An experimental comparison of software-based power meters

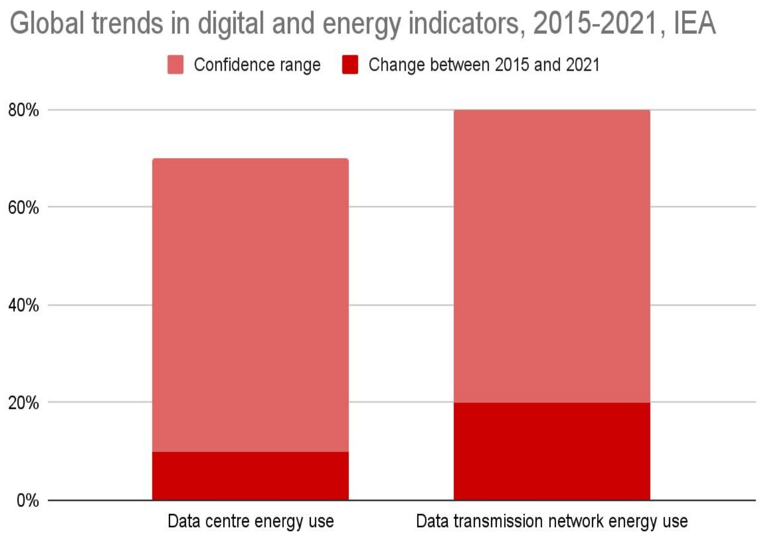
Mathilde Jay [1]
Vladimir Ostapenco [2]

Laurent Lefevre [2]
Anne-Cécile Orgerie [3]
Denis Trystram [1]
Benjamin Fichel [4]

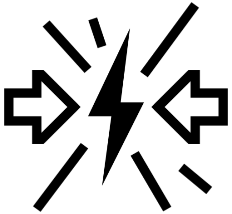
[1] Univ. Grenoble Alpes, CNRS, Inria, Grenoble INP, LIG
[2] Univ Lyon, EnsL, UCBL, CNRS, Inria, LIP
[3] Univ Rennes, Inria, CNRS, IRISA
[4] OVHcloud

CCGRID 2023

Why measuring energy?



https://www.iea.org/reports/data-centres-and-data-transmission-networks



European regulations

- Climate and environmental impact report standard (CSRD)
- Product Environmental Footprint (PEF)

Scope: services

Data center

Cloud services

Artificial intelligence

Why?

- Center of digital impact growth
- Now in most industrial sectors
- Environmental cost of **services** remain mostly unknown

Challenges

Data center

- High energy demand
- Cooling and power delivery overhead

Cloud services

- Virtualized environment
- Over provisioning

Artificial intelligence

- Computationally intensive
- Specialized equipment (GPUs)

→ **Specific tools are needed to measure energy consumption**

External measurement

Most commonly used way to measure energy consumption at the electrical outlet

External devices :

- ITECH power meters
- Eaton power meters
- OmegaWatt power meter
- Raritan intelligent power distribution units

✓ Accurate and used for a long time
✓ Has little impact on the measured system

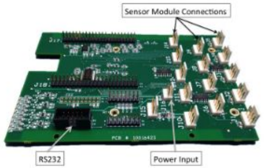
× Difficult to implement on a large scale and can be expensive
× Usually has low measurement rates, which gives too coarse-grained results
× Reports only the consumption of the entire computing node

Intra-node devices

Equipment placed inside computing nodes capable of reporting energy consumption metrics

BMC

PowerMon



PowerInsight

- ✓ Provide information on the consumption of individual computing node components
- × Cost-ineffective
- × Lack user-friendliness

Hardware sensors and software interfaces

Digital sensors, onboard measurement circuits, and interfaces integrated in computing nodes

Intel RAPL **Nvidia NVML** **AMD APM/RAPL**

Allow to measure the power of :

- Entire system
- Processor socket
- Memory
- Other components of a computing node

✓ More cost-effective and user-friendly

✓ Already available on the majority of modern computing nodes

× Direct use of these interfaces is relatively complicated

Power and energy modeling

Usage-based modeling

- Based on resource usage
- Computing node usage-based power modeling
- Component-specific usage-based power modeling
  - CPU (utilization rate)
  - RAM (last-level-cache-misses)
  - Disk (number of read/writes)
  - Network interface (in/out traffic)
- Process level usage-based power modeling

Usage-based modeling with TDP

- Thermal design power (**TDP**) is provided by the manufacturer and represents the **maximum amount of heat** that is generated by a component under a steady workload
- Can be used as a reasonable approximation of component power consumption at maximum use

Energy = TDP x Avg. Utilization x Execution Time

✓ Relatively simple to implement
✓ Can be done offline
✗ Accuracy depends on workload

How to use all this?

- **Most of the methods introduced earlier**
  - Need specific equipment or detailed knowledge of the underlying hardware being used
  - Require additional implementation
    - Implement and tune models
    - Query integrated technologies and process data to obtain energy metrics
- Fortunately, a number of **software-based power meters** were developed to simplify the process

Scaphandre

PowerAPI

Carbon Tracker Initiative

CODE CARBON

Breakend/**experiment-impact-tracker**

ML CO2 Impact

Green Algorithms

Energy scope

Scope of our study

Which tool is the most adapted to one's need?

- Experimental comparison of a selection of software-based power meters
- On both qualitative and quantitative criteria

Selected tools

- Power profiling
  - **PowerAPI** with Smartwatts - Inria Spirals
  - **Energy Scope** - Inria Bordeaux
  - **Scaphandre** - Hubblo Org
  - **Perf** - IBM
- Energy measurement
  - **Code Carbon** - MILA
  - **Experiment Impact Tracker** - Univ. Standford
  - **Carbon Tracker** - Univ. Copenhagen
- Energy calculators
  - **ML CO2 Impact** - MILA
  - **Green Algorithms** - Univ. Cambridge

Qualitative analysis

		Measures		Power profiling software				Energy measurement software packages			Energy calculators	
		OmegaWatt	BMC	Power API	Scaphandre	Energy Scope	Perf	Code Carbon	Experiment Impact Tracker	Carbon Tracker	Green Algorithms	ML CO2 Impact
Development	Citation			Inria - Spirals	Hubblo	Inria Bordeaux	IBM	MILA	University of Standford	University of Copenhagen	University of Cambridge	MILA
	First (latest) release date			Jul. 2019 (Aug. 2022)	Dec. 2020 (May 2021)	2021	Sept. 2009 (Jan. 2023)	Nov. 2020 (Sept. 2022)	Dec. 2019 (Jan. 2020)	Apr. 2020 (Jul. 2021)	Jul. 2020 (Jun. 2022)	Aug. 2019 (Jul. 22)
Environment	Hardware compatibility	Any	Any	Intel RAPL	Intel RAPL	Intel RAPL, Nvidia NVML	Intel RAPL	Any	Intel RAPL, Nvidia NVML	Intel RAPL, Nvidia NVML	Any	Any
	Scope	Machine	Machine	CPU, DRAM, process	CPU, DRAM, process	CPU, DRAM, GPU	CPU, DRAM	CPU, DRAM, GPU	CPU, DRAM, GPU, process	CPU, DRAM, GPU	Machine	Machine
	Virtualization support			Yes	Yes	No	No	No	No	No		
	Job management support			No	No	OAR, SLURM	No	No	No	No		
Functional	Hardware technology used			RAPL	RAPL	RAPL, NVML	RAPL	RAPL, NVML, TDP	RAPL, NVML	RAPL, NVML	TDP	TDP
	Software power model used			Regression based on perf events	CPU usage based				GPU, CPU and RAM usage based			
	Default sampling frequency (Hz)	1	0.2	1	0.1	2	10	1/15	1	0.1		
	Online reporting	Yes	Yes	Yes	Yes	No	Yes	No	No	No	No	No
	Power profiling	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No	No	No
User-friendliness	Availability of source code (License)			Yes (BSD 3-Clause)	Yes (Apache 2.0)	No	Yes (GNU GPL)	Yes (MIT)	Yes (MIT)	Yes (MIT)	Yes (CC-BY-4.0)	Yes (MIT)
	Ease of use	Poor	Poor	Poor	Fair	Good	Good	Good	Good	Good	Very good	Very good
	Quality of documentation			Good	Very good	Good	Good	Quite good	Quite good	Good	Good	Quite good
	Configurability	Fair	Poor	Good	Good	Good	Good	Poor	Fair	Poor	Poor	Poor
	Resulting data format	HTTP endpoint	HTTP endpoint	MongoDB, InfluxDB, Prometheus, CSV, Socket, File	Prometheus, Warp10, Riemann, JSON, Stdout	JSON	CSV, Stdout, File	CSV	JSON, Code	File, Code	Web	Web, Latex
	Data visualisation possibilities	Grafana (Kwollect)	Grafana (Kwollect)	Grafana (InfluxDB, Prometheus)	Grafana (Prometheus)	Custom Dashboard		Comet			Graphs on the web page	

Experimental methodology

Environment: Grid'5000 - cluster Gemini

- 2 CPUs Intel **Xeon E5-2698 v4** – 20 cores/CPU
- 8 GPU Nvidia **Tesla V100** (32GiB)
- Physical power meters (OmegaWatt)
- Baseboard Management Controller (BMC)

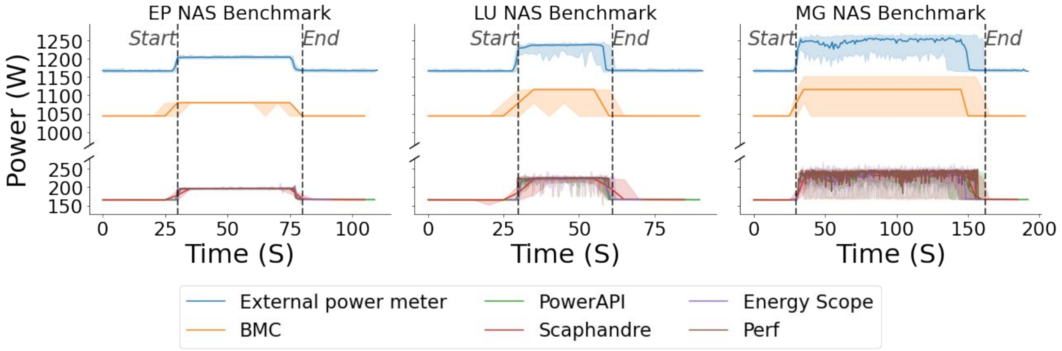
Benchmarks

- **CPU**: *NAS Parallel Benchmarks* (NPB) in *OpenMP/Fortran*
- **GPU**: *NAS Parallel Benchmarks* (NPB) in *CUDA*

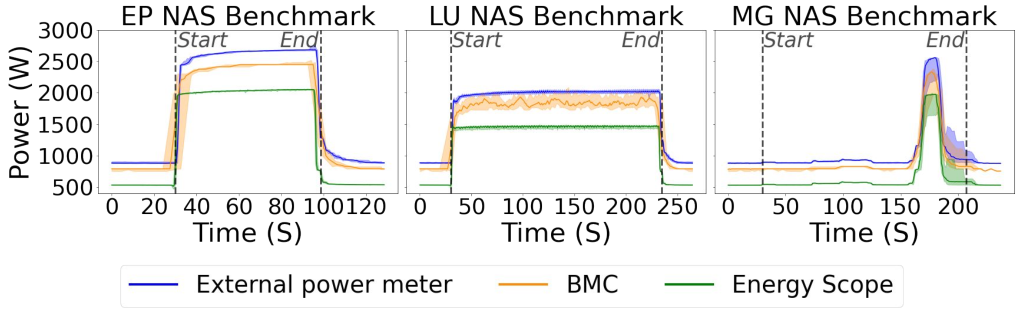
Benchmark application	Abbreviation	Stress target
Embarrassingly Parallel	EP	CPU/GPU
Multi-Grid on a sequence of meshes	MG	RAM
Lower-Upper Gauss-Seidel solver	LU	RAM + CPU/GPU

Power Profile

CPU Benchmarks



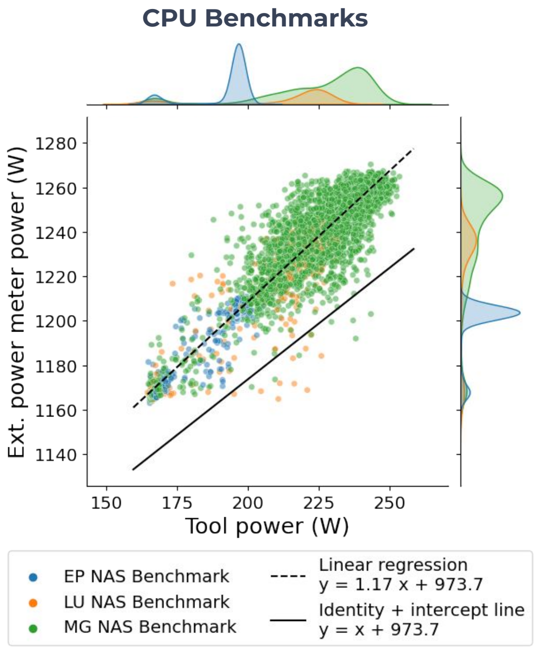
GPU Benchmarks

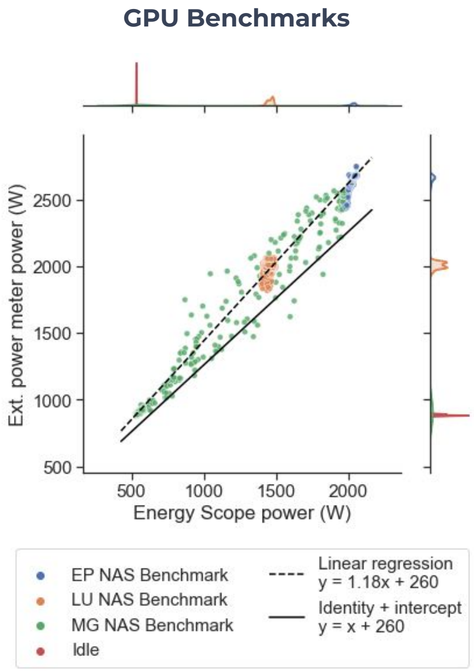


Origin of **offsets**

- **Tools**
  - CPU, RAM, GPU
- **Power meter**
  - CPU, RAM, GPU
  - Fans
  - Power supply
  - Disk
  - Network card
  - ....

Correlation with external power meter





- **Bias in power** between the external power meter and tools
- Pearson correlation coefficient **> 0.95**

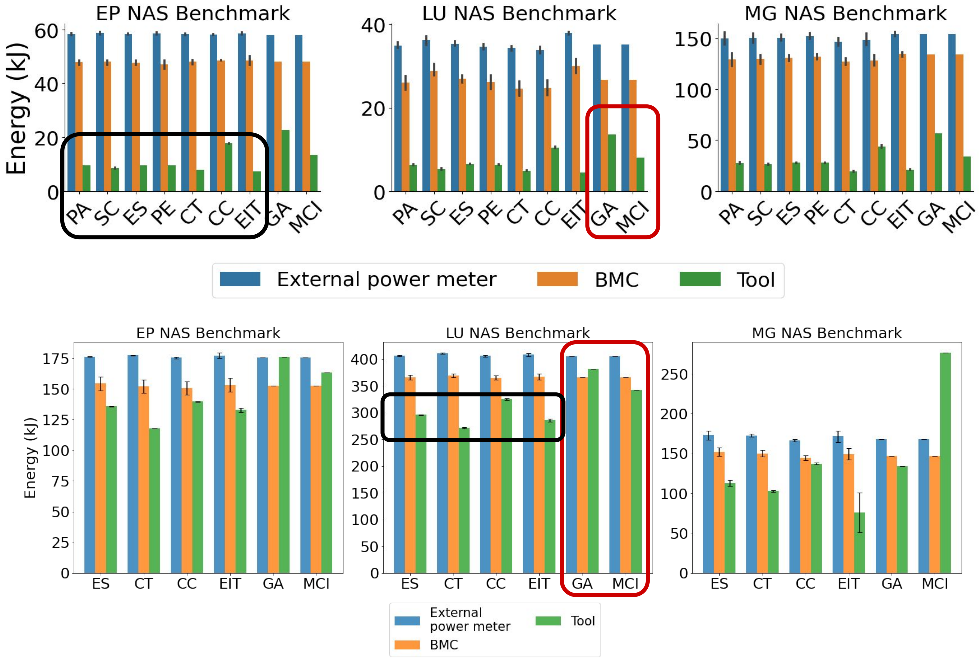
✓ Very strong correlation

Energy consumed

- **PA** - PowerAPI
- **SC** - Scaphandre
- **ES** - Energy Scope
- **PE** - Perf
- **CT** - Carbon Tracker
- **CC** - Code Carbon
- **EIT** - Experiment Impact Tracker
- **GA** - Green Algorithms
- **MCI** - ML CO2 Impact

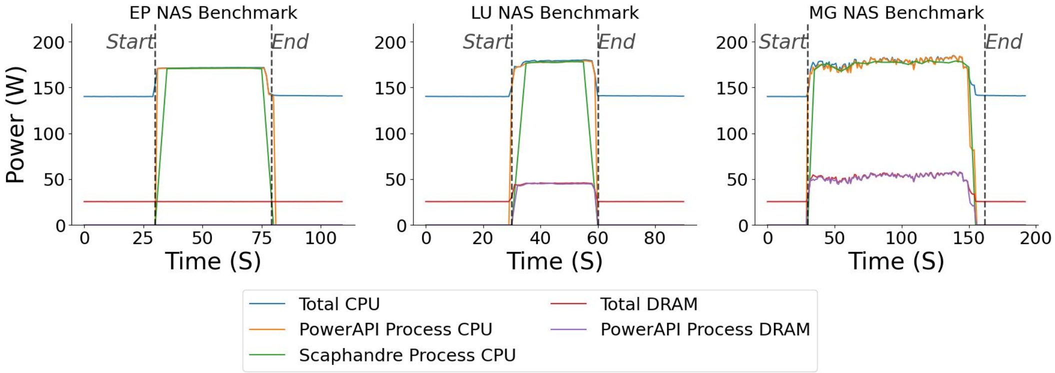
CPU Benchmarks

GPU Benchmarks



- **Similar estimations** for tools based on internal interfaces for both CPU and GPU benchmarks
- **Good estimations** from Green Algorithms and ML CO2 Impact on constant workloads
- Higher variability on non-constant workloads

By process power profile



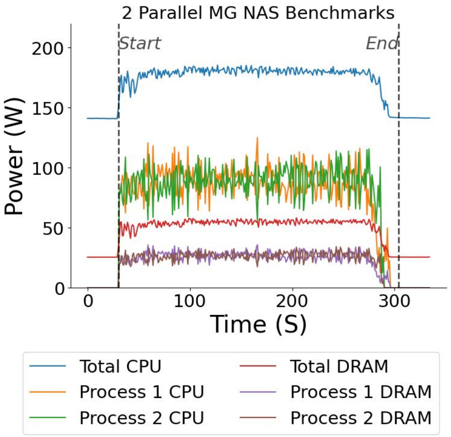
Power profiles given by PowerAPI and Scaphandre for a single process

- **Scaphandre**
  - *Usage-based process-level modeling*
  - Based on ***CPU utilization rate*** and ***RAPL*** counters metrics
- **PowerAPI**
  - *Performance event based regression modeling*
  - Based on ***hardware performance counter events*** and ***RAPL*** counter metrics

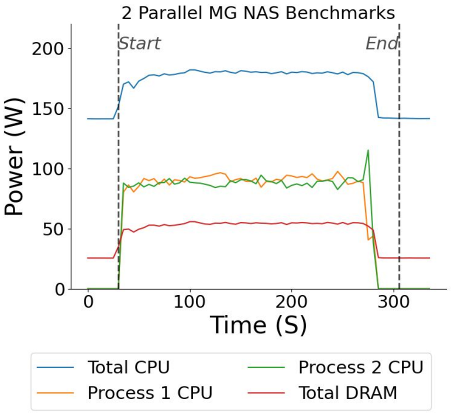
Multiple processes

All is good

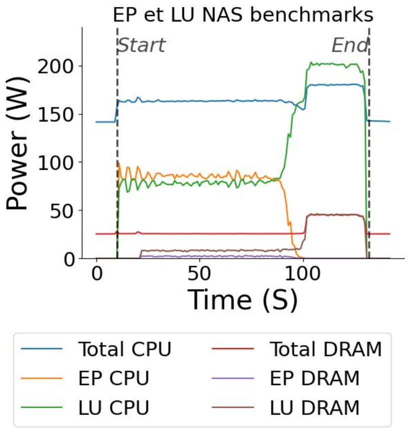
PowerAPI

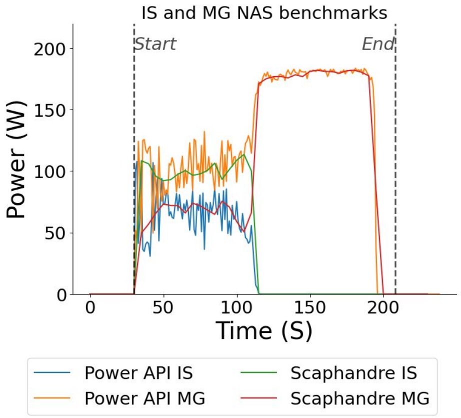


Scaphandre



Problems and differences





- × PowerAPI overestimates CPU
- × PowerAPI and Scaphandre give different estimates

Limits

Maximum number of simultaneously profiled processes

- **PowerAPI**
  - Tested version < **6** processes
  - New version (Smartwatts v0.9.2)
    - ~ **18** processes without delay
    - Estimates for 50 processes but with delay
- **Scaphandre**
  - **+ 100** processes

Maximum supported sampling frequency

- **PowerAPI**
  - Default: 1 Hz
  - Maximum: **10 Hz**
- **Scaphandre**
  - Default: 0.1 Hz
  - Maximum: **0.5 Hz**
- **Energy Scope**
  - Default: 2 Hz
  - Maximum: **50 Hz**
- **Perf**
  - Default: 10 Hz
  - Maximum: **1000 Hz**

Recommendations

	Power profiling software				Energy measurement packages			Energy calculators	
	PowerAPI	Scaphandre	Energy Scope	Perf	Code Carbon	Experiment Impact Tracker	Carbon Tracker	Green Algorithms	ML CO2 Impact
Measure **exact** power consumed									
Reduce energy needed to execute a workload	✔	✔	✔	✔	✔	✔	✔	✔	✔
Monitor power consumption in **real time**	✔	✔		✔					
Monitor **cloud** services	✔	✔							
Monitor **GPUs**			✔		✔	✔	✔	✔	✔
Long-running jobs	✔	✔	✔	✔	✔	✔	✔	✔	✔
Short executions	✔		✔	✔					

Conclusion

- *Software-based power meters* relying on **TDP** modeling, **Intel RAPL** or **Nvidia NVML** can be used to estimate energy consumption with ***low overhead***
- Outputs are **consistent** with power meters
- Main differences between them
    - Supported sampling frequencies
    - User-friendliness
    - Environment in which they can be used
    - Ability to estimate the power at various granularities

No excuse not to use them!

Thank you!

Artifacts

https://inria.hal.science/hal-03974900

For our study we have consumed a total of 481 kWh.

Contacts

Mathilde Jay
mathilde.jay@univ-grenoble-alpes.fr

Vladimir Ostapenco
vladimir.ostapenco@inria.fr