Pinpoint the Joules

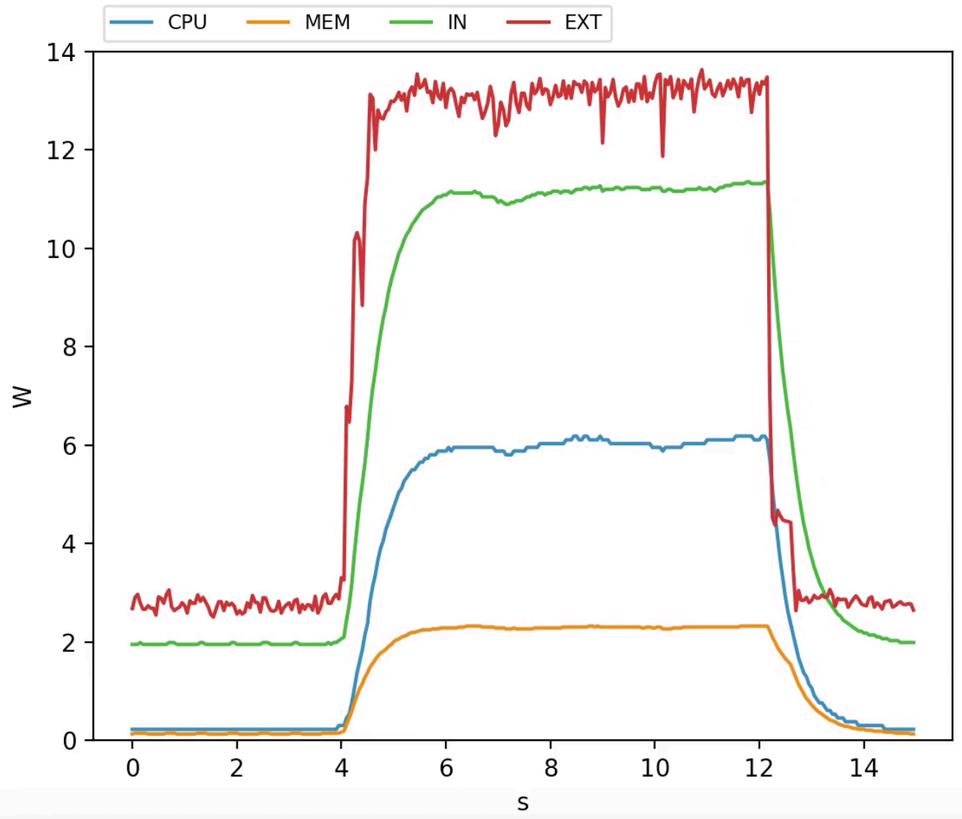
Unifying Runtime-Support for Energy Measurements on Heterogeneous Systems

Sven Köhler[1], Benedict Herzog[2], Timo Hönig[3], Lukas Wenzel[1], Max Plauth[1], Jörg Nolte[4], Andreas Polze[1], and Wolfgang Schröder-Preikschat[2]

International Workshop on Runtime and Operating Systems for Supercomputers (ROSS 2020)

[1] Hasso Plattner Institute for Digital Engineering, University of Potsdam, Germany {firstname.lastname}@hpi.de
[2] Friedrich-Alexander University Erlangen-Nürnberg (FAU), Germany {firstname.lastname}@cs.fau.de
[3] Ruhr University Bochum (RUB), Germany timo.hoenig@rub.de
[4] Brandenburg University of Technology Cottbus–Senftenberg (BTU), Germany joerg.nolte@b-tu.de

jetson-tx2-01:live



jetson-tx2-01: ~

```
jetson-tx2:~$ pinpoint -i 50 -e CPU,MEM,IN,EXT ./heat 1000 1000 10000 in.csv
Energy counter stats for './heat 1000 1000 10000 in.csv':
[interval: 50ms, before: 0 ms, after: 0 ms, delay: 0 ms, runs: 1]

        44668.45 mJ     CPU
        18967.85 mJ     MEM
        87475.30 mJ     IN
        109235.00 mJ    EXT

        8.61876501 seconds time elapsed

jetson-tx2:~$
```

1 Motivation

Power and energy demand are a critical operating resource …

SUPERCOMPUTER FUGAKU - SUPERCOMPUTER FUGAKU, A64FX 48C 2.2GHZ, TOFU INTERCONNECT D

Site:	RIKEN Center for Computational Science
System URL:	https://www.r-ccs.riken.jp/en/fugaku/project
Manufacturer:	Fujitsu
Cores:	7,299,072
Memory:	4,866,048 GB
Processor:	A64FX 48C 2.2GHz
Theoretical Peak (Rpeak)	513,855 TFlop/s
Power Consumption	
Power:	28,334.50 kW (Submitted)
Power Measurement Level:	2

[1]

[2]

Original research article

Digitalisation, energy and data demand: The impact of Internet traffic on overall and peak electricity consumption

Janine Morley [a], Kelly Widdicks [b], Mike Hazas [b]

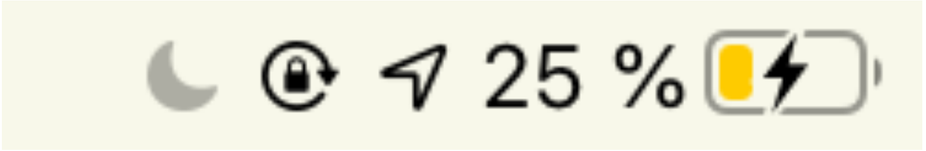
Show more

https://doi.org/10.1016/j.erss.2018.01.018

Get rights and content

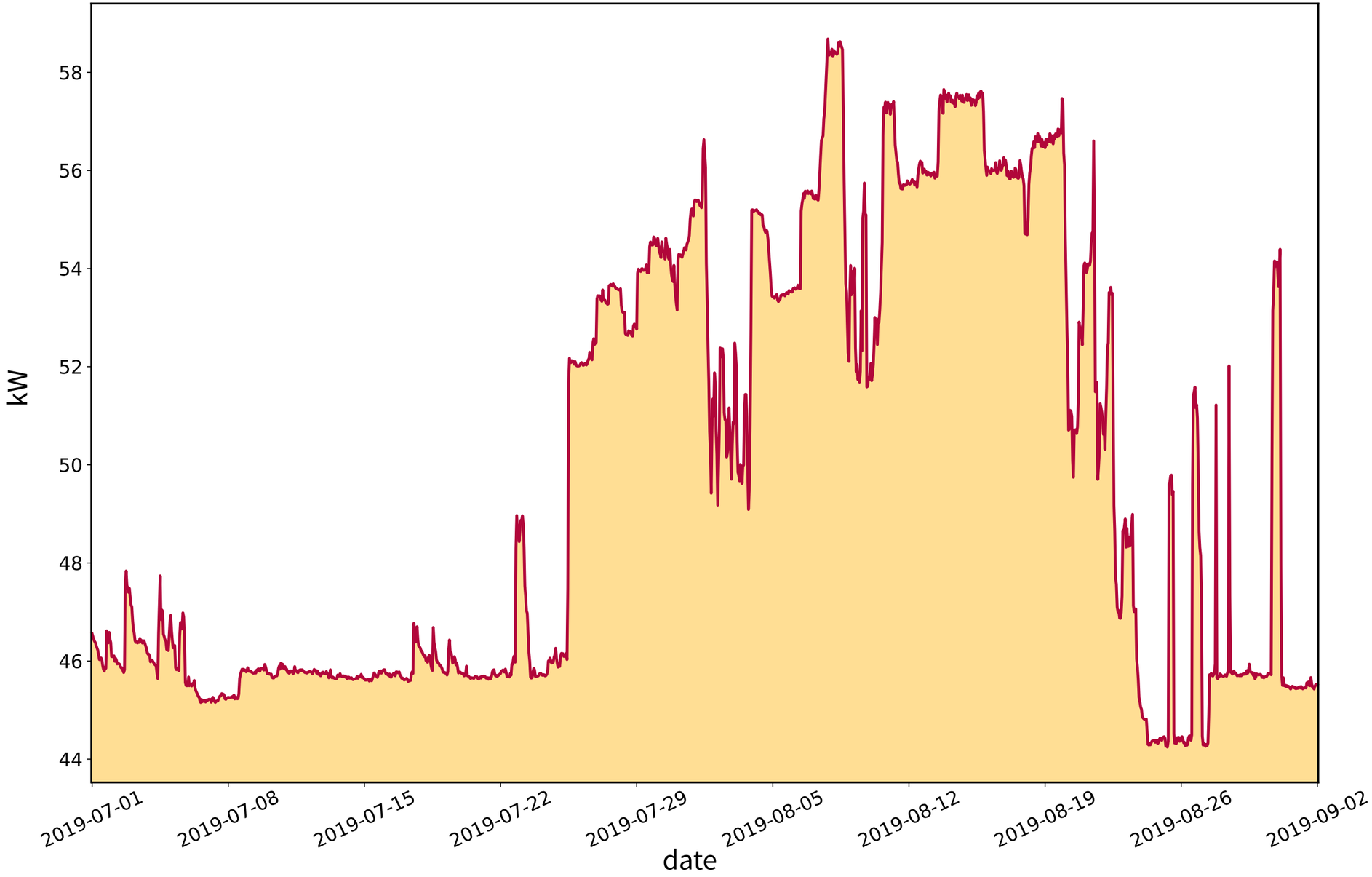
Abstract

Over the last decade, concerns have been raised about increases in the electricity used by information technologies, other consumer electronic devices, data centres



[1] Fugaku Supercomputer. Acc.2020-08-13. https://www.top500.org/system/179807
[2] Morley, J., Widdicks, K., & Hazas, M. (2018). Digitalisation, energy and data demand: The impact of Internet traffic on overall and peak electricity consumption. Energy Research & Social Science, 38, 128-137.

Power demand fluctuates over time …



[HPI Future SOC Lab Potsdam, UPS power supply, hourly log]

How does the power behavior of a workload change …

… over **time**?
… when **configured** differently?
… when **implemented** differently?
… when ported to **another platform**?

Questions related to performance analysis, but performance does not always correlate and energy/power characteristics.

Hönig, T., Janker, H., Eibel, C., Mihelic, O., & Kapitza, R. (2014). Proactive Energy-Aware Programming with PEEK. In 2014 Conference on Timely Results in Operating Systems (TRIOS 14).

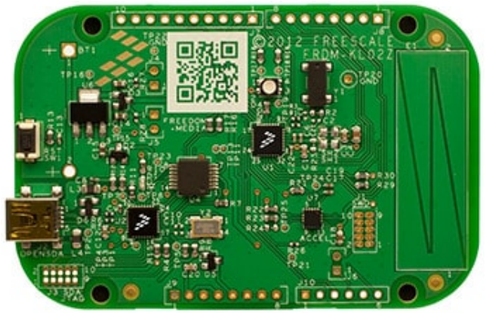
2

Hardware Platforms

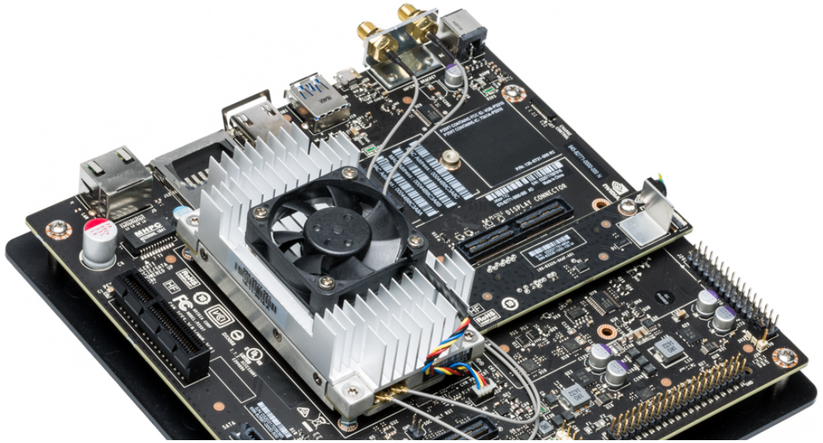
Coral USB: TPU

Desktop PC: CPU

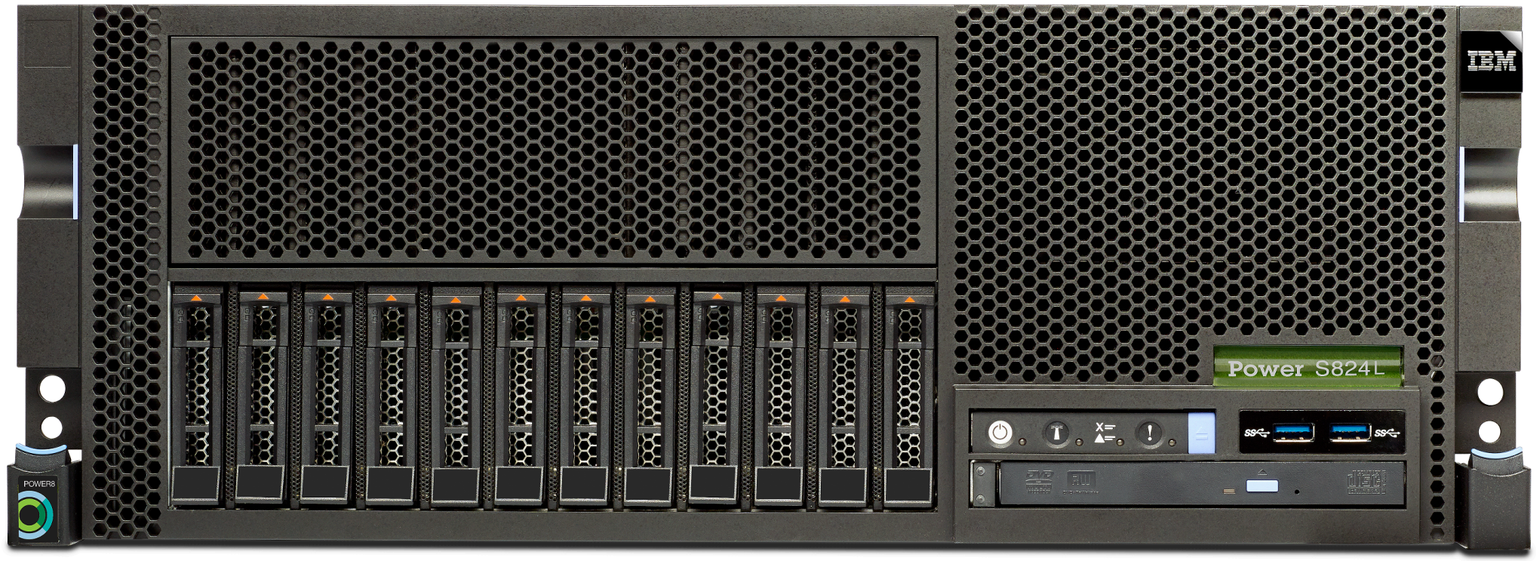
HPE BL460c: 2x CPU



FRDM-KL02Z: Microcontroller

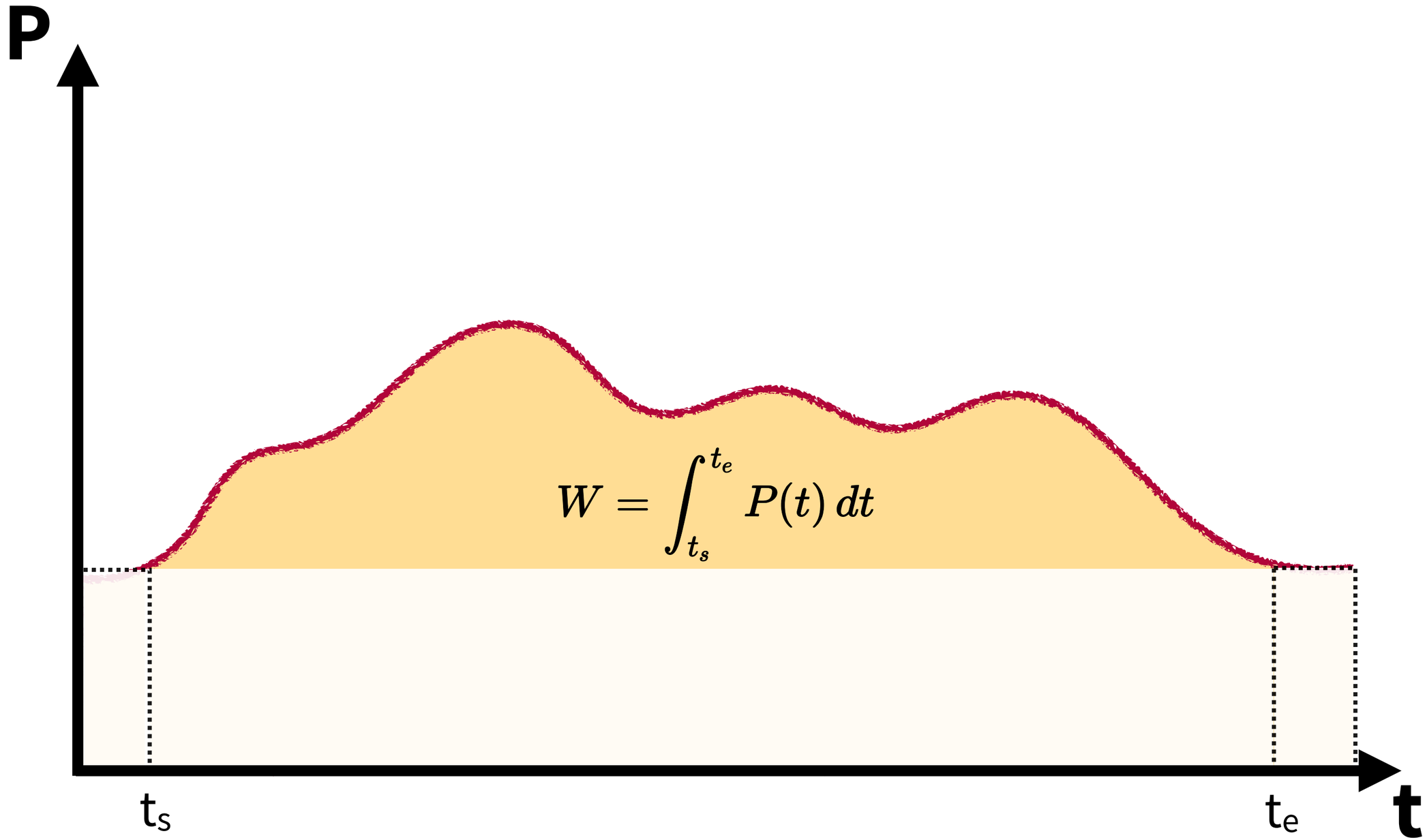


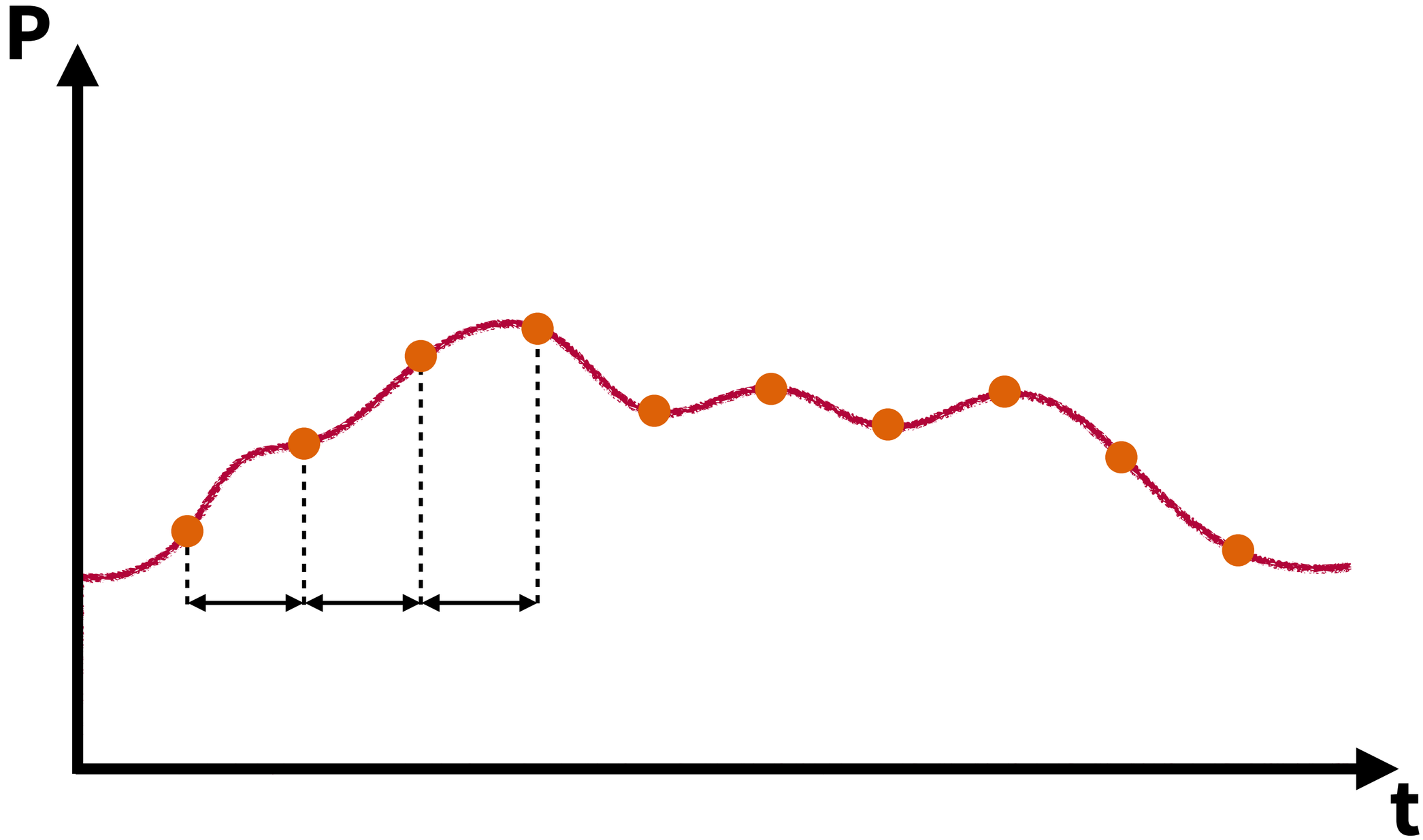
Jetson TX2: CPU, iGPU

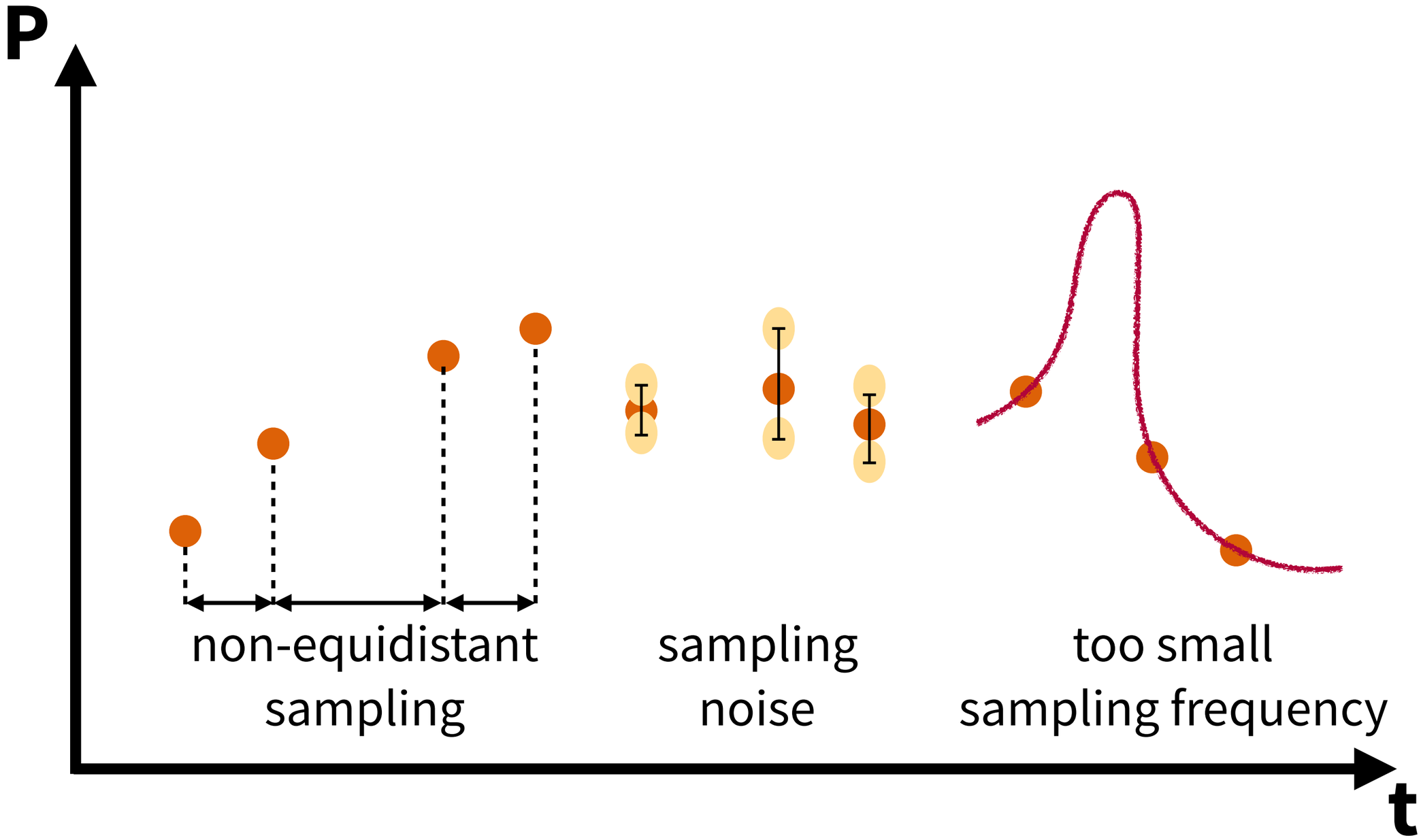


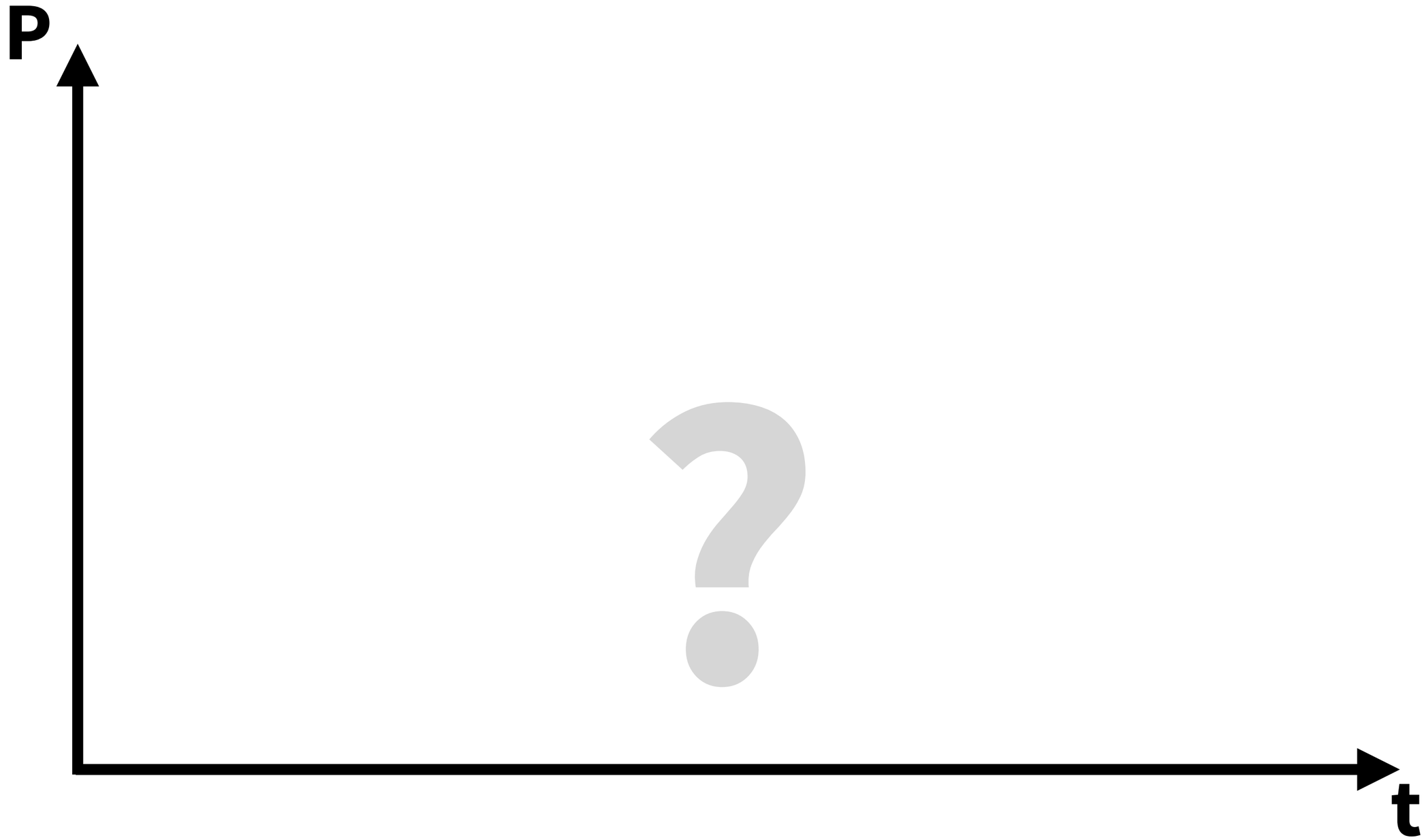
S824L: 2x CPU, 2x dGPU, 1x FPGA

Intermission: Power Sampling Basics


P
$W = \int_{t_s}^{t_e} P(t)\,dt$
t_s
t_e
t


P
t


P
non-equidistant
sampling
sampling
noise
too small
sampling frequency
t


P
?
t

3

Measuring Power & Energy

Measurement Facilities

On-Board

i.e., methods that are integrated with the individual hardware platform

BMC

IPMI

RAPL

NVIDIA Jetson TX2

External

i.e., methods that use standalone devices

LTC2991

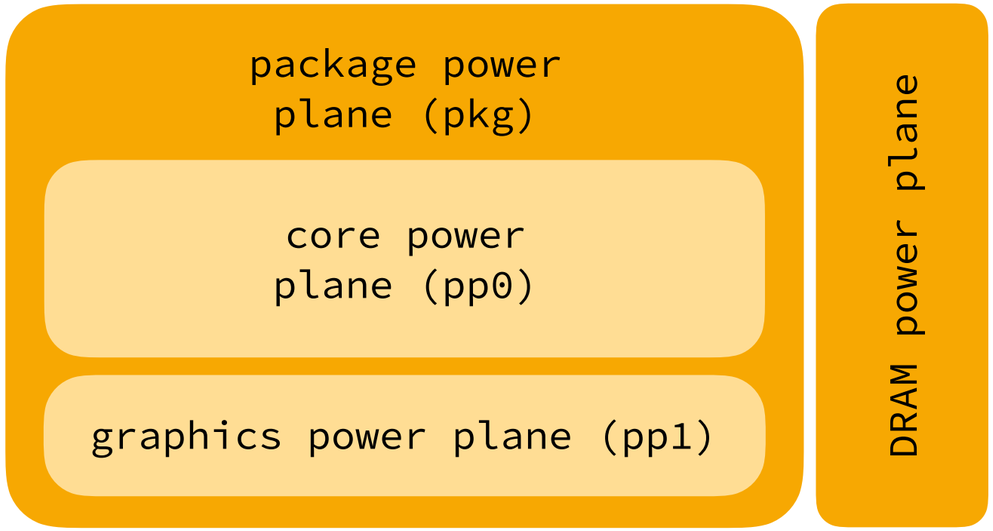
MeasureAlot

INA260

MCP29F511N

Running Average Power Limit (RAPL)

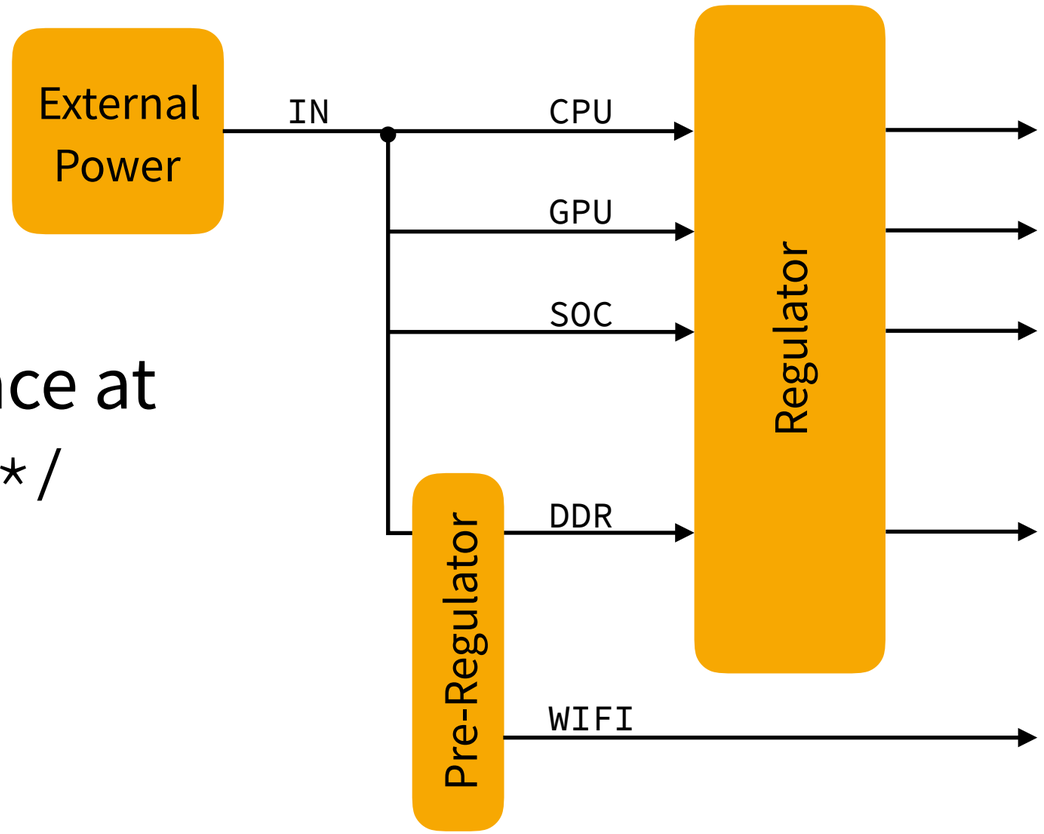
- Available for Intel platforms, since Sandy Bridge
- Registers capture cumulative energy consumption (not power draw), at ~1 ms resolution (wrap around after ~60s)
- Accessible via MSRs, Linux `sysfs`, or `perf_event_open`
- Semi-compatible AMD implementation since Ryzen Gen 3



Intel. 64 and IA-32 Architectures Software Developer's Manual (Volume 3). Acc. 2020-06-12
https://software.intel.com/content/dam/develop/public/us/en/documents/325384- sdm- vol- 3abcd.pdf

NVIDIA Jetson TX2

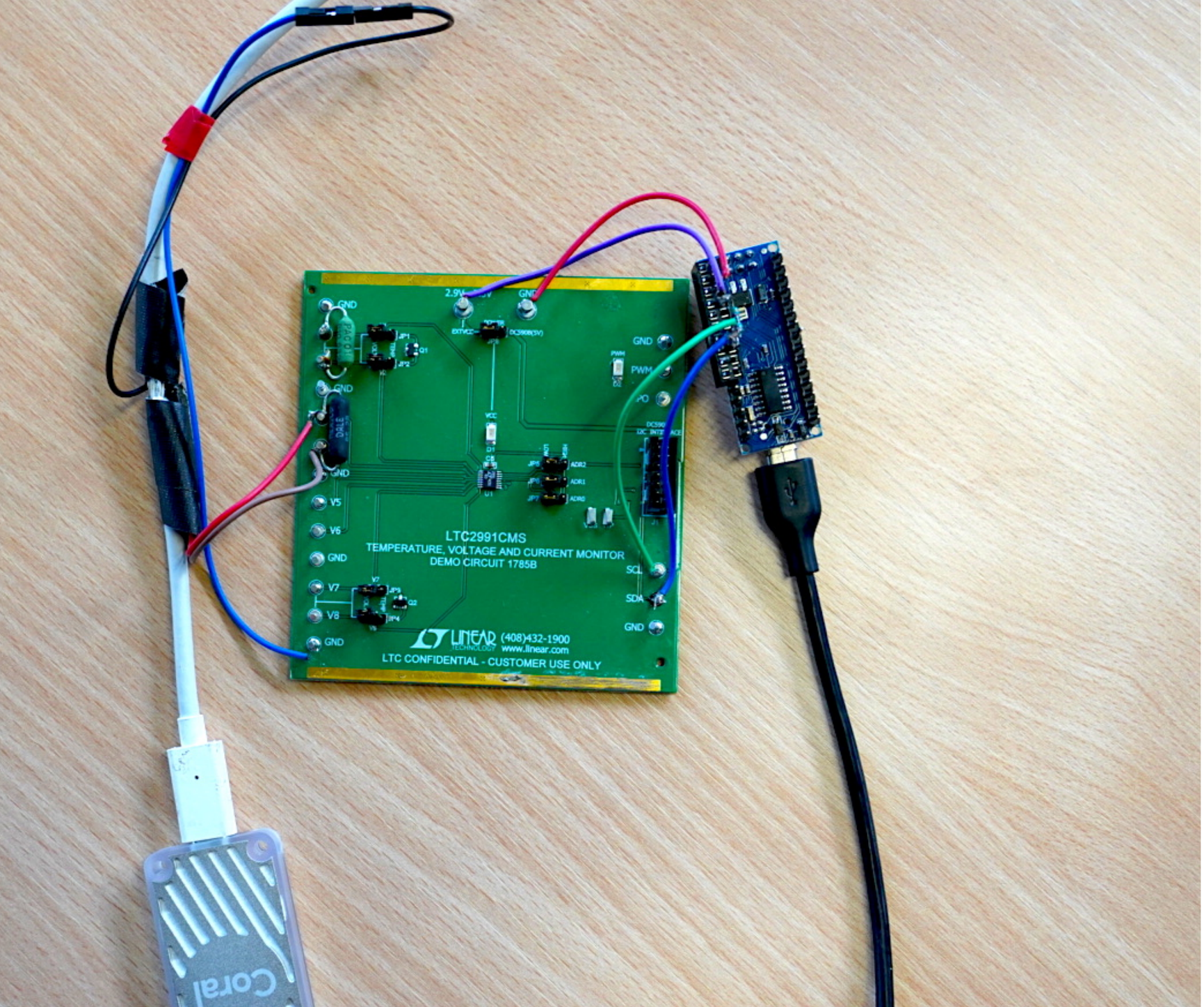
- Two triple-channel INA3221 power monitors
- report averaged power draw, voltage and current
- estimated 5% sample accuracy, 20 Hz sampling frequency
- I²C exposed via Linux `sysfs`-interface at `/sys/bus/i2c/drivers/ina3221x/*/iio_device/in_power`



NVIDIA. Jetson TX2 Series Thermal Design Guide. Acc. 2020-08-25.
http://developer.nvidia.com/embedded/dlc/jetson- tx2- series- thermal- design-guide

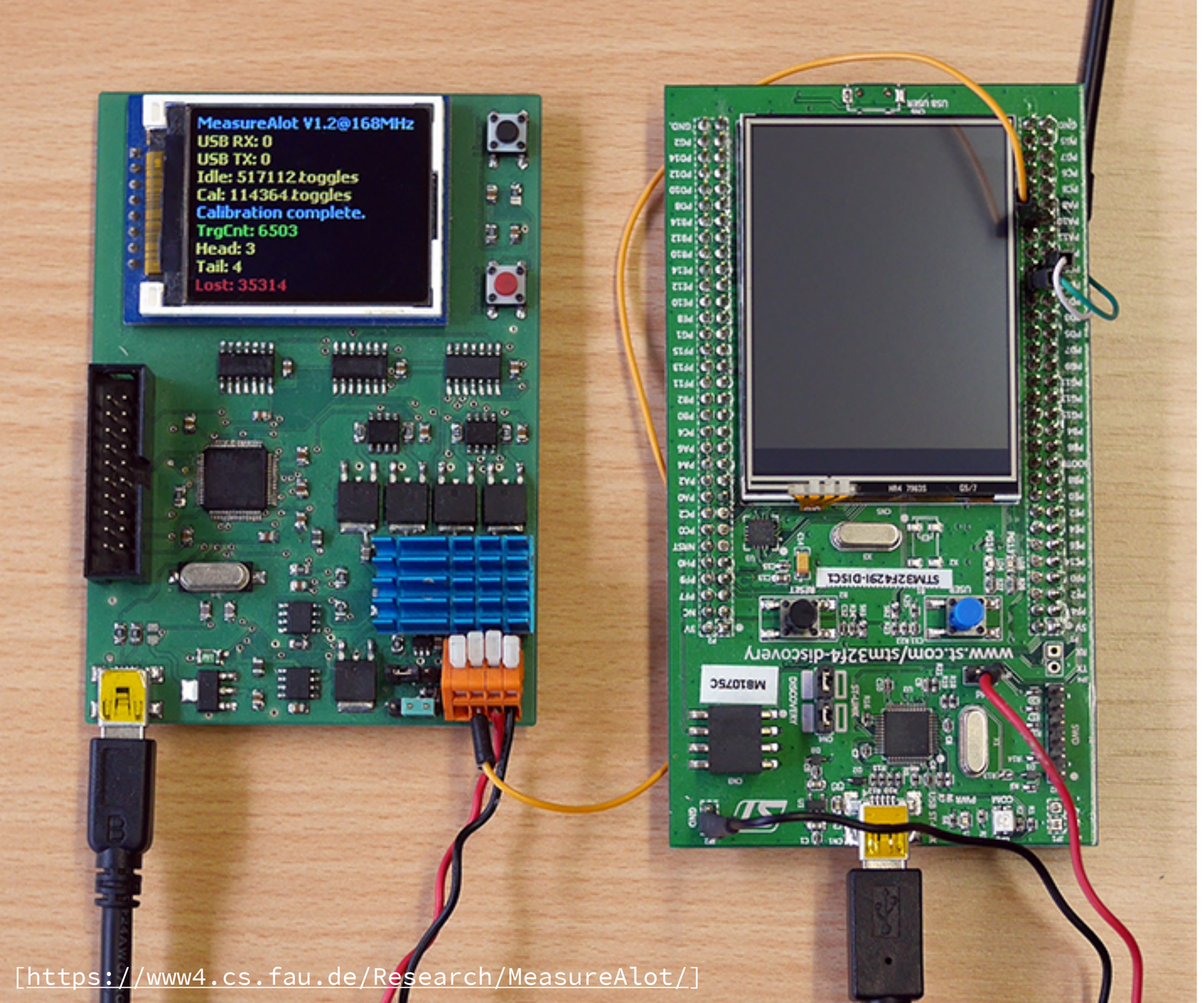
Microchip MCP29F511N

- Dual-channel power measurement device
- Intercepts a system's power supply, exports data via USB
- 200-240 Hz sampling rate (phase-locked to line frequency)
- Up to **15 A** at **230 V**, 24-bit ADC, 0.5 % accuracy



LTC2991

- Four-channel monitor for voltage, current, and temperature
- **3.3-5 V** operating voltage
- 250 Hz sampling rate
- 14-bit ADC,
  1 % accuracy
- I²C data transfer
  (right: custom build USB adapter)



[https://www4.cs.fau.de/Research/MeasureAlot/]

MeasureAlot

- Measures energy demand between two points in time (no power sampling)
- Two capacitors alternatingly charge & discharge, number of rounds is reported
- Energy Demand = Rounds * Energy to charge one capacitor
- up to **1 A** at **5V**

Hardware Platform	Measurement Facility						
	Jetson	RAPL	IPMI	M.Alot	LTC2991	INA260	MCP.
Coral USB	—	—	—	—	✓	✓	—
FRDM-KL02Z	—	—	—	✓	✓	✓	—
SAMA5D3	—	—	—	—	✓	✓	—
Odroid-C2	—	—	—	—	✓	✓	✓
Jetson TX2	✓	—	—	—	—	✓	✓
Desktop PC	—	✓	—	—	—	—	✓
HPE BL460c	—	✓	✓	—	—	—	—
IBM S824L	—	—	✓	—	—	—	✓

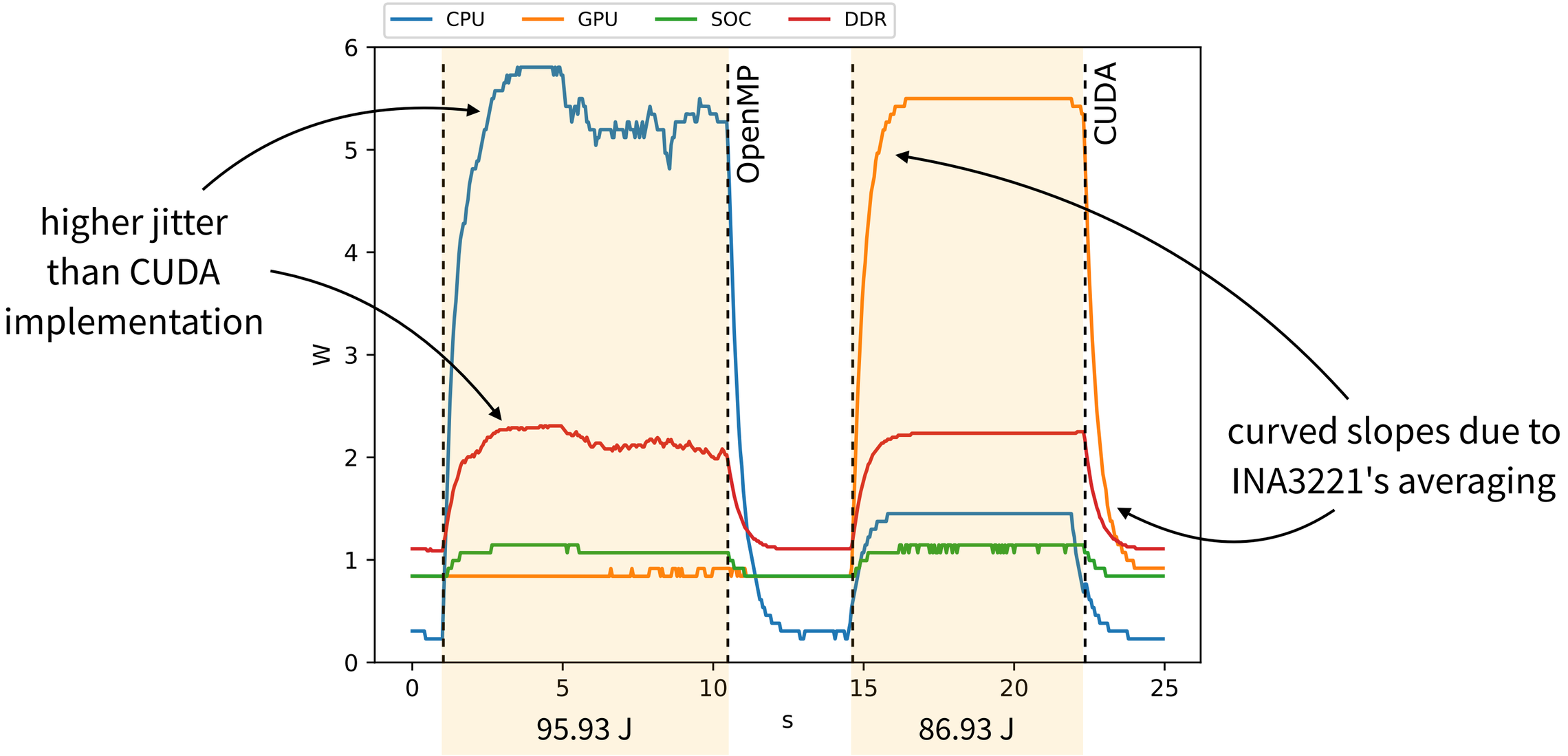
No measurement facility can be used for all hardware platforms.

We introduce PINPOINT to unify interfaces.

4

Evaluating Power Behavior with PINPOINT

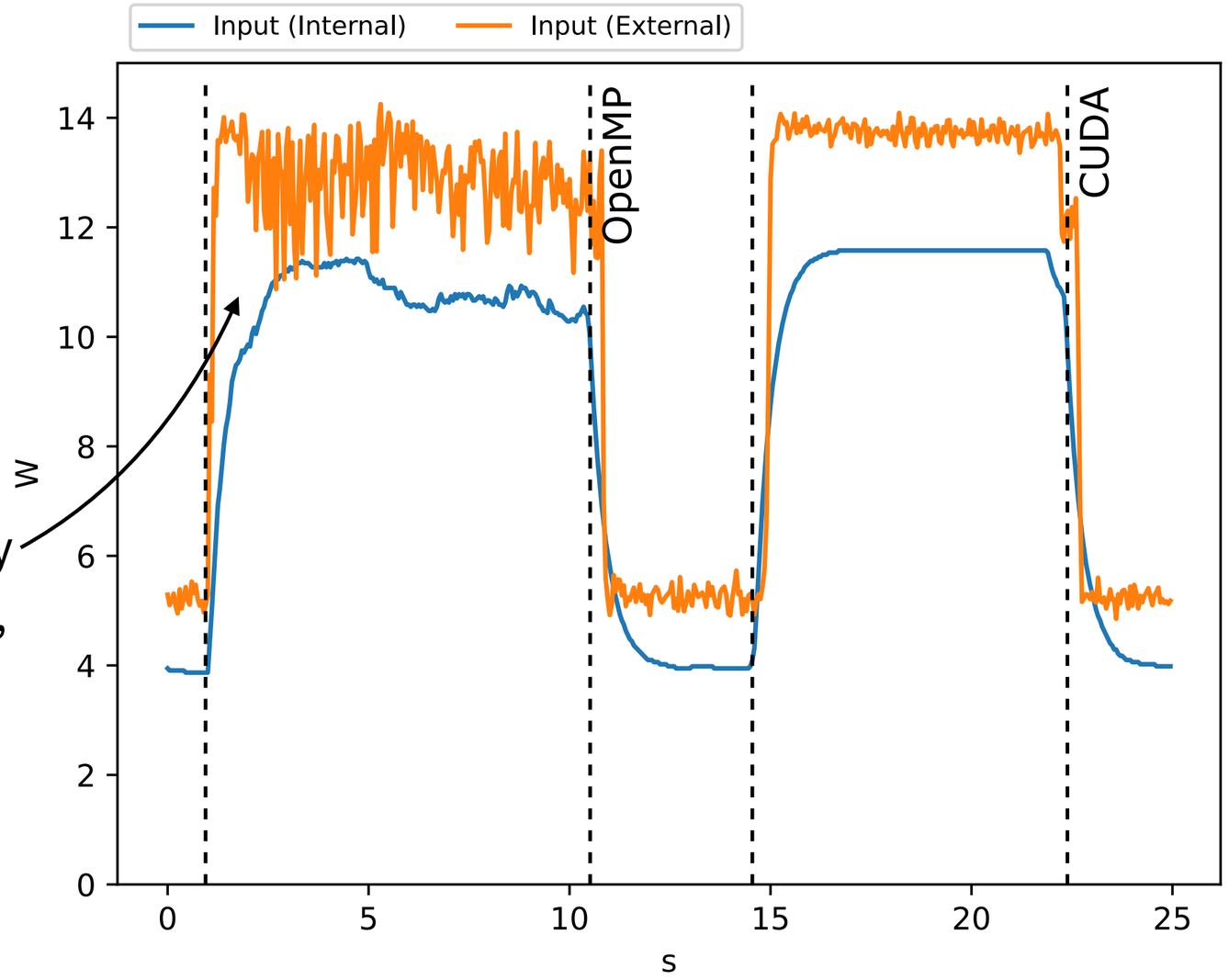
Comparing different workload implementations



OpenMP vs CUDA implementation of heatmap simulation on NVIDIA Jetson TX2. Jetson counters, sampling at 50ms.

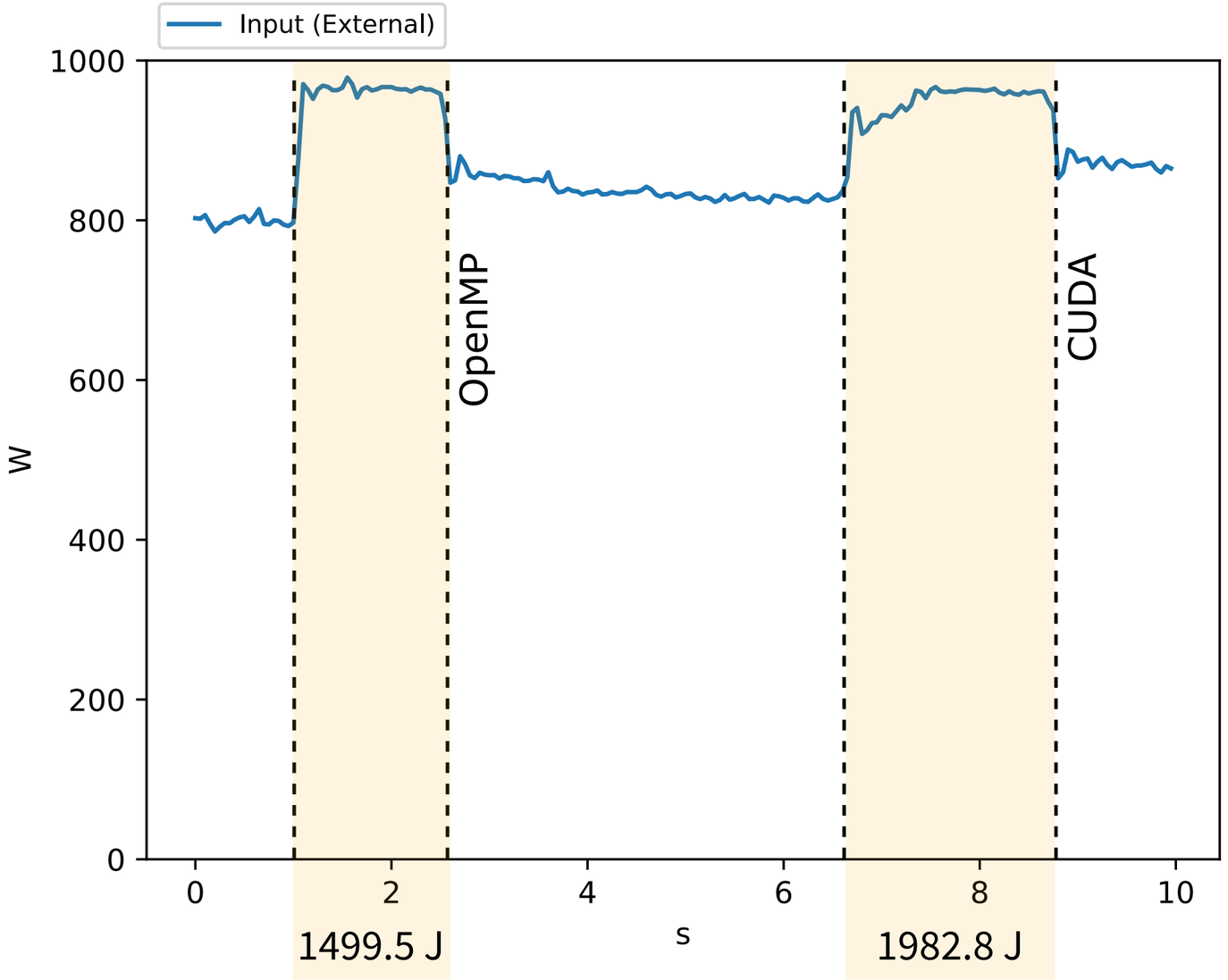
Comparing different measurement facilities

~2.5 W discrepancy (carrier board, fan, power supply)



OpenMP vs CUDA implementation of heatmap simulation on NVIDIA Jetson TX2. Jetson & MCP counters, sampling at 50ms.

Comparing different computing platforms

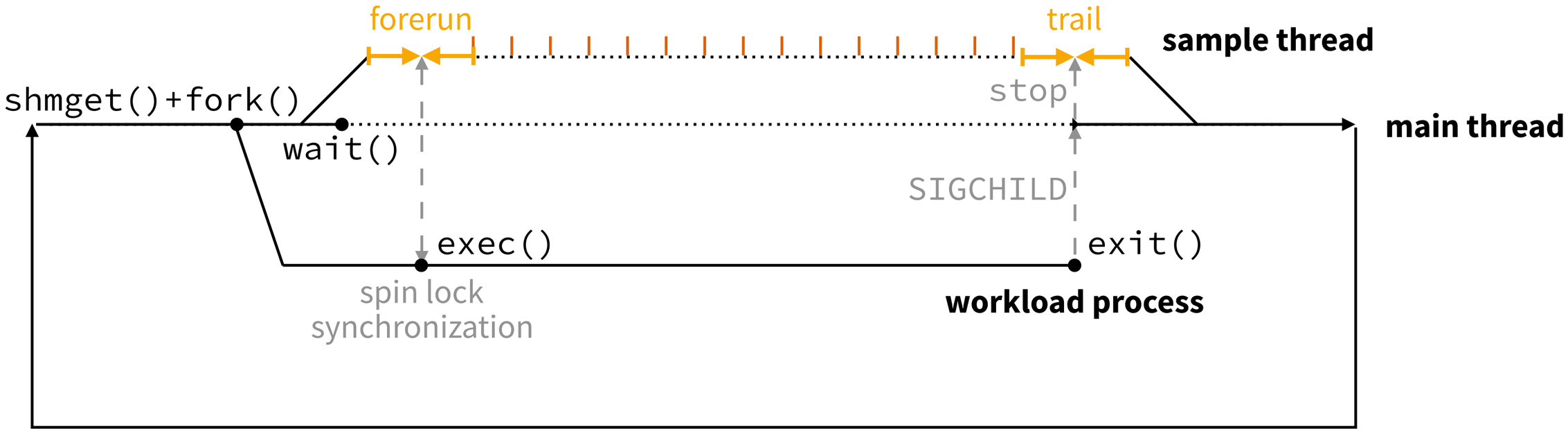


OpenMP vs CUDA implementation of heatmap simulation on IBM S824L, sum of four MCP channels, sampling at 50ms.

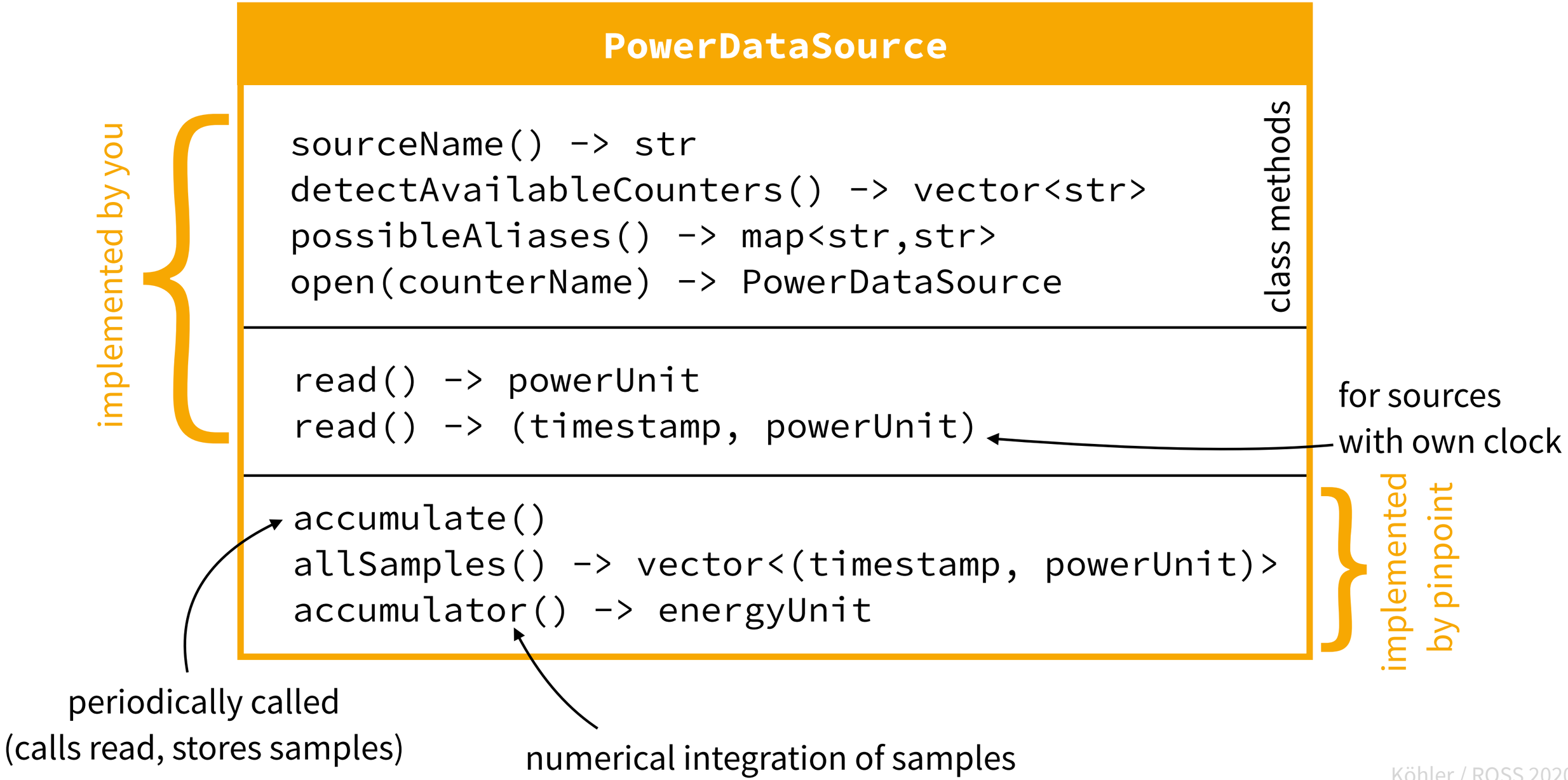
<demo>

5

Imple-mentation


forerun
trail
sample thread
shmget()+fork()
stop
main thread
wait()
SIGCHILD
exec()
exit()
spin lock
synchronization
workload process

How to add your counters to pinpoint



With PINPOINT we can now evaluate how the power and energy demand of a workload change

… over time
… when configured differently
… when implemented differently
… when ported to another platform

`https://github.com/osmhpi/pinpoint`