JACET

J. ADV COMP ENG TECHNOL, 6(3) Summer 2020

A Method for Measuring Energy Consumption in IaaS Cloud

Derdus M. Kenga[1], Vincent O. Omwenga[2], Patrick J. Ogao[3]
1,2- Faculty of Information Technology, Strathmore University Nairobi, Kenya.(derduskenga@gmail.com)
3- School of Computing and Information Technology, Technical University of Kenya, Nairobi, Kenya

Received (2020-04-24) Accepted (2020-07-15)

Abstract: *The ability to measure the energy consumed by cloud infrastructure is a crucial step towards the development of energy efficiency policies in the cloud infrastructure. There are hardware-based and software-based methods of measuring energy usage in cloud infrastructure. However, most hardware-based energy measurement methods measure the energy consumed system-wide - including the energy lost in transit. In an environment such as the cloud, where energy consumption can be a result of different components, it is important to isolate the energy, which is consumed as a result of executing application workloads. This information can be crucial in making decisions such as workload consolidation.*

In this paper, we propose an experimental approach of measuring power consumption as a result of executing application workloads in IaaS cloud. This approach is based on Intel's Running Average Power Limit (RAPL) interface. Application workload is obtained from Phoronix Test Suite (PTS)' 7zip and aio-stress. To demonstrate the feasibility of this approach, we have described an approach, which can be used to study the effect of workload consolidation on CPU (7zip) and data (aio-stress)-intensive application workloads over power performance by varying the number of Virtual Machines (VMs) in a server. Power is measured in watts. Performance of CPU (as a result of processing CPU intensive workloads) in Performance is measured in Million Instructions per Second (MIPS) and I/O performance (as a result of processing data intensive) is measured in MB/s. Our results on the effect of workload consolidation has been compared with previous research and was found to be consistent. This shows that the proposed method of measuring power consumption is accurate.

Keywords: *Energy measurement, IaaS cloud, energy efficiency, Phoronix Test Suite, RAPL.*

How to cite this article:
Derdus M. Kenga, Vincent O. Omwenga, Patrick J. Ogao. A Method for Measuring Energy Consumption in IaaS Cloud. J. ADV COMP ENG TECHNOL, 6(3) Summer 2020 : 145-1

I. INTRODUCTION

THE data centers, which run the cloud host thousands of physical servers and consume a lot of energy. Considering the increase in costs of running data centers and addressing the problem of global warming, energy efficiency in a data center is a very important subject [1] [2] [3] [4]. Being able to measure energy usage in the data center servers is an essential starting point towards achieving energy efficiency.

There are two categories of techniques, which are used to measure energy usage [5]; those which make use of software such as Powertop and those which make use of hardware devices such as Wattmeters and Power Distributing Units (PDUs) [6]. PDUs must be connected between the physical server and Uninterrupted Power Supplies (UPS) [7]. A PDU is just a smart meter, which measures current at a time interval and multiplies it with voltage to get the watt value. The problem with hardware-based methods of energy measurement is that they monitor all the power that goes into the physical server thus is unable to isolate energy consumed as a result of executing application workloads and the one lost on transit.

On the other hand, software-based methods of energy measurements use computationally obtained power models to accurately infer the power consumption by a physical server [8]. Software-based techniques involve collecting data on server power consumption patterns and analysis, which leads to the creation of a power consumption model that estimates power from server data. Example of data that can be used to generate a power consumption model includes power consumption of each computing resource for idle server, the power consumption of each computing resource at different server utilization levels, the power consumption of each computing resource at peak utilization, the power consumption of each computing resource at different levels of utilization of the different computing resources and power consumption of each computing resource when processing different types of resources.

The energy measurement technique we propose in this paper is software-based and combines other tools for generating application workload and measuring the performance of specific components of a physical server. For instance, we can generate a CPU-intensive workload, measure CPU performance, and measure power consumption over time. The technique is based on Intel's Running Average Power Limit (RAPL) [9] [10]. We have used Phoronix Test Suite (PTS) as a source of application workloads [11]. PTS is an open-source benchmarking tool with a collection of test profiles with different but known behaviors. For instance PTS has a microbench application known as 7zip [12]. 7zip is a file compression application using 7-zip compression, which puts a maximum load on the CPU, and hence it is referred to as CPU-intensive workload. PTS also has microbenchmarks targeting other computing resources such as I/O. For instance aio-stress is an asynchronous I/O benchmark created by SuSE [13]. The current profile (version 1.1.1) uses a single thread and constantly reads and writes a 2048 MB test file and a 64KB block size from and to the hard disk. For this reason, the system subcomponent that is being tested is the I/O. PTS has been used successfully in testing computer performance, hardware validation, and as a source of cloud workload [8] [14] [15].

The main contributions of this work are:

1. A software-based method for measuring power consumption in cloud data center server.

2. A method for showing the effect of server consolidation on power consumption and power performance on CPU and data intensive workloads.

3. An assembly of tools for automating application microbenchmarks orchestration and measuring of CPU and I/O performance.

The remainder of this paper is organized as follows: in section two, we present related works. Section three describes our experiment setup. In section four, our experimental results and discussion are presented, and the paper is concluded in section five.

II. RELATED WORK

The main challenge in monitoring energy usage in cloud-based systems for purposes of improving energy efficiency is the reliance on hardware PDU to give accurate results. A software-based solution is required to be able to measure even energy consumed by a process [5]. Software-based alternatives require that there exists software to measure energy as well as a suitable application workload to simulates application execution [5].

In [16], the authors describe an assembly of components used to measure and compare energy consumption across three different virtualization technologies namely, Docker, Kubernetes, and virtual machines. To trigger the processes in the systems under investigation, they have used a workload generator, JMeter with MOODLE portal, to simulate incoming user activity into the system. Power is then measured using powermeter, which measures the consumption of electrical energy via information from the CPU state and usage. Although the authors report success in showing that similar workloads executed in the different virtualization technology consume different amounts of energy, the workability of their software-based powermeter is not clear.

In [9], the authors have carried out extensive experiments to investigate the accuracy, performance, and granularity of Intel's RAPL energy measurement of a physical server. In the experiments, they have used Stream, Stress-ng, and ParFullCMS application-level micro-benchmarks as a source of application workloads. In their analysis, the authors have leveraged a production-level power measurement dataset from the Taito to make their analysis realistic [17]. The results from their experiments show that RAPL is accurate in predicting server energy consumption. Similar results have been reported in [18]. Unfortunately, the execution of the micro-benchmarks has been done manually.

In [19], the authors have proposed an architecture for profiling energy usage in virtual machines in a cloud infrastructure. The infrastructure is based on the fact that different types of application workloads dominate particular types of computing resources. For instance, data-intensive tasks would depend more on the disk storage thus requiring high I/O bandwidth and on the other hand, computation-intensive tasks use more of the processor. The proposed architecture in this work consists of two main components, namely Resource Monitoring Unit (RMU) and Energy Profiling Unit (EPU). RMU is designed to measure the energy consumed by the physical machines using WattsUp watt meters, which is then sent to Zabbix. Zabbix is the monitoring infrastructure that has been used in the proposed architecture [20]. VM energy apportioning is done in the EPU component, which uses a set of algorithms. When testing the set-up, Cloud9 has been used as the source of cloud workloads. The main drawback of the proposed architecture is the use of a hardware-based method to record energy usage by the physical servers. Besides, the setup is complex.

Energy consumed by datacenter servers can also be measured to create a power prediction model for servers. For instance, in [21], the author collects power consumption data from three components of a server (processor, memory, and network interface card), which is used to create a server power consumption model for predicting energy usage at different levels of server utilization. The power model is created by analyzing the data using Lasso model linear regression. Power is measured by running a bash script using ipmitool tool, which reads power consumption directly from the power supply units while the performance of the components such as the CPU is measured using perf, a UNIX utility. Three different applications are used as a source of application workloads. A C program, which allocates large blocks of memory and fills them with 1 is executed to create load on memory. A load on the processor is created by running a tress tool in the server, which spawns threads spinning on the sqrt(). Finally, to simulate a network load, iperf tool is used. The model created is reported to predict server energy consumption with 5.33% median error.

The abovementioned related works have successfully measured energy consumption by different entities such as servers, virtual machines, or processes, for different purposes using different components. However, there are weaknesses such as using hardware-based power measurement approaches [19], manual execution of benchmarks applications [9], use of non-standard benchmark applications [21] and complexity [19]. Table 1 shows a summary of the approaches power measurement the related works.

III. EXPERIMENT SETUP

For our experiment, we have used an HP ProLiant DL380 server. The server (Physical Machine (PM)) is equipped with one Intel Core i7-4790 at 4.00GHz CPU, which has 8 cores and 8192 KB cache size. The RAM is DDR4 of size 12 GB. The server runs Ubuntu Linux OS 18.04 and 4.15.0-36-generic (x86_64) kernel version. We have used KVM hypervisor to create 8 virtual machines (VMs) in the server. Each VM is equipped with 1 vCPU at 3.59GHz, 1280MB of memory, and 45 GB of disk storage. Each VM runs a similar OS to that of the PM. It is in these VMs that we execute application workloads to simulate application load.

TABLE 1
A SUMMARY OF RELATED WORKS

Reference	Power measurement approach	Source of application workload	Weakness
[16]	Software-base	JMeter and Moodle – a fully working application	-Workload orchestration is manual -Workability of their software-based powermeter is not clear.
[9]	Software-based	Stream, Stress-ng, and ParFullCMS – all are microbenchmark applications.	-Workload orchestration is manual
[19]	Mixed – both software-based and hardware-based	Cloud9	-Proposes a complex setup -Relies upon the hardware-based approach of power measurement
[21]	Software-based	A C-program, C programming language's sqrt() and iperf.	-Workload orchestration is manual -The C-program used is written by the author and there is no proof that it is tested to ensure it gives the desired results.

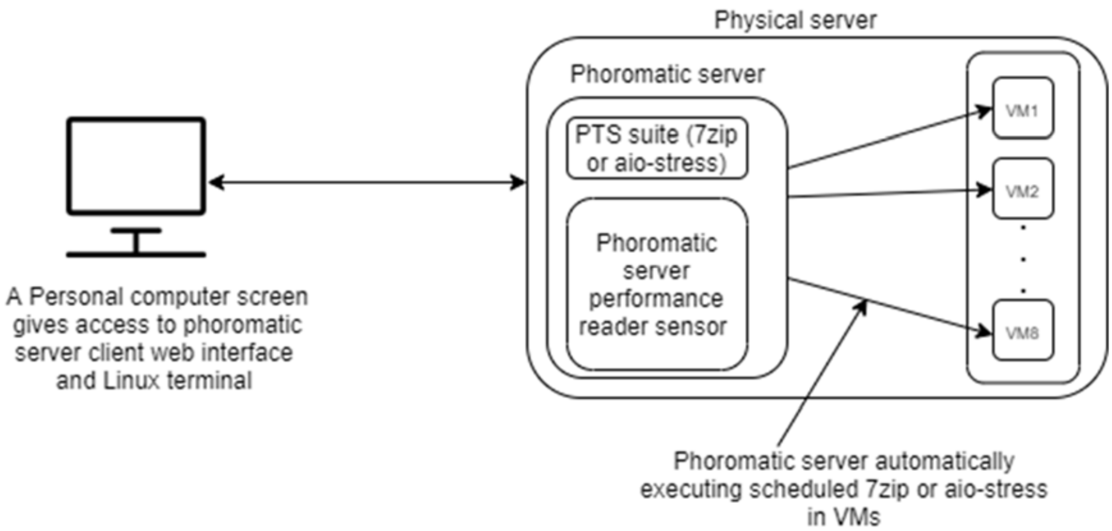

Fig 1: Experiment Setup

We execute 7zip microbenchmark application (version 1.7.0) in the VMs to create a load on the CPU. 7zip is installed in each of the 8 VMs and can be executed by running a command in the Linux terminal [22]. However, to automate the process of executing 7zip in the VMs, we have used Phoromatic Server (PS) [23]. With PS, we schedule a time when 7zip is executed in 1 VM, 2VMs, 3VMs upwards till 8 VMs (number of VMs represents the level of consolidation). With each execution, PS collects the performance of the CPU for each VM. Performance is measured in Million Instructions per Second (MIPS). Fig. 1 shows how we have set up our components. PS has a web client with an easy to use graphical user interface, which we have used to schedule the execution of 7zip in the VMs. In each execution, we repeat the measurements (performance measurement) 3 times to check for consistency. By default, PS repeats an execution 3 times. Hence each execution is repeated 9 times and the average vCPU performance for each active VM is recorded. We have computed and recorded the total CPU performance by adding up vCPU performance of each active VM. We have also computed and recorded the average CPU performance by diving the total CPU performance by the number of active VMs. The whole procedure ie repeated with aio-streass microbenchmark instead of 7zip. Aio-stress is a data intensive workload thus it put a load on I/O. I/O perfoamnce is measured in Megabits Per Second (MB/s).

To collect power consumption, we have a bash script, which we have scheduled to run at the same time as 7zip or aio-stress. As mentioned earlier, 7zip or aio-stress is scheduled and automatically orchestrated via PS. The bash script runs powerstat command, which reads power consumption by the physical server via RAPL. Additional statistics such as standard deviation (Std Dev.) is computed by powerstat. The bash script runs 3 times and each run takes 60 seconds. In each run, 60 samples of power consumption (60 samples at 1-second intervals) are captured from RAPL. Thus, in each execution of 7zip or aio-stress in a set of VMs, there are 180 samples. We have computed and recorded the average power consumption of the server. Power is measured in Watts. We have also represented energy efficiency as Performance per Watt (PPW) by dividing the total CPU or I/O performance by power consumption at different levels of consolidation. To demonstrate the feasibility of the proposed approach, we have recorded the standard deviation of power measurement readings as well as plotted a chart showing how VM consolidation affects server power consumption and power performance for both 7zip and aio-stress. To check the accuracy of our outcome, we compare it with previous research in power measurements, which uses hardware-based methods, which is considered as ground truth [10].

IV. RESULTS AND DISCUSSION

Table 2 & 3 summarizes the data collected from the experiment conducted. We recorded the average CPU and I/O performance, average power consumption, power consumption measurement Std Dev. and computed server PPW at different levels of VM consolidation for both 7zip and aio-stress microbenchmarks. When no VM (zero VM) is running in the server, the server still consumes power (Table 2). This is called idle power. The use of PPW as a preferred metric for measuring energy efficiency is because it shows the amount of IT load that is processed per watt of power drawn.

Fig. 2 shows the effect of increasing the number of VMs (VM consolidation) on power consumption and power performance when using 7zip microbenchmark. It can be observed that as the number of resident VMs increases in the server, the amount of power consumption increases but flattens at some point (No. of VMs = 6). It is also observed that an increase in the number of resident VMs, decreases the average vCPU performance per VM. Fortunately, the power performance increases. This is the advantage of server consolidation. On the other hand, Fig. 3 shows the effect of increasing the number of VMs (VM consolidation) on power consumption, power performance and I/O performance when using aio-stress microbenchmark. It can be observed that an increase in consolidation (No. of VMs), does not cause a significant increase in power consumption. However, there is degradation in I/O performance, which in turn causes a slight decrease in power performance. The decrease in

TABLE 2

SERVER POWER CONSUMPTION, CPU PERFORMANCE AND POWER PERFORMANCE AT DIFFERENT LEVELS OF CONSOLIDATION. CONSOLIDATION IS REPRESENTED BY THE NUMBER OF VMS EXECUTING 7ZIP.

No. of VMs running 7zip	Total CPU performance (MIPS)	Average vCPU performance (MIPS)	Average server power consumption in Watts (w)	Power consumption measurement Std Dev.	CPU performance Per Watt (MIPS/w)
0	0	0	7.88	1.35	0
1	3737	3737	23.39	2.25	159.77
2	6658	3329	32.55	2.57	204.55
3	9647.01	3215.67	40.61	2.21	237.55
4	11940	2985	46.48	1.76	256.88
5	14450	2890	41.58	2.19	347.52
6	15637.0002	2606.17	48.014	1.62	325.68
7	17629.997	2518.57	49.26	1.88	357.9
8	16570	2071.25	51.085	1.34	324.36

TABLE 3

SERVER POWER CONSUMPTION, I/O PERFORMANCE, AND POWER PERFORMANCE AT DIFFERENT LEVELS OF CONSOLIDATION. CONSOLIDATION IS REPRESENTED BY THE NUMBER OF VMS EXECUTING AIO-STRESS.

No. of VMs running 7zip	Total I/O performance (MB/s)	Average I/O performance (MB/s)	Average server power consumption in Watts (w)	I/O performance Per Watt (MB/s/w)
1	57.57	57.57	11.71	4.92
2	30.76	15.38	9.07	3.39
3	31.71	10.57	9.645	3.29
4	41.4	10.35	10.36	4
5	44.95	8.99	9.008	4.99
6	39.3	6.55	9	4.37
7	28	3.9	9.45	2.96
8	31.84	3.98	9.67	3.29

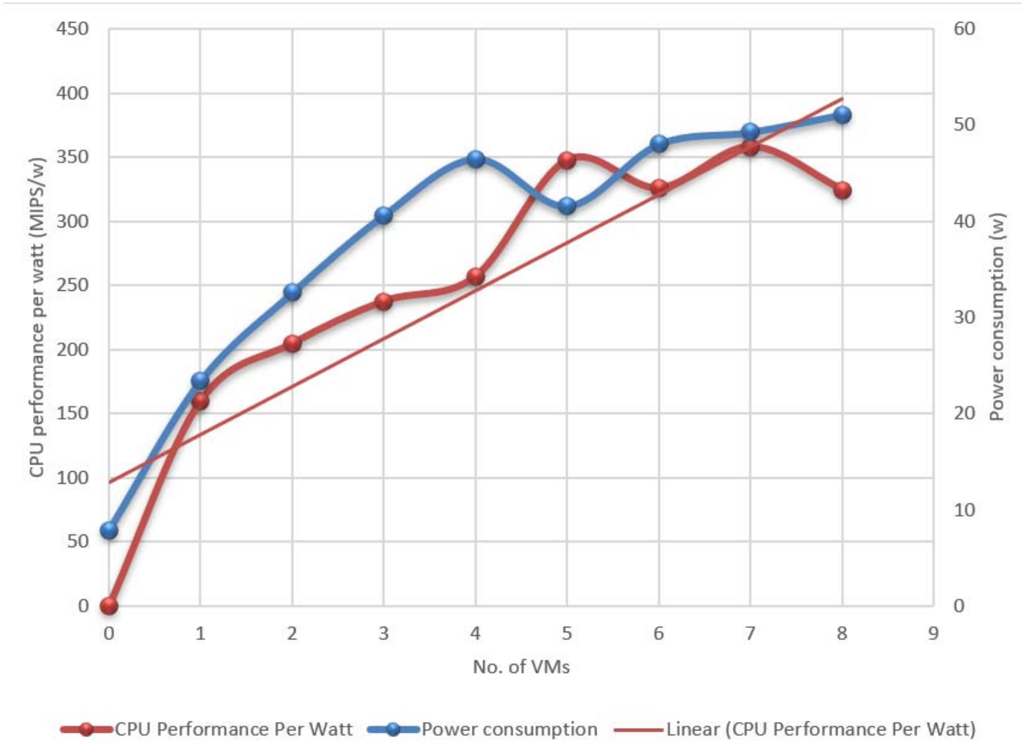

Fig. 2: The effect of increasing the number of VMs executing 7zip on power consumption and power performance

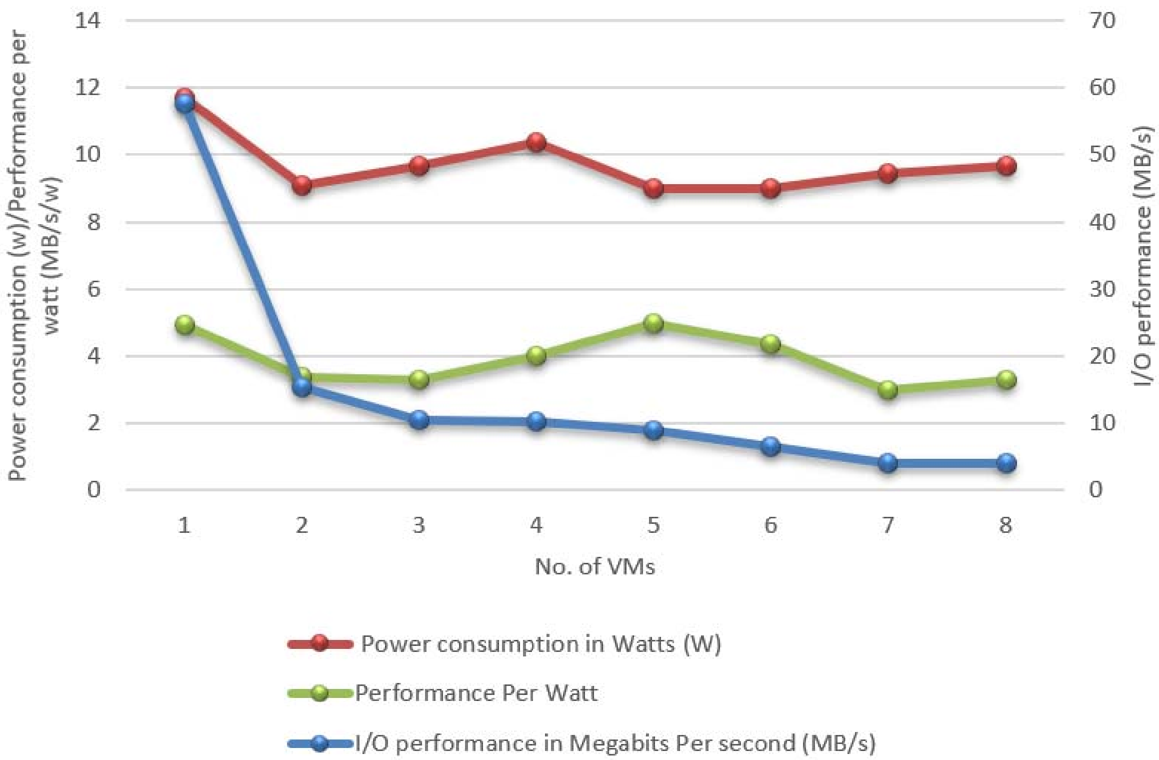

Fig. 3: The effect of increasing the number of VMs executing aio-stress on power consumption, power performance and I/O performance

I/O performance results from more VM instances competing for the physical resources as well as the hypervisor capacity resulting to performance loss. Thus, it can be concluded that workload consolidation limiting factor for CPU intensive workloads and I/O intensive workloads is power consumption and performance, respectively. This shows that different types of workload will behave differently from a power consumption and performance perspective.

To conclude that our conclusions are accurate, we have compared them with those of previous research work, where power measurement was undertaken by using physical powermeter, which is considered as ground truth. Table 4 summarizes the comparisons. Therefore, we can conclude that our approach to measuring power consumption is accurate. This conclusion is also supported by the low power consumption measurement Std Dev.

TABLE 4
PREVIOUS RESEARCH WITH SIMILAR OUTCOMES AS OURS, BUT HAVING USED HARDWARE-BASED POWER MEASUREMENT APPROACH, WHICH IS CONSIDERED AS GROUND TRUTH.

Outcome	Reference
For CPU intensive workloads, increase in number of VMs in a server results in a linear increase in power consumption to a given point, then flattens.	[24] [25]
Different workload types behave differently. Thus, workload consolidation limiting factor for CPU intensive workloads and I/O intensive workloads is power consumption and performance, respectively	[26], [19]
Although power consumption is higher with increase in consolidation, power performance is better	[7]

These results are similar to those obtained in [24] [27] [25] and [26]

Although we report an advantage of aggressive consolidation, our work has not investigated the side effects of increased consolidation. For instance, in [28], the authors discovered that at high levels of VM consolidation, there is interference between co-resident workloads in multitenant virtual environments, which leads to poor Quality of Service (QoS). Thus, our work does not authoritatively conclude that aggressive server consolidation is necessarily good – there should be a trade-off between aggressive server consolidation for energy efficiency and QoS. Furthermore, our work has only used two type of workload i.e. CPU and I/O intensive workload. Thus, other types of workloads such as memory-intensive and network intensive workloads may behave differently.

V. CONCLUSION

In this paper, we have proposed a simple and accurate approach that can be used to measure power consumption in a server. The approach is composed of a workload generator, which executes workloads automatically and reading power consumption of a server via RAPL via powerstat command. To show the feasibility of this approach, we have described an experiment, which can be used to study the effect of workload consolidation on CPU and I/O intensive application workloads. The results of our work are similar to those of previous research work. We have thus concluded that our approach is accurate. As future work, we intend to extend our work to investigate how our approach of measuring power performs when used on other types of workloads.

REFERENCES

1. I. Salam, R. Karim and M. Ali, "Proactive dynamic virtual-machine consolidation for energy conservation in cloud data centres," Journal of Cloud ComputingAdvances, Systems and Applications, vol. 7, no. 1, 2018.

2. D. Sara and M. Hicham, "An Adaptive Autonomic Framework for Optimizing Energy Consumption in the Cloud Data Centers," International Journal of Intelligent Engineering and Systems, vol. 12, no. 4, pp. 111-129, 2019.

3. I. Azlan, "Energy-driven cloud simulation: existing surveys, simulation supports, impacts and challenges," Cluster Computing - Springer, 2020.

4. Y. Maede, G. R. Akbar and H. F. Mohammad, "emperature and energy-aware consolidation algorithms in cloud computing.," Journal of Cloud Computing , vol. 8, no. 13, 2019.

5. M. Made and B. Behzad, "A Simplified Method of Measurement of Energy Consumptionin Cloud and Virtualized Environment," in International Conference on Big Data and Cloud Computing, Sydney, 2014.

6. A. Noureddine, "Adel Noureddine. Towards a Better Understanding of the Energy Consumption of Software Systems.," Lille University of Science and Technology, Lille , 2014.

7. G. Chaima, "Energy efficient resource allocation in cloud computing Enviroment," Institut National des T´el´ecommunications, Paris, France , 2014.

8. J. Smith and I. Sommerville, "Workload Classification & Software Energy Measurement for Efficient Scheduling on Private Cloud Platforms," in Conference'10 University of St Andrews, 2011.

9. N. K. Kashif, M. Hirki, N. Tapio, N. Jukka and O. Zhonghong, "RAPL in Action:Experiences in Using RAPL for Power measurements," ACM Transactions on Modeling and Performance Evaluation of Computing Systems, vol. 9, 2018.

10. F. Muhammad, S. Arsalan, R. M. Ravi and L. Alexey, "A Comparative Study of Methods for Measurement of Energy of Computing," Energies - mdpi, vol. 12, no. 11, 2019.

11. Phoronix Test Suite, "Phoronix Test Suite - Linux Testing and Benchmarking Platform, Automated Testing, Open-Source Benchmarking," Phoronix Media, 2018. [Online]. Available: https://www.phoronix-test-suite.com/. [Accessed 01 February 2020].

12. Phoronix Test Suite, "7-Zip Compression Test profile," Phoronix Test Suite, 2020. [Online]. Available: https://openbenchmarking.org/test/pts/compress-7zip. [Accessed March 01 2020].

13. Phoronix Media, "AIO-Stress test profile," Phoronix Media, 2020. [Online]. Available: https://openbenchmarking.org/test/pts/aio-stress. [Accessed May 05 2020].

14. G. Rishu and T. S., "A systematic review on the various approaches used for achieving Energy consumption in Cloud," Test Engineering and Management, vol. 82, pp. 3936 - 3953, 2020.

15. M. Alexis and M.-M. Vania, "Automatic benchmark profiling through advanced workflow-based trace analysis," Practice and Experience, Wiley, vol. 48, no. 6, pp. 1195-1217, 2018.

16. M. Made and Y. Pujianto, "Evaluating Energy Consumption in a Different Virtualization within a Cloud System," in 4th International Conference on Science and Technology , Yogyakarta, Indonesia, 2018.

17. CSC, "Taito supercluster," CSC, 2020. [Online]. Available: https://research.csc.fi/guides. [Accessed 15 MArch 2020].

18. S. Yu, H. Yang, R. Wang, Z. Luan and D. Qian, "Evaluating architecture impact on system energy efficiency," PLoS ONE, vol. 12, no. 11, 2017.

19. A. Ibrahim, D. Karim, A. Django and K. Richard, "Energy-Aware Profiling for Cloud Computing Environments," Electronic Notes in Theoretical Computer Science, vol. 318, pp. 91-108, 2015.

20. Zabbix LLC, "Zabbix:: The Enterprise-Class Open Source Network Monitoring Solution," Zabbix , 2020. [Online]. Available: https://www.zabbix.com/. [Accessed March 01 2020].

21. T. Makris, "Measuring and Analyzing Energy Consumption," Aalto University, Espoo, Finland, 2017.

22. Phoronix Test Suite , "Phoronix Test Suite - Benchmarking Linux with Phoronix Test Suite - Lots of command line examples," Phoronix Test Suite , 2018. [Online]. Available: https://wiki.mikejung.biz/Phoronix_Test_Suite . [Accessed March 01 2020].

23. Phoronix Test Suite, "Phoromatic: Automated Linux Benchmark Management & Test Orchestration," Phoronix Test Suite, 2018. [Online]. Available: http://www.phoronix-test-suite.com/index.php?k=phoromatic#phoromatic. [Accessed 01 March 2020].

24. C.-Z. Mar, S. Lavinia, O. Anne-Cécile and P. Guillaume, "An experiment-driven energy consumption model for virtual machine management systems," Sustainable Computing: Informatics and Systems, vol. 18, pp. 163-174, 2018.

25. C. HeeSeok, L. JongBeom, Y. Heonchang and E. Lee, "Task Classification Based Energy-Aware Consolidation in Clouds," Scientific Programming , vol. 2016, 2016.

26. A. Mirabel and R. Siddiqui, "Energy Aware Consolidation in Cloud Computing," California State University, 2015.

27. M. David, G. Brian and W. Thomas, "PowerNap: Eliminating Server Idle Power," in Proceedings of

the 14th International Conference on Architectural Support for Programming Languages and Operating Systems, ASPLOS 2009, Washington, DC, USA, 2009.

28. X. Chen, L. Rupprecht, R. Osman, P. Pietzuch, F. Franciosi and W. Knottenbelt, "CloudScope: Diagnosing and Managing Performance Interference in Multi-tenant Clouds," in 2015 IEEE 23rd International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems, 2015.