Common Errors and Assumptions in Energy Measurement and Management

Jóakim v. Kistowski

University of Würzburg

Symposium on Software Performance, November 5th 2015, Munich, Germany
What is this Talk about?

- Measurement methodologies for energy efficiency
  - Focus on server systems

- Some pitfalls: Energy efficiency measurements can be unrepresentative or inaccurate if done incorrectly

- SPEC power methodology [1]: A methodology for standardized energy efficiency benchmarking

- Some results that challenge common implicit assumptions on energy efficiency of servers
Energy Efficiency of Servers

- Relationship of Performance and Power

- For transactional workloads:

\[
\frac{\text{transactions}}{\text{energy}} \left[ \frac{1}{J} \right] = \frac{\text{throughput}}{\text{power}} \left[ \frac{1/s}{W} \right]
\]

- Comparison of efficiency of different workload types is difficult
  - Different scales of transaction-counts / throughput
  - \(\Rightarrow\) normalization
How to do it wrong…

PITFALLS IN POWER MEASUREMENT
A typical server …

- has an average utilization between 10% and 50%,
- is provisioned with additional capacity (to deal with load spikes).
- is not energy efficient at low utilization, more efficient at high utilization

Power consumption depends on server utilization.
Bad Practice for…

- Full system power characterization
- Comparison of server systems intended for transactional workloads (most of them)

Good Practice for…

- HPC energy efficiency benchmarking
- Power meters have power measurement ranges
  - Lose measurement accuracy outside of range
  - Switching ranges takes time (~ 1 s)

- Example

![Load Profile Power](chart.png)
Lessons:

- Auto-Ranging is bad for varying loads
  - Lose measurements

- But:
  - Disabling auto-ranging decreases accuracy

- Measurement uncertainty depends on power meter
  - SPEC PTDaemon supported ➔ Less than 1% at optimal range

- Also:
  - Good load calibration is important
How to do it right...

SPEC POWER METHODOLOGY
Methodology for benchmarking of energy efficiency

Goal:
- Benchmarking at multiple load levels
- Taking the quality criteria for benchmarks into account [3]:
  - Relevance
  - Reproducibility
  - Fairness
  - Verifiability
  - Usability

Used in the following SPEC products:
- SPECpower_ssj2008 [4]
- SPEC SERT [5]
- ChauffeurWDK

Other Benchmarks that follow the methodology:
- SAP Power Benchmark [6]
- TPC Energy [7]
Goal: For a given workload, achieve a load level of n% of system “utilization”.

Utilization = \( \frac{t_{busy}}{t_{busy} + t_{idle}} \)

DVFS increases CPU busy time at low load
- \( \rightarrow \) increases utilization
- Power over load measurements need to compensate
  How to compare?

Our solution: Machine utilization
- 100% utilization at calibrated maximum throughput

Load level = \( \frac{\text{current throughput}}{\text{max. throughput}} \)
- Controller System runs
  - SPEC Director: Chaffeur
  - Reporter

- PTDaemon
  - Network-capable power and temperature measurement interface
  - Can run on controller system or separate machine

- SUT runs
  - Host, which launches
  - Pinned SERT clients
Transactional workloads are dispatched in “Intervals”:

- Warmup
- Calibration
  - Multiple intervals
  - Maximum transaction rate
- Graduated Measurement Series
  - Multiple intervals at decreasing transaction rate
  - Target transaction rate is percentage of calibration result
  - Exponentially distributed wait times between transactions
- Separate measurement intervals at stable states
  - 10 second sleep between intervals
  - 15 second pre-measurement run
  - 15 second post-measurement run
  - 120 second measurement

- Temperature analyzer for comparable ambient temperature

- Power Measurements: AC Wall Power
Throughput results from load level definition
- Throughput variation is measure of benchmark driver stability
- Throughput coefficient of variation $> 5\% \Rightarrow$ invalid interval

Power consumption results from SUT response to load
- Power variation is measure of SUT stability
- CVs often $< 1\%$ on state-of-the-art x86 systems
Workloads can be anything, as long as...

... they have a measurable throughput

... allow for result validation

Common Workloads:

- SPEC SERT: “Worklets”
  - 7 CPU Workets
  - 2 HDD Worklets
  - 2 Memory Worklets
  - 1 Hybrid Worklet (SSJ)
- SPECpower_ssj2008: Business Transactions
- TPC Energy
- ChauffeurWDK: Allows custom workload creation
Motivating future work…

SOME MEASUREMENT RESULTS
The Software Stack Matters! (1/2)

(With differing extent)

- Operating System [8]
  - Impact on base consumption and power scaling behavior

![Graph: RHEL6.4_E5-2690_8x8GB Power vs Load Level (%)]

![Graph: W2012_E5-2690_8x8GB Power vs Load Level (%)]
(With differing extent)

- JVM [8]
  - Little impact through secondary effects
Energy Efficiency depends on multiple factors
- Hardware
- Software Stack
- Workload
- Load Distribution

Maximum Energy Efficiency is often reached at < 100% load

Result: Load Consolidation is not most efficient load distribution strategy [9]
Conclusions

- Power and energy efficiency measurements have many pitfalls
  - Can lead to inaccurate or missing results

- SPEC power methodology is an established standard to avoid errors in energy efficiency benchmarking
  - Goal: Energy efficiency characterization at multiple load levels

- Results demonstrate that energy efficiency and energy efficiency scaling depend on many factors, including hardware, software stack, workload, etc.
Thanks for listening!

joakim.kistowski@uni-wuerzburg.de
http://se.informatik.uni-wuerzburg.de
The SPEC logo, SPEC, and the benchmark and tool names, SPECpower_ssj, SERT, PTDaemon are registered trademarks of the Standard Performance Evaluation Corporation. Reprint with permission, see spec.org.

The opinions expressed in this tutorial are those of the author and do not represent official views of either the Standard Performance Evaluation Corporation, Transaction Processing Performance Council or author’s company affiliation.
References


