POET: A Platform for O-RAN Energy Efficiency Testing

N. K. Shankaranarayanan, Zhuohuan Li, Ivan Seskar, Prasanthi Maddala
Rutgers University WINLAB,
NJ, USA

Sarat Puthenpura, Alexandru Stancu
Aether (previously ONF)
CA, USA

Anurag Agarwal
Cognizant
Bengaluru, India

***Abstract*—This paper presents a platform for measuring, evaluating and modeling the energy efficiency aspects of an O-RAN 5G wireless network. We describe our open-source based O-RAN testbed which includes both bare-metal and Kubernetes network functions, in addition to physical network components. We focus on measuring power consumption of servers and workloads of cloudified and virtualized network functions. We show that a combination of different power measurements can be used successfully to achieve the accuracy and granularity required for energy efficiency measurement, evaluation and modeling.**

Keywords—O-RAN, Open RAN, Energy Efficiency, Test Methodology

I. INTRODUCTION

The energy consumption of mobile networks is a significant and growing concern for mobile network operators, especially with plans for dense 5G/6G deployments based on Open RAN technologies. Energy efficiency (EE) and energy savings (ES) use cases are a priority for mobile network operators, thus driving ongoing standardization discussions in the O-RAN Alliance and 3GPP. Energy efficiency testing, metrics, and estimation models provide the foundation for solutions for monitoring and optimizing energy usage. In this paper, we present our vision and platform for researching energy efficiency of O-RAN based 5G networks and provide results from our initial tests with emphasis on energy consumption measurements.

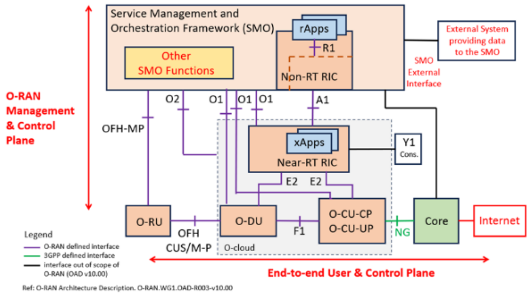
II. ENERGY EFFICIENCY IN THE CONTEXT OF O-RAN

A. Wireless Network Energy Efficiency

Improving wireless network energy efficiency not only helps mobile network operators (MNOs) in reducing operational costs but is also vital for environmental sustainability. The first step towards this is to measure energy consumption accurately with sufficient granularity, then use those measurements to model energy efficiency, and finally use the models to optimize energy efficiency without compromising network performance.

B. O-RAN Architecture and Standardization

Fig. 1 shows a high level O-RAN Architecture diagram. Two key aspects of O-RAN networks – disaggregation and virtualization – both have a major impact on energy efficiency. The energy profile characteristics of physical network functions (such as O-RU), virtualized and cloudified network functions (such as O-CU, O-DU, Core) and the supporting infrastructure (O-Cloud) are all different. Understanding and profiling the power consumption characteristics and the interplay under various network and traffic conditions is thus very important. There are several ongoing discussions to standardize tests and features in the EE/ES area in the O-RAN Alliance [1-4]. A good amount of foundational work has been done but there is ongoing work on topics such as: details of energy metrics to be supported for cloudified NFs (O-RAN WG6), gaps in overall EE/ES specifications (O-RAN SuFG), and EE/ES test specifications (TiFG) [2-4]. An example of energy savings test methodology is covered in the Network Energy Savings section 9.3 of the O-RAN TIFG End-to-End Test Specification [3]. This focuses on network energy savings with carrier and cell switch on/off, while assuming that accurate energy measurements are available.



Ref: O-RAN Architecture Description. O-RAN.WG1.OAD-R003-v10.00

Fig. 1. O-RAN Architecture

III. ENERGY EFFICIENCY TESTING PLATFORM

We describe our platform for O-RAN energy efficiency testing (POET) in this section and the testbed in section IV. We describe: (1) the methodology used for testing energy efficiency, (2) metrics and KPIs for energy consumption and efficiency, and (3) models for estimating energy consumption and efficiency.

A. Energy Testing Methodology

Our broad objective is to research the best methodology for measuring energy consumption, energy efficiency, and energy savings. While energy measurement is not a new field, there is relatively little in terms of standardized methodology, especially when addressing cloudified NFs. Our objective is to determine the test methodology rather than compare the efficiency and performance of available equipment. We intend to explore different methods as part of our work, including:

(a) Testing under different user traffic and load scenarios, and different radio scenarios (e.g., frequency bands, channel bandwidth, path loss, MIMO modes, etc.)
(b) Testing different types of O-RU (e.g., indoor/outdoor,) and server architectures (e.g., with and without accelerators)
(c) Automating the collection of power and performance metrics and test metadata, and reporting of results

B. Energy Efficiency Metrics

The metrics and KPIs for O-RAN energy consumption and energy efficiency are covered in specifications from O-RAN [1-4]. and these refer back to 3GPP and ETSI specifications [5-9].

This work was done as part of a Test R&D project funded by U.S. NTIA Public Wireless Supply Chain Innovation Fund Grant #06-60-IF008

For Energy Efficiency, we essentially have a high-level relationship where *EE* (Energy Efficiency) is given by:

$$EE = \frac{\text{Desired network performance}}{\text{Energy consumed in relevant portion of network}}$$

For the energy consumption (denominator) we plan to measure power consumption and energy consumption using the multi-pronged approach described in section III.C. Physical Network Functions (PNF, e.g. O-RU), Virtualized/Cloudified NF (VNF/CNF, e.g. O-DU, O-CU) require different approaches. In all cases we plan to measure the energy supplied by the power supply to an NF (e.g. $E_{NF,PS}$) and compare that to various types of internal estimated/reported measures of energy consumption ($E_{NF,INT}$), and any standard/pre-standard measures reported on the northbound interfaces ($E_{NF,NB}$). We expect partially-available information and seek to develop models to estimate energy consumption based on available measures. For VNF/CNF, the internal estimates are based on estimates of CPU usage, trained models and server power consumption, and this is an ongoing area of research and standardization.

The numerator for the EE metric can be specified in different ways depending on the objectives of the whole system as given in 3GPP TS 28.554 [6]. We expect our initial focus to be on energy efficiency based on PDCP SDU data volume as per 3GPP TS 28.554, clause 6.7.1 [6] which is given by:

$$EE_{MN,DV} = \frac{\text{Total SDU data volume (bits)}}{\text{Energy consumed by participating network elements (Joules)}}$$

The desired network performance is a “loaded term”, and depending on the context, its definitions can vary. For example, MNOs broadly use throughput & cell-edge performance as well as accessibility & retainability to measure performance. We believe that there are relevant variations in network state which may not be reflected in the energy efficiency metrics currently included in the 3GPP specifications. For example, the Data Volume based EE metric may be influenced by factors such as the channel efficiency (e.g., MCS distribution), fairness criteria (e.g., cell-edge performance), and traffic mix (e.g., video versus file transfer). Our plan is to collect performance measures below and explore correlations and ideas relevant to energy efficiency:

- Number of UEs / RRC connections per cell
- DL and UL PRB utilization per cell
- DL and UL throughput (Mbps) per cell and per UE
- DL and UL data volume (bytes) per cell and per UE
- Latency: per UE and aggregate across UEs per cell
- MCS value: per UE and aggregate across cell
- Characteristics of UE traffic mix

C. *Energy Consumption and Performance Measurement*

The collection of energy, power, and performance metrics is a central aspect of our platform. Fig. 2 shows our multi-pronged approach for collection of power and performance measurements:

1. As depicted in Fig. 2 with label 1, the northbound management and control interfaces for the O-RU, O-DU, O-CU (OFH-MP, O1, E2) can be used to collect network performance metrics and self-reported power/energy metrics. We rely on this for performance KPIs and await more standardized support for energy metrics.
2. For NFs hosted on O-Cloud servers, the power/energy metrics is estimated by the cloud infrastructure. This is an ongoing area of standardization, and we expect these power metrics (label 2) to be provided by the O-Cloud IMS and DMS over the O2 interface to the SMO.
3. While the actual power supplied by the physical equipment (servers, physical O-RU) may be outside the scope of the O-RAN specification, it is a critical ground truth measurement which is necessary for energy efficiency testing. This metric (label 3) can be gathered by using monitored PDUs (power distribution units).
4. In a controlled test environment, we will also have end-to-end performance measurements from the UEs and test servers (e.g., iPerf).

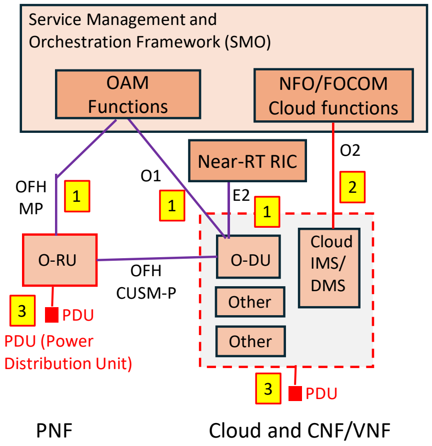


Fig. 2. Multi-pronged O-RAN EE/KPI metric collection approach

The full specification and available support for O-RAN interfaces is still evolving. For the metrics listed in items 1 and 2 above, we expect to use a combination of standardized and proprietary methods. For example, we use E2-based metrics and a pre-standard O1-like measure for performance KPIs for label 1 items, and methods based on IPMI, Kepler, Scaphandre for label 2 items (see details in section IV.B). One of the objectives of our Test R&D project is to use results from research with pre-O-RAN interfaces to provide feedback and insight into testing methodology and contribute to O-RAN for improvement of interface and test case specifications.

D. *Overall Energy Consumption Modeling*

We plan to use our test results to develop models to estimate the power consumption of O-RAN networks, including:

1) Models to provide estimates based on available energy and performance metrics when a given deployed system may not support all the latest self-reported energy metrics.
2) Models which can be calibrated with initial lab-based tests of equipment which are later deployed in the network

3) Models to provide input for rApp/xApp ES applications.

E. *Role of Open Source*

Our testbed leverages open-source components to implement end-to-end O-RAN scenarios. Open-source software is free and openly accessible, thus enabling teams to replicate results in a common reference framework. This aligns with our project objective of contributing our results to influence the industry.

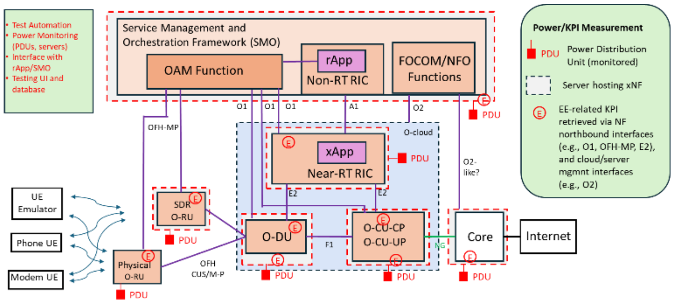


Fig. 3. O-RAN energy-efficiency test system architecture

IV. POET TESTBED

A. *System Description*

The industry trend is to have a Kubernetes cloud implementation of the O-Cloud and Core, as feasible. O-RAN NF modules may also be hosted on bare-metal servers, especially to achieve performance requirements for the O-DU. Accordingly, our testbed includes bare metal servers as well as a Kubernetes cluster for the O-DU, O-CU, and Core. Fig. 3 depicts our test system architecture leveraging the knowhow and capabilities of the COSMOS testbed [10]. The testbed includes physical O-RUs as well as SDR-based RUs. The Kubernetes system includes a FlexRIC near-RT RIC, Kepler power monitoring and support for SDR RU and simulated RAN. The testbed includes an Amarisoft UE-emulator and modem/phone UEs. Power supplied to all components is monitored using smart PDUs (Server Technology PRO3X and STV-6521V.) Power and performance metrics are captured and exported to Prometheus and displayed on Grafana dashboards. Our initial implementation results are based on software from OAI and we plan to also use software from ONF/LF, srsRAN, and O-RAN Software Community. The Kubernetes servers are Dell R730xd (Intel Xeon E5-2680 v4 2.4GHz, 56 CPUs) while the bare-metal servers are Dell R740 (Intel Xeon Gold 6226 2.70GHz, 24 CPUs) and Penguin Relion 1900 (Intel Xeon E5-2620 v4 2.10GHz, 32 CPUs).

B. *Power Measurement*

Measuring the power and energy consumption is at the heart of the energy efficiency testing and validation of energy optimization solutions. We have used the following different approaches for power measurement:

(a) PDU: The power, current, voltage supplied to PNFs and servers is obtained by querying the PDUs regularly (nominally 10 secs) and exported to Prometheus/Grafana. This provides a critical ground-truth measurement of power consumption.

(b) IPMI: Power and environment variables reported by the servers are also monitored using the Intelligent Platform Management Interface (IPMI). Queries are made to the server Baseboard Management Controller (BMC) and exported to Prometheus/Grafana.

(c) Scaphandre [11]: We deployed the Scaphandre open-source energy monitoring functionality on bare-metal servers running OAI software. Scaphandre measures process utilization (based on Intel Running Average Power Limit (RAPL)) and estimates the power consumption using an estimation model.

(d) Kepler [12]: The Kubernetes deployment in the testbed uses Kepler exporters on each server sending power metrics to Prometheus/Grafana. Kepler is an open-source energy monitoring functionality for Kubernetes systems. It measures node and container utilization (based on Intel RAPL) and estimates the power consumption using an estimation model.

C. *O-RAN Performance KPI Measurement*

Energy efficiency testing and energy optimization monitoring require accurate and convenient performance KPI measurements such as the ones listed in section II.B. We have the following approaches for KPI metrics:

(a) E2E: We use iPerf end-to-end application to obtain the data volume, throughput, latency of a session between the UE and a test server in the packet data network.

(b) O1-based KPI: We use an “O1-like” approach using a telnet-based solution which provides uplink/downlink throughout, and downlink PRB load for OAI DU [13]. This is a preliminary version of ongoing open-source work to implement an O1 interface for OAI CU/DU.

(c) E2-based KPI: The E2 interface from the O-CU and O-DU to the Near-RT RIC and xApps can provide several KPIs using the E2-SM KPM. Our Kubernetes-based O-RAN system includes a KPM xApp on FlexRIC with an E2 interface to the OAI CU/DU.

O-RAN energy testing is currently a very active research area [14-15]. Our primary focus is to develop test methodologies and estimation models to provide a foundation for optimizing network energy consumption.

V. RESULTS

In this paper, we present our initial energy test results from the multi-year project we have initiated. Here we focus on the power measurement capability, with some O-RAN power and KPI measurements to demonstrate the testbed capability. In future papers, we hope to show more details of O-RAN performance and power consumption.

A. *PDU and IPMI Power Measurements*

We first present comparisons of power measurement utilities, using the PDU measurement as the ground truth. We deployed the *stress* utility function to vary the CPU load and compare the power consumption measurements from the four approaches in IV.B. Fig. 4 shows a Grafana panel with the PDU (green), IPMI (yellow) and Kepler (blue) power consumption measures increasing with CPU load when the stress utility was used to add CPU-load in steps of 8CPUs every 240secs up to 40CPUs (the machine has 56CPUs). There is a saturation at 330 watts a little beyond the 32CPU stress level. We take the PDU power to be the ground truth of total power consumed by the server. The power reported by IPMI tracks this fairly well except for some granularity and is lower by an average of 10 watts. The IPMI curve shows the effect of longer averaging and takes about 60 secs to stabilize.

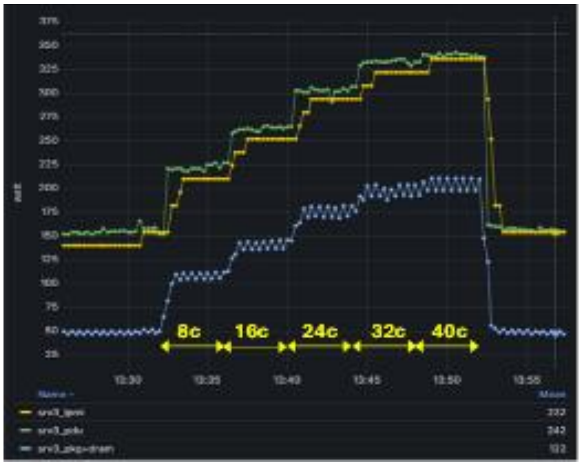


Fig. 4. Characterization and comparison of power measurements from PDU (green), IPMI (yellow), and Kepler (kepler_node) utility. Server load request increases from 8CPUs to 40CPUs, in steps of 8CPUs every 240secs.

Fig. 5. Top: Power measurements from PDU (green), IPMI (yellow), and Kepler for small and large increases of CPU load. Bottom left/right: Current and voltage from PDU and IPMI, showing the relatively coarse granularity of the IPMI metrics.

The IPMI measure shows a granularity in steps of about 14 watts. The reason for this is that the voltage and current measures are noticeably quantized. Fig. 5 shows another stress test with small as well as large step increases in CPU-load. We can see the three (PDU, IPMI, Kepler) power measures in the upper figure and the current and voltage measures for PDU and IPMI in the lower two figures. The current granularity of 0.2A and voltage granularity of 1.0 V leads to the granularity in the IPMI power measurement. The behavior of the PDU, IPMI, and Kepler metrics are similar to the Watt-meter, IPMI, and RAPL observations in [16]. We conclude that IPMI is a good estimate of the total power supplied to the server by the PDU. This has major ramifications on the ability of O-RAN networks to track the total power consumption without requiring an external power measurement. The IPMI metrics should be available in most modern servers and can be included in the cloud management system specifications (e.g., O2-IMS/DMS).

B. *RAPL-based Process Power Measurements*

We have used both Kepler and Scaphandre for estimating the power consumed by O-RAN network functions as well as other supporting processes. Both make use of Intel RAPL which can have different levels of support in different servers. The servers we used supported most of the RAPL measures, but do not support the more-recent RAPL PSYS domain. The sum of the power reported by the Kepler node metrics in Fig. 4 (*kepler_node_package* + *kepler_node_dram*) is about 120-to-130/110-to-120 watts below the PDU/IPMI powers. We attribute this to baseline power drawn by the server fan and other elements other than the Kubernetes containers. This offset can be calibrated at zero load and used to adjust the Kepler estimates. We have similar results from a CPU load test for both types of the bare-metal servers while measuring power using Scaphandre using *scaph_host_power*, *scaph_socket_power*, *scaph_process_power* and *scaph_domain_power*. We observed that *scaph_host_power* was the sum of *scaph_socket_power* and *scaph_domain_power* (expected behavior when the PSYS domain is not supported, and similar to the Kepler results) Compared to the PDU power, the *scaph_host_power* is lower by about 95 (Relion1900) and 195 (DellR740) watts for the two servers. The Dell R740 is more powerful and has GPU cards which accounts for the larger overhead.

C. *O-RAN Power and Performance KPI Measurements*

We explored the use of both Scaphandre and Kepler to zoom in and filter on various containers, processes, namespaces to obtain the power estimates for O-RAN related workloads. This includes the O-DU, O-CU, Core, as well as estimates of other supporting O-RAN related functions in the Kubernetes system. We focused on O-RAN deployments with rfsim simulated UEs for convenience, and also verified the same behavior for an OAI bare-metal O-RAN deployment with an N310 SDR gNB. Future work will explore physical O-RU and other O-RAN scenarios as described in section III.

Fig. 6 shows Scaphandre power measurements and the O1-like KPIs from an all-OAI bare-metal deployment (rfsim) for an iPerf throughput test, first with one UE and then with two UEs. For one and two UEs respectively, the DU PRB Load is 70% and 100%, and the aggregate DU Downlink (DL) throughput is 65 Mb/s (1 UE) and 84 Mb/s (42 Mb/s per each of two UEs). We can see the corresponding changes in power consumed using scaphandre metrics for: (a) *scaph_host* (green), *scaph_* *socket* (yellow), *scaph_domain_dram* (blue) (b) total power of all *processes* (orange), (c) power for OAI O-RAN/Core processes (red) (O-DU, O-CU, Core), (d) power for the simulated UE process (blue). The total PDU power is about 95 watts higher than the *scaph_host* power. The total O-RAN/Core power correlates with the total DU DL PRB utilization and throughput. Besides the O1-like kpis, we have also used the FlexRIC kpm xApp to obtain similar E2-based KPIs from an OAI CU/DU.

Fig. 6. Scaphandre-based power measurements and O1-like performance KPIs for a bare-metal OAI O-RAN system and iperf sessions with 1 and 2 UEs.

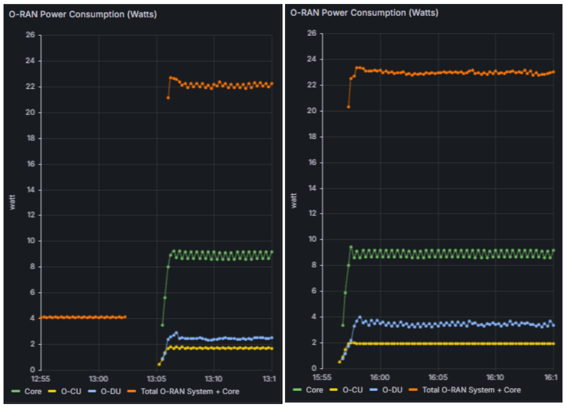


Fig. 7. Kepler-based power measurements for a multi-node Kubernetes O-RAN system with 1 UE and 2 UEs supported by the DU/CU/Core.

Fig. 7 shows Kepler-based power measurements for the Kubernetes O-RAN system. The figure shows the use of *kepler_container* power metrics filtered by the workloads for the OAI Core (green), O-CU (yellow), O-DU (blue). Initially there is no Core/CU/DU deployed, and the overall O-RAN system consumes 4.13 watts. Then the Core/CU/DU is deployed, with first 1 UE and then 2 UEs supported by the DU/CU/Core. The total O-RAN system power consumptions increase to 22.2 and 22.9 watts for 1 UE and 2 UEs respectively. The power change to support one extra UE is small (as expected). The takeaway from Figs. 6 and 7 is that we are able to measure the system and NF powers at sufficient granularity. This will enable us to explore various methods in future work.

VI. Conclusion & Future Work

We have presented POET, our platform for O-RAN energy efficiency testing, and a multi-year project plan to develop energy efficiency test methodology, metrics, and models using our open-source based testbed. The PDUs provide a ground-truth accurate measurement of total power supplied to equipment. We find that IPMI offers a very good estimate of the PDU measures. For individual workloads on servers (e.g., O-DU, O-CU, Core), we show that we can use Kepler for Kubernetes O-RAN system and Scaphandre for a bare-metal O-RAN system to get measurements of sufficient granularity. We can measure the increase in O-RAN NF power consumption which correlates with the increase in DU PRB utilization and throughput.

We will continue executing our Test R&D project plan to determine the best energy test methodology and estimation models which can be used to improve network energy optimization. In order to validate and improve our models, we aim to pursue collaborations (especially with other NTIA-funded labs) to gather results from different network deployment scenarios to refine our models.

References

[1] O-RAN.WG1.Network-Energy-Savings-Technical-Report-R003-v02.00: O-RAN Network Energy Saving Use Cases Technical Report 2.0, June 2023

[2] O-RAN.WG6.O-Cloud-ES-v01.00: O-RAN Study on O-Cloud Energy Savings 1.0, June 2024

[3] O-RAN.TIFG.E2E-Test.0-R003-v06.00: O-RAN End-to-end Test Specification 6.0, June 2024

[4] O-RAN.SuFG.NES-Analysis-v00.00.02: O-RAN Sustainability Focus Group; Energy Measurements Analysis Report, May 2024

[5] 3GPP TS 28.310: "Management and orchestration; Energy efficiency of 5G".

[6] 3GPP TS 28.554. Management and orchestration; 5G end to end Key Performance Indicators (KPI)

[7] 3GPP TS 28.552: Management and orchestration; 5G performance measurements.

[8] ETSI TS 103 786 Environmental Engineering (EE); Measurement method for energy efficiency of wireless access network equipment Dynamic energy performance measurement method of 5G Base Station (BS). (Note: Annex C defined traffic mix)

[9] ETSI ES 203 228, Environmental Engineering (EE); Assessment of mobile network energy efficiency.

[10] Cloud Enhanced Open Software Defined Mobile Wireless Testbed for City-Scale Deployment (COSMOS): https://www.cosmos-lab.org/

[11] Scaphandre - https://github.com/hubblo-org/scaphandre

[12] Kepler - https://sustainable-computing.io/

[13] https://gitlab.eurecom.fr/oai/o1-adapter

[14] V. Gudepu et al., "EARNEST: Experimental Analysis of RAN Energy with Open-Source Software Tools," 2024 16th International Conference on COMmunication Systems & NETworkS (COMSNETS), India, 2024.

[15] V. Gudepu, R. R. Tella, C. Centofanti, J. Santos, A. Marotta and K. Kondepu, "Demonstrating the Energy Consumption of Radio Access Networks in Container Clouds," NOMS 2024-2024

[16] Kavanagh R, Djemame K. Rapid and accurate energy models through calibration with IPMI and RAPL. Concurrency Computat Pract Exper. 2019; 31:e5124. https://doi.org/10.1002/cpe.5124