evaluation Setup

- Three workloads (each with batch of 12 jobs):
  1. Homogenous
  2. Additive (arithmetic progression)
  3. Multiplicative (factors of 2)

- Three algorithms (all with α=1):
  1. PS (epitomizes fairness)
  2. FSP-PS (decoupled speed scaling; improves mean response time while retaining fairness)
  3. YDS (minimizes power consumption)

UNIVERSITY OF CALGARY

Experimental Evaluation Results

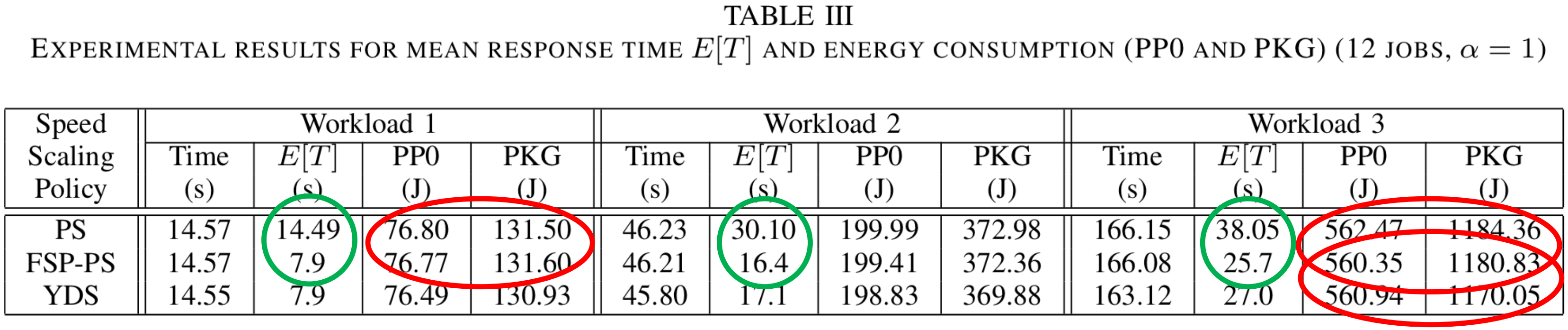


TABLE III

EXPERIMENTAL RESULTS FOR MEAN RESPONSE TIME $E[T]$ AND ENERGY CONSUMPTION (PP0 AND PKG) (12 JOBS, $\alpha = 1$)

Speed Scaling Policy	Workload 1				Workload 2				Workload 3			
	Time (s)	$E[T]$ (s)	PP0 (J)	PKG (J)	Time (s)	$E[T]$ (s)	PP0 (J)	PKG (J)	Time (s)	$E[T]$ (s)	PP0 (J)	PKG (J)
PS	14.57	14.49	76.80	131.50	46.23	30.10	199.99	372.98	166.15	38.05	562.47	1184.36
FSP-PS	14.57	7.9	76.77	131.60	46.21	16.4	199.41	372.36	166.08	25.7	560.35	1180.83
YDS	14.55	7.9	76.49	130.93	45.80	17.1	198.83	369.88	163.12	27.0	560.94	1170.05

- Observation 1: Decoupled speed scaling (FSP-PS) provides a significant response time advantage over PS, for the “same” energy costs
- Observation 2: The response time advantage of FSP-PS decreases as job size variability increases
- Observation 3: FSP-PS has a slight energy advantage over PS because of fewer context switches between jobs
- Observation 4: YDS has the lowest energy consumption among these policies (even better than expected due to discretization effect, and no speed changes)

Simulation Results

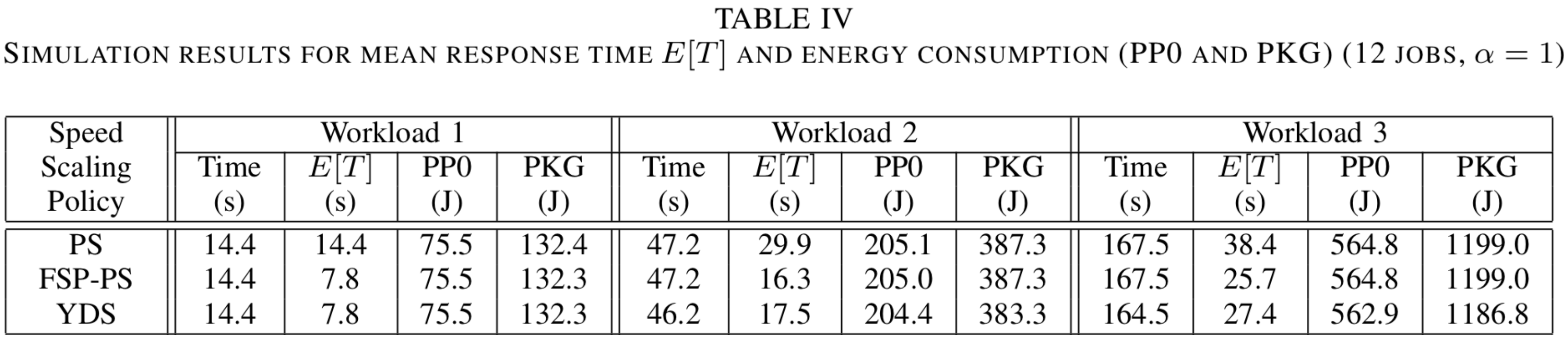


TABLE IV

SIMULATION RESULTS FOR MEAN RESPONSE TIME $E[T]$ AND ENERGY CONSUMPTION (PP0 AND PKG) (12 JOBS, $\alpha = 1$)

Speed Scaling Policy	Workload 1				Workload 2				Workload 3			
	Time (s)	$E[T]$ (s)	PP0 (J)	PKG (J)	Time (s)	$E[T]$ (s)	PP0 (J)	PKG (J)	Time (s)	$E[T]$ (s)	PP0 (J)	PKG (J)
PS	14.4	14.4	75.5	132.4	47.2	29.9	205.1	387.3	167.5	38.4	564.8	1199.0
FSP-PS	14.4	7.8	75.5	132.3	47.2	16.3	205.0	387.3	167.5	25.7	564.8	1199.0
YDS	14.4	7.8	75.5	132.3	46.2	17.5	204.4	383.3	164.5	27.4	562.9	1186.8