Towards a General Metric for Energy Efficiency in Cloud Computing Data Centres: A Proposal for Extending of the ISO/IEC 30134-4

Carlos Juiz[1] [a], Belen Bermejo[1] [b], Alejandro Fernández-Montes[2] [c] and Damián Fernández-Cerero[2] [d]

[1]*Department of Computer Science, Universitat de les Illes Balears, Palma, Spain*

[2]*Department of Computer Languages and Systems, University of Seville, Spain*

Keywords: Server Consolidation, Management, Energy Efficiency Metrics, ISO/IEC 30134-4.

Abstract: For some years now, energy efficiency has been one of the concerns of cloud system administrators. To improve energy efficiency, in recent years standards such as ISO/IEC 30134-4 and ISO/IEC 21836 have emerged. Both standards are focused on the evaluation of physical servers, taking into account the power consumed and the maximum peak of performance, under running a SPEC benchmark. In this way, the server consolidation through virtualization is not considered in these standards, being the consolidation of servers as one of the most applied techniques to improve energy efficiency in cloud data centres. This work proposes an extension of the methodology proposed in these standards to measure energy efficiency in consolidated servers. As a result, it has been demonstrated through real experimentation that the proposed generical methodology considers the consolidation of servers in any type of virtualization environment. This methodology helps system administrators to manage cloud data centres and servers more efficiently.

1 INTRODUCTION

Currently, IT represents around 15% of the worldwide emissions of greenhouse gas emissions. There are a large number of contributors to these emissions, but the main contributors to these emissions are cloud data centres. Besides, the use of IT and data centres is increasing day by day since the population is demanding more services through the Internet and cloud. For this reason, the performance demand for data centres is increasing day by day.

Nevertheless, increasing the demand of data centres implies an increment in the power consumption of data centres and, particularly, the servers that compose them. Specifically, the physicals servers are the most power-demanding devices, demanding 50% of the power of the whole cloud data centre. On this point, it is important to highlight that the reduction in power consumption in cloud data centres can be achieved by lowering the server's performance. However, if users' services or applications consume more time to be performed due to lowering the server's performance, the servers will consume more energy. Then, cloud data centres aim to maximize energy efficiency without interfering with performance, finding a trade-off between them.

1.1 Server Consolidation

To achieve better energy efficiency in cloud data centres, the Green IT techniques appeared several years ago, the server consolidation is one of the most important techniques with this aim (Juiz and Bermejo, 2020). The server consolidation is based on allocating the maximum workload possible in the minimum number of physical servers. Then, by grouping the workload, under-used physical servers can be switched off, saving in power consumption.

1.2 ISO/IEC 30134-4

The ISO/IEC 30134-4 is a standard that specifies data centre energy effectiveness KPIs to help cloud data centre operators measure and improve specific aspects of energy effectiveness. Specifically, this standard provides a metric, called ITEEsv, reflecting the energy effectiveness capability of servers (ISO/IEC,

[a] https://orcid.org/0000-0001-6517-5395
[b] https://orcid.org/0000-0002-9283-2378
[c] https://orcid.org/0000-0002-2998-4950
[d] https://orcid.org/0000-0002-9403-111X

Juiz, C., Bermejo, B., Fernández-Montes, A. and Fernández-Cerero, D.
Towards a General Metric for Energy Efficiency in Cloud Computing Data Centres: A Proposal for Extending of the ISO/IEC 30134-4.
DOI: 10.5220/0012707600003711

In *Proceedings of the 14th International Conference on Cloud Computing and Services Science (CLOSER 2024)*, pages 239-247
ISBN: 978-989-758-701-6; ISSN: 2184-5042

2017). ITEEsv defines the method to obtain average energy efficiency for servers. Also, the scope of this metric is reduced to physical servers. It is not considering the server consolidation. Moreover, this metric is calculated based on a SPEC benchmark. Then, this standard does not consider other current benchmarks or workloads, such as CPU or micro-services-based workloads. Besides, the consolidation is not considered in any case in this standard. Consequently, there is no standard considering server consolidation to balance energy efficiency and performance degradation.

1.3 ISO/IEC 21836

The ISO/IEC 30134 series specifies data centre energy efficiency KPIs to help data centre operators measure and improve specific aspects of data centre energy effectiveness. ISO/IEC 30134-4 in particular defines a method to measure the peak capacity and utilization of servers operating in a data centre using operator-selected benchmarks. However, it does not provide a method for comparing individual server energy efficiency across data centres, as stated in ISO/IEC 30134-4. The standard ISO/IEC 21836 provides a server energy efficiency metric (SEEM) to measure and report the energy effectiveness of specific server designs and configurations.

Since the ISO/IEC 30134-4 does not provide a method for comparing individual server energy efficiency across data centres, the ISO/IEC 21836 appears to fill the mentioned gap (ISO/IEC, 2020). Even though the ISO/IEC 21836 provides a methodology to measure energy efficiency in data centres, it is not devoted to virtualization. That is, this standard does not consider the server consolidation server. Then, it lacks one of the main features of current data centres.

2 SERVERS ENERGY EFFICIENCY METRICS

2.1 Energy-Delay Product (EDP)

The Energy-Delay Product (EDP) considers the total energy consumption of cores and the amount of time for executing applications (Gonzalez and Horowitz, 1996). Then, the EDP is calculated as the product of the energy consumption and the response time. It is important to note that the EDP, originally, was created for microprocessors. But it can be also applied to servers, where the energy consumption corresponds to the energy consumed by the physical server performing a workload; and the response time corresponds to the needed time to execute the work.

2.2 CiS^2 Index

The CiS^2 is a metric aiming to quantify the "goodness" of a consolidation configuration (Juiz and Bermejo, 2020). That is, it measures the trade-off between the energy consumption and the performance degradation of a physical server which is consolidating a set of virtual machines (or containers) (Juiz and Bermejo, 2024).

Since the trade-off is based on the comparison between a physical server and the same physical server with virtual machines consolidated, the CiS^2 is calculated as Eq. 1 shows.

$$CiS^2 = SP_p \cdot SP_e \tag{1}$$

Where the SP_p is the speed-up of the performance of consolidating a server, and the SP_e is the ratio of the energy consumption of the physical server and the consolidated one.

Considering this definition, the CiS^2 metric can be applied to any type of server, any type of workload nature, and any virtualization platform.

2.3 ITEEsv Index

ITEEsv is a metric that describes the maximum performance per kW of all servers or a group of servers in the data centre based upon a specification or potential performance of these servers (ISO/IEC, 2017). Then, the server energy effectiveness is a combination of the capacity to do work per unit energy (capability), the amount of time the server is doing work (utilization), and the ability of the server to reduce the energy use when the workload is reduced (power management).

The ITTEsv metric is calculated as Eq. 2 shows.

$$ITEE_s v = \frac{\sum_{i=1}^{n} SMPE_i}{\sum_{i=1}^{n} SMPO_i} \tag{2}$$

Where $SMPE_i$ is the maximum performance of a server i, and, $SMPO_i$ is the maximum power consumption of a server i. For this metric, the maximum performance is obtained from a specific benchmark execution, which is based on transactional workload, being the throughput of the performance metric.

2.4 Relation Between Metrics

In this work, some statements regarding the metrics should be considered:

- The metric of ITTEsv is to provide the relationship between the maximum performance and the maximum power consumption of a physical server. In any case, the virtualization is not considered in this metric.

- The ITTEsv considers a transactional-based workload. Then, if the server workload has a different nature, the metric calculation should be modified. For example, if the workload is CPU-based and the maximum performance is measured in time units, the maximum performance should be calculated as $\frac{1}{\sum_{i=1}^{n} SMPE_i}$(time and throughput are inverse). Besides, for this metric, the maximum power consumption is expressed in kilowatts.
- As the ITEEsv shows, it does not consider energy consumption. It just considers the power consumed by a specific scenario of the physical servers. Considering only the maximum power consumption (or the power consumption) is not realistic at all. For this, considering the energy consumption would be more suitable.
- If we want to use the ITEEsv to compare the physical server and the consolidated one, it is necessary to use the following algebra, where "c" corresponds to the consolidated server and "p" corresponds to the physical one (see Eq. 3).

$$\frac{ITEE_{sv}^{c}}{ITEE_{sv}^{p}} = \frac{\frac{\frac{1}{R^c}}{power^c}}{\frac{\frac{1}{R^p}}{power^p}} = \frac{\frac{power^c}{R^c}}{\frac{power^p}{R^p}} = \frac{R^p \cdot power^c}{R^c \cdot power^p} \quad (3)$$

- The EDP metric considers energy consumption. Nevertheless, in the same manner, as ITEEsv, the EDP considers the state of the energy consumption and the performance of a system in a specific time and conditions. For this reason, the EDP is not suitable for comparing different scenarios for server consolidation. If we would like to compare different scenarios of server consolidation, the use of the CIS^2 provides the relation with EDP (see Eq. 4).

$$CiS^2 = S_p \cdot S_e = \frac{R^c}{R^p} \cdot \frac{E^c}{E^p} = \frac{R^c}{R^p} \cdot \frac{R^c \cdot power^c}{R^c \cdot power^c} = \frac{EDP^c}{EDP^p} \quad (4)$$

3 THEORETICAL VIEW OF THE METRICS

To evaluate the EDP, the CiS^2, and the ITEEsv metrics in a theoretical manner, we proceed to take into account the index and metrics requirements definitions from (Potts, 2012), and then, apply them to our set of metrics (Juiz and Bermejo, 2020).

These requirements are: quantifiability, sensitivity, linearity, reliability, efficiency, and improvement oriented. It is important to highlight that in this paper, a change in the system's state is a variation in the number of consolidated virtual machines/containers.

After performing this analysis, we can state that all the metrics meet the requirements for being a suitable metric. Nevertheless, from the semantic point of view, the ITEEsv and EDP metrics do not reflect the reality of performance and energy trade-offs in server consolidation. Consequently, the CiS^2 index is the most suitable metric for this aim.

4 CURRENT METHODOLOGY: EVALUATION OF ITEESV METRIC

In this section, we attempt to demonstrate if the ITEEsv is suitable for server consolidation from the empirical point of view.

4.1 Experimental Setup

The experiment methodology is based on the comparison of CPU-intensive workloads (Juiz et al., 2023), in this case, the Sysbench benchmark (Casalicchio, 2019). Besides, the workload is distributed in a balanced manner over N physical machines against the same workload distributed over N virtual machines or containers consolidated in a single physical machine (see Figure 1). The comparison is done in terms of the ITEEsv metric.

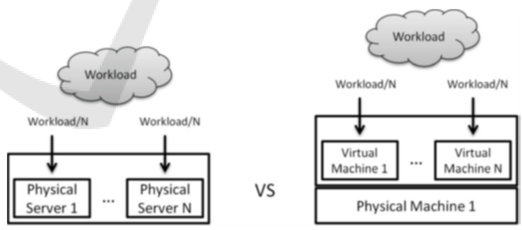

Figure 1: Workload execution comparison.

In this work, they take the system under test (SUT) as a black box, applying benchmarking and monitoring as performance engineering techniques ((Jain, 1991), (Molero et al., 2004)). The SUT executes the workload, and its performance and power consumption are monitored during the execution. We selected the response time as a performance metric. Also, the experiment setup is as follows:

- Server: Dell Power Edge T430. Number of CPUs 16, 8GB RAM size, and Ubuntu Server 16.04 as OS.

Table 1: ITEEsv value for physical machines.

N	R (s)	1/R (max perf)	Power (KW)	Energy (W·s)	ITEEsv
1	24,2532	0,04123	0,0937	2272,6218	0,440020
2	9,2707	0,1078	0,1871	1735,4101	0,576233
3	5,2972	0,1887	0,2804	1485,7390	0,673066
4	3,5626	0,2806	0,3712	1322,5724	0,756102
5	2,6253	0,3809	0,4635	1216,8265	0,821810
6	2,0386	0,4905	0,5560	1133,6161	0,882133
7	1,5668	0,6382	0,6509	1019,9209	0,980468
8	1,3093	0,7637	0,7441	974,3433	1,026332
9	1,1296	0,8852	0,8374	945,9438	1,057145

- Power meter: we measured the power consumption of the SUT with a Chroma 66200 device.
- Virtual Machine Monitor: Kernel-based Virtualization (KVM) (Type-I), Virtual Box (Type-II), and Docker (container-based).
- Virtual Machines and containers: Ubuntu 16.04, 1GB of RAM, and the same number of physical CPUs, in any case.

4.2 ITEEsv for Physical Machines

In Table 1 we can observe the behaviour of N parallel physical machines executing a workload in a distributed manner. When the number of distributed physical machines, the mean response time decreases. Then, the performance is increasing. Besides, when the number of distributed physical machines increases, the power consumption increases in the same manner.

Since the performance is measured in time units instead of throughput, the ITEEsv metric was calculated considering the maximum performance as the inverse of the mean response time.

Then, when the number of physical machines increases, the performance also increases, as a result, the ITEEsv value grows among the number of physical machines.

4.3 ITEEsv for Consolidated Virtual Servers

4.3.1 Type-I Hypervisor

In Table 2, we depict the behaviour of the mean response time of N virtual machines consolidated in one physical machine. In this case, the consolidation is done by a Type-I hypervisor, achieving less overhead than other solutions (Bermejo and Juiz, 2022). In this case, the workload is distributed among the virtual machines, with the portion of the load to be executed being smaller as the number of consolidated machines increases.

As the number of consolidated machines increases, performance increases. On the other hand, the power consumed by the physical machine remains stable, since there is always a single physical machine to consolidate the virtual machines.

Table 2: ITTEsv for Type-1 consolidation.

N	R (s)	1/R (s^{-1})	Power (KW)	Energy (W·s)	ITEEsv
1	25,42600	0,0393	0,1113	2831,4139	0,353180
2	19,2885	0,0518	0,1118	2156,6086	0,463691
3	15,8380	0,0631	0,1126	1784,7842	0,560292
4	13,9777	0,0715	0,1111	1554,0182	0,643493
5	12,5415	0,0797	0,1112	1395,0098	0,716841
6	5,1560	0,1939	0,1089	561,8493	1,779837
7	5,0785	0,1969	0,1048	532,6507	1,877403
8	2,8964	0,3452	0,1044	302,4701	3,3061
9	2,2928	0,4361	0,0977	224,0304	4,4636

In the ITEEsv we can see how its value increases as the degree of parallelism increases. However, its behaviour is different from the ITEEsv calculated for the physical machine (without consolidation). In this case, practically only the value of performance varies, with the value of the power consumed being constant. In this way, the value of the ITEEsv increases (to a greater extent than in the case of the physical machine) as the degree of consolidation increases.

At this point, it is important to answer the following question: could we use the ITEEsv metric to compare different consolidation configurations of virtual machines?

Considering the ITEEsv metric definition, it shows the relationship between the maximum performance and the maximum peak power consumed. If the aim is to compare the physical server with the consolidated one, it is necessary to divide the ITEEsv(type-I) by the ITEEsv(physical server).

In Figure 2 we can observe the relationship between the ITEEsv of physical and consolidated servers, and each one separately. For each number of physical machines and consolidated virtual machines, the ITEEsv value is depicted. Regarding the ITEEsv relationship, we only see how the consolidated server behaves concerning the physical server, for a given configuration.

When $N > 6$, the ITEEsv of the consolidated server is greater than the ITEEsv of the physical server, that is, the ITEEsv of the consolidated server is greater than that of the physical server. This means that, in these cases, the consolidated server consumes less electrical power than the physical server.

In addition, when $N \leq 6$ the ITEEsv of the physical server is lesser than the ITEEsv of the consolidated server. This is because the performance of the physical server is much better than the performance of the consolidated server, for a specific case.

Regarding the ratio between the ITEEsv, from Table 3 we can extract what relationship there is between a degree of parallelism represented in physical or virtual machines. However, the ITEEsv does not provide information on whether it is more suitable to consolidate the workload on virtual machines or not. This is because the ITEEsv does not consider the energy, and only takes into account the maximum peak power and performance but does not take into account the entire workload execution time.

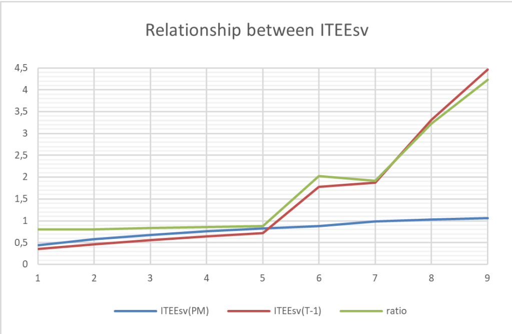

Figure 2: Relationship between ITEEsv of physical and virtual machines.

Table 3: ITEEsv ratio.

N	ITEEsv(PM)	ITEEsv(T-1)	Ratio
1	0,4400	0,3531	0,8026
2	0,5762	0,4636	0,8046
3	0,6730	0,5602	0,8324
4	0,7561	0,6434	0,8510
5	0,8218	0,7168	0,8722
6	0,8821	1,7798	2,0176
7	0,9804	1,8774	1,9148
8	1,0263	3,3061	3,2212
9	1,0571	4,4636	4,2223

4.3.2 Type-II Hypervisor

Table 4 shows the behaviour of a physical machine with N consolidated virtual machines using a Type-II hypervisor. Also, the workload is distributed and executed proportionally. In this way, the workload is distributed among all the consolidated machines, with the portion of the load to be executed being smaller as the number of consolidated machines increases.

As the number of consolidated machines increases, performance also grows. Nevertheless, the power consumed by the physical machines remains stable, due to there is always a single physical machine allocating all the consolidated virtual machines.

4.3.3 Type-I and Type-II Hypervisor Comparison

Once we obtain the values of the ITEEsv for Type-I and Type-II hypervisors, it is important to determine

Table 4: ITTEsv for Type-II consolidation.

N	R (s)	1/R (s-1)	Power (KW)	ITEEsv
1	26,2150	0,0381	0,0934	0,4082
2	10,0690	0,0993	0,0938	1,0581
3	5,8810	0,1700	0,0937	1,8134
4	3,9040	0,2561	0,0930	2,7535
5	2,8750	0,3478	0,0929	3,7407
6	2,3030	0,4342	0,0930	4,6647
7	1,7330	0,5770	0,0929	6,2108
8	1,4780	0,6765	0,0935	7,2329
9	1,2840	0,7788	0,0930	8,3661

Table 5: Comparison between ITEEsv in Type-I and Type-II hypervisors.

	Type-I				Type-II			
N	R (s)	1/R (s^{-1})	Power(KW)	ITEEsv	R (s)	1/R (s^{-1})	Power (KW)	ITEEsv
1	25,4260	0,0393	0,1113	0,3531	26,2150	0,0381	0,0934	0,4082
2	19,2885	0,0518	0,1118	0,4636	10,0690	0,0993	0,0938	1,0581
3	15,8380	0,0631	0,1126	0,5602	5,8810	0,1700	0,0937	1,8134
4	13,9777	0,0715	0,1111	0,6434	3,9040	0,2561	0,0930	2,7535
5	12,5415	0,0797	0,1112	0,7168	2,8750	0,3478	0,0929	3,7407
6	5,1560	0,1939	0,1089	1,7798	2,3030	0,4342	0,0930	4,6647
7	5,0785	0,1969	0,1048	1,8774	1,7330	0,5770	0,0929	6,2108
8	2,8964	0,3452	0,1044	3,3061	1,4780	0,6765	0,0935	7,2329
9	2,2928	0,4361	0,0977	4,4636	1,2840	0,7788	0,0930	8,3661

if the ITEEsv metric could be used to compare both hypervisors. That is, to answer the following question: Could we determine which hypervisor is more suitable to consolidate using the ITEEsv metric?

Table 5 shows the data for Type-1 and Type-II hypervisors: response time, maximum performance (1/R), power consumption, and the ITEEsv, calculated from the maximum performance and power consumption.

Regarding the response time, we can observe that the Type-II hypervisor is lower than that of Type-1. Therefore, the performance of Type II is better than the performance of Type I. From the power consumption point of view, the server with a Type-I hypervisor consumes more than the server with a Type-II hypervisor. Then, under these conditions, the ITEEsv of the server with Type-II hypervisor is higher than the ITEEsv of the server with Type-I.

If we want to compare the consolidated server with Type-I and with Type-II, we could only use the ITEEsv metric to compare with the given number of consolidated virtual machines. As in the previous case, this comparison would be made without considering energy consumption, disregarding the temporal behavior of power consumption. In this way, we could not determine the suitability of one hypervisor or another to consolidate virtual machines.

As a conclusion, the ITEEsv metric is not suitable for the comparison between Type-I and Type-II consolidation.

Table 6: ITEEsv for container consolidation.

N	R	1/R	Power (KW)	ITEEsv
1	247,6244	0,0040	0,0965	0,0418
2	95,0726	0,0105	0,0963	0,1091
3	53,7403	0,0186	0,0965	0,1927
4	32,6270	0,0306	0,0964	0,3177
5	23,9344	0,0417	0,0970	0,4306
6	18,5042	0,0540	0,0980	0,5512
7	10,6251	0,0941	0,0984	0,9561
8	8,7577	0,1141	0,0980	1,1643
9	6,5180	0,1534	0,0970	1,5812

4.3.4 Container-Based Hypervisor Consolidation

In previous sections, we considered the Type-I and Type-II hypervisors for server consolidation. However, since the containers are used as a lightweight alternative to consolidate servers, it is important to consider them in this paper.

In Table 6, the performance, the power consumption, and the ITEEsv metric are shown in container consolidation. As the number of containers increases, the mean response time decreases. As a consequence, the performance increases due to the workload being divided among the number of containers increases. In any case, the performance of containers is worse than Type-I and Type-II consolidation due to the overhead (Bermejo and Juiz, 2022).

Regarding the power consumption, we can observe that its value remains constant among the different numbers of containers. Moreover, considering the maximum performance and the maximum power consumption, the ITEEsv metric is also shown in Table 5.

To compare the ITEEsv from the three different ways to consolidate we can observe the results depicted in Figure 3. Since the containers have the worst performance, the ITEEsv values are smaller than the Type-I and Type-II. In addition, Type-I's performance is worse than the Type-II's performance. Then, the ITEEsv of Type-II is higher than the Type-I ITEEsv value, for any number of consolidated virtual machines or containers.

In the same manner, if we want to compare the consolidated server with Type-I, Type-II, and containers, we could only use the ITEEsv metric to compare with the given number of consolidated virtual machines. As in the use of virtual machines, this comparison would be made without considering energy consumption, disregarding the temporal behaviour of power consumption. In this way, we could not determine the suitability of one hypervisor or another to consolidate virtual machines or containers. Therefore, the ITEEsv metric is not suitable for the comparison between Type-I, Type-II, and container consolidation.

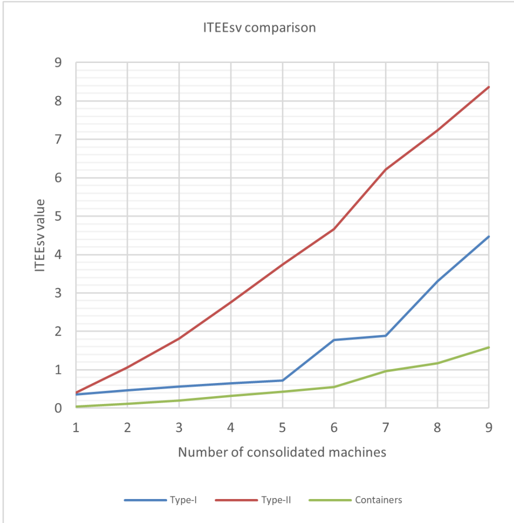

Figure 3: ITEEsv of Type-I, Type-II and containers.

5 PROPOSED METHODOLOGY

In this section, we proposed a generic methodology to measure the energy efficiency in server consolidation based on the CiS^2 metric (Juiz and Bermejo, 2020). As (Juiz and Bermejo, 2020) states, the CiS^2 metric quantifies the performance-energy trade-off of server consolidation, helping system administrators to decide about the servers' efficiency through benchmarking.

The general methodology is composed of the following phases. It is important to highlight that the workload execution in physical and virtual machines (or containers), was done following the architecture depicted in Figure 1.

- Phase 1: Set up the physical machines. In this stage, the physical machine should be set up with the corresponding workload.
- Phase 2: execution in physical machines with monitoring. After the set-up, the workload division starts with its execution. In parallel, the monitoring system starts to recover data from the performance and the power consumption. All of these data should be stored to be used in the last stage of this methodology.
- Phase 3: Set up the virtual machines or containers in the same physical machine. At this point, we will use the same physical machine as the previous one. However, deploying a set of virtual

machines or containers is necessary to execute the workload.

- Phase 4: execution in virtual machines or containers with monitoring. After the set-up, the workload division starts with its execution in the virtual machines or containers. In parallel, the monitoring system starts to recover data from the performance and the power consumption. It is important to highlight that the performance is measured from the virtual machines or containers, but the power consumption is monitored from the physical machine. All of these data should be stored to be used in the last stage of this methodology.
- Phase 5: CiS^2 calculation and compassion. After the workload execution in the physical machines and the consolidated one, the CiS^2 index is calculated as Eq. 5 shows.

$$CiS^2 = \frac{S_p}{S_e} = \frac{R^c}{R^p} \cdot \frac{E^c}{E^p} \tag{5}$$

In this stage, the CiS^2 can be calculated using the EDP metric as Eq. 6 depicts.

$$CiS^2 = \frac{EDP^c}{EDP^p} \tag{6}$$

For any number of consolidated machines, a CiS^2 index is obtained. Moreover, the CiS^2 index can be represented graphically as Figure 4a shows. Also, it is important to note that the CiS^2 index has a reference diagonal aiming to represent the ideal case for consolidation: the linearity of performance and energy (Juiz and Bermejo, 2020). If the value CiS^2 index is on the diagonal, it means that this consolidation configuration is the ideal one.

Nevertheless, the values in the green area indicate that the consolidation is efficient. On the contrary, the CiS^2 values in the red area are not efficient for server consolidation. Also, the position in the area influences the efficiency or the inefficiency of the CiS^2 value. For example, in Figure 4b, there are four represented points. Considering the reference diagonal, the point 2 is more efficient than the point 1. However, the point 4 is more inefficient than the point 3.

5.1 Application of the Methodology to Virtual Machines

After proposing the methodology, it is important to evaluate it using the same physical servers, virtual machines, and containers as the previous ITEEsv evaluation.

For the case of virtual machines, in Figure 5 the CiS^2 values are depicted for the Type-I and Type-II

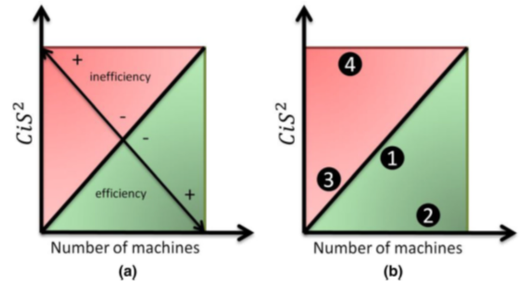

Figure 4: CiS^2 index graphical representation.

hypervisor. We can observe that for every number of consolidated virtual machines, there is a corresponding value of the CiS^2 index. For example, 6 consolidated virtual machines are more efficient in terms of performance and energy efficiency using Type-I hypervisor.

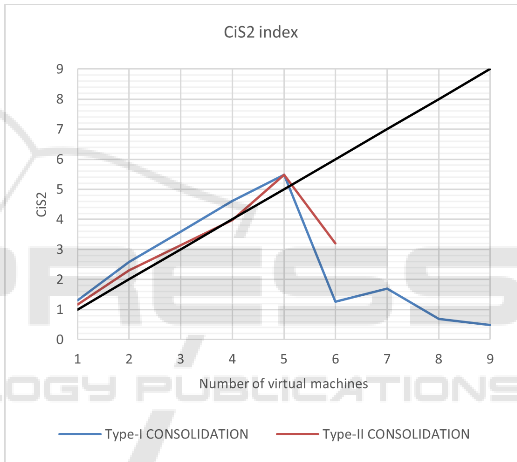

Figure 5: CiS^2 index for virtual machines.

5.2 Application of the Methodology to Containers

Moreover, it is important to consider server consolidation using containers. In Figure 6 we can observe the value of the CiS^2 index for any number of consolidated containers in a physical machine. In this case, we can observe that the container consolidation is not efficient, since all the CiS^2 values are in the inefficient area.

6 DISCUSSION

Once the evaluation of the ITEEsv metric has been carried out and an alternative methodology has been proposed using the existing CiS^2 metric, we can observe a series of phenomena to take into account.

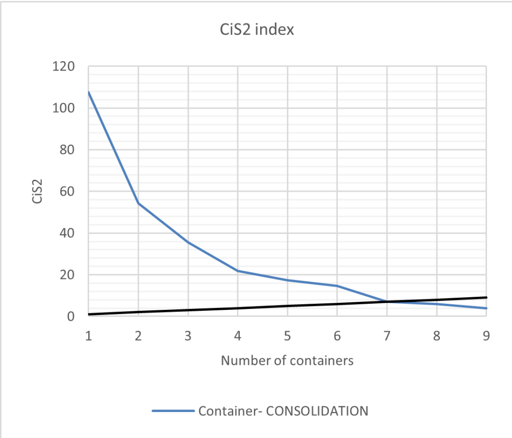

Figure 6: CiS^2 index for containers.

The first is related to the advantages provided by the proposed methodology. This is not restricted by the use of any load (such as those of SPEC) but can be applied to any work scenario. Even the type of virtualization technology used is irrelevant. This is because in this methodology the server is considered as a black box.

The second fact is related to the application of the methodology. Deploying virtual machines and containers has a higher cost than just working with physical machines. However, to deploy the scenario required to apply this methodology, it can be easily automated with current tools.

The last refers to the distribution of the load that uses the proposed methodology. As has been seen, it is necessary to use a load that can be distributed among the N physical and virtual servers. However, dividing the load is not a trivial task, and in this case, it would be necessary to correctly select said load.

To conclude, we could say that although the deployment of the proposed methodology requires some investment, the truth is that by taking into account the consolidation of virtual machines and containers, a current need would be satisfied.

7 CONCLUSIONS AND FUTURE WORK

Throughout this work, the current standards (ISO/IEC 30134-4 and ISO/IEC 21836) have been evaluated to determine their suitability for the consolidation of servers. It has been proven through real experimentation that both standards are not suitable for server consolidation.

As a result, a generic methodology based on the calculation of the CiS^2 metric has been proposed to cover this knowledge gap. In this way, the generic methodology could be applied to any server consolidation scenario.

As a line of future work, the formal standardization of the proposed methodology stands out.

ACKNOWLEDGEMENTS

This research work is part of the TED2021-132695B-I00 project, financed by MCIN/AEI/10.13039/501100011033 and by the European Union "NextGenerationEU"/PRTR.

This research is part of the project PID2021-122208OB-I00, PROYEXCEL_00286 funded by MCIN / AEI / 10.13039 / 501100011033 and by Andalusian Regional Government.

REFERENCES

Bermejo, B. and Juiz, C. (2022). A general method for evaluating the overhead when consolidating servers: performance degradation in virtual machines and containers. *The Journal of Supercomputing*, 78(9):11345–11372.

Casalicchio, E. (2019). A study on performance measures for auto-scaling cpu-intensive containerized applications. *Cluster Computing*, 22(3):995–1006.

Gonzalez, R. and Horowitz, M. (1996). Energy dissipation in general purpose microprocessors. *IEEE Journal of solid-state circuits*, 31(9):1277–1284.

ISO/IEC (2017). Iso/iec 30134-4:2017 information technology — data centres — key performance indicators — part 4: It equipment energy efficiency for servers (iteesv).

ISO/IEC (2020). Iso/iec 21836:2020 information technology – data centres – server energy effectiveness metric.

Jain, R. (1991). *The art of computer systems performance analysis: techniques for experimental design, measurement, simulation, and modeling*, volume 1. Wiley New York.

Juiz, C. and Bermejo, B. (2020). The c i s 2: a new metric for performance and energy trade-off in consolidated servers. *Cluster Computing*, 23(4):2769–2788.

Juiz, C. and Bermejo, B. (2024). On the scalability of the speedup considering the overhead of consolidating virtual machines in servers for data centers. *The Journal of Supercomputing*, pages 1–49.

Juiz, C., Capo, B., Bermejo, B., Fernández-Montes, A., and Fernández-Cerero, D. (2023). A case study of transactional workload running in virtual machines: The performance evaluation of a flight seats availability service. *IEEE Access*.

Molero, X., Juiz, C., and Rodeño, M. (2004). *Evaluación y modelado del rendimiento de los sistemas informáticos*. Pearson Educación London.

Potts, P. J. (2012). Glossary of analytical and metrological terms from the international vocabulary of metrology (2008). *Geostandards and Geoanalytical Research*, 36(3):231–246.