

Productive Efficiency of Energy-Aware Data Centers

Authors:

Damián Fernández-Cerero, Alejandro Fernández-Montes, Francisco Velasco

Date Submitted: 2018-09-21

Keywords: return-to-scale, cloud computing, efficiency, energy policies, data envelopment analysis

Abstract:

Information technologies must be made aware of the sustainability of cost reduction. Data centers may reach energy consumption levels comparable to many industrial facilities and small-sized towns. Therefore, innovative and transparent energy policies should be applied to improve energy consumption and deliver the best performance. This paper compares, analyzes and evaluates various energy efficiency policies, which shut down underutilized machines, on an extensive set of data-center environments. Data envelopment analysis (DEA) is then conducted for the detection of the best energy efficiency policy and data-center characterization for each case. This analysis evaluates energy consumption and performance indicators for natural DEA and constant returns to scale (CRS). We identify the best energy policies and scheduling strategies for high and low data-center demands and for medium-sized and large data-centers; moreover, this work enables data-center managers to detect inefficiencies and to implement further corrective actions.

Record Type: Published Article

Submitted To: LAPSE (Living Archive for Process Systems Engineering)

Citation (overall record, always the latest version):

LAPSE:2018.0575

Citation (this specific file, latest version):

LAPSE:2018.0575-1

Citation (this specific file, this version):

LAPSE:2018.0575-1v1

DOI of Published Version: <https://doi.org/10.3390/en11082053>

License: Creative Commons Attribution 4.0 International (CC BY 4.0)

Article

Productive Efficiency of Energy-Aware Data Centers

Damián Fernández-Cerero ^{1,*}, Alejandro Fernández-Montes ¹ and Francisco Velasco ²

¹ Department of Computer Languages and Systems, University of Seville, 41012 Sevilla, Spain; afdez@us.es

² Department of Applied Economy I, University of Seville, 41018 Sevilla, Spain; velasco@us.es

* Correspondence: damiancerero@us.es; Tel.: +34-954-559-531

Received: 17 July 2018; Accepted: 1 August 2018; Published: 8 August 2018

Abstract: Information technologies must be made aware of the sustainability of cost reduction. Data centers may reach energy consumption levels comparable to many industrial facilities and small-sized towns. Therefore, innovative and transparent energy policies should be applied to improve energy consumption and deliver the best performance. This paper compares, analyzes and evaluates various energy efficiency policies, which shut down underutilized machines, on an extensive set of data-center environments. Data envelopment analysis (DEA) is then conducted for the detection of the best energy efficiency policy and data-center characterization for each case. This analysis evaluates energy consumption and performance indicators for natural DEA and constant returns to scale (CRS). We identify the best energy policies and scheduling strategies for high and low data-center demands and for medium-sized and large data-centers; moreover, this work enables data-center managers to detect inefficiencies and to implement further corrective actions.

Keywords: data envelopment analysis; return-to-scale; cloud computing; efficiency; energy policies

1. Introduction

Data centers, which constitute the computational muscle for cloud computing, can be compared in energy consumption to many industrial facilities and towns. The latest trends show that these infrastructures represent approximately 2% of global energy consumption [1], with a 5% annual growth rate [2].

The data envelopment analysis mathematical model enables the management organizational divisions to measure the performance of an organization by providing the relative efficiency of each organizational unit. This relative efficiency measurement can be applied to a set of decision-making units, also known as DMUs, or for productive efficiency. The productive efficiency, also called technical efficiency, involves a collection of inputs (the resources needed for the production) and outputs (the production achieved). To this end, DEA constructs an “efficiency frontier” which places the relative performance of all units so these can be contrasted. This method is notably well-suited for the examination of the behavior of complex relations, even unknown, between numerous inputs and outputs, where the decisions made are affected by a level of uncertainty [3]. Moreover, DEA has been used both in private [4,5] and in public contexts [6–9].

Many initiatives have emerged looking for the decrease of the consumption of energy and the CO₂ trace of data-centers, especially those of a medium and large size. These facilities are composed of thousands and even tens of thousands of machines.

A substantial part of these initiatives focuses on the improvement of the Power Usage Effectiveness (PUE), that is the amount of energy consumed in non-computational tasks, such as power supply, cooling and networking components. This accounts for more than half of the energy consumption of an Internet data-center (IDC).

Several strategies are proposed to significantly improve energy efficiency in large-scale clusters [10]: cooling and temperature management [11,12]; power proportionality for CPU and

memory hardware components [13,14]; fewer energy-hungry and non-mechanical hard disks [15]; and new proposals for energy distribution [16].

On the other hand, almost 50% of energy is consumed by computational servers to satisfy the incoming workload. The job arrival is not stable over time, but usually presents correlative low and high periods, such as those present in day/night and weekday/weekend workload patterns.

Such scenarios present a huge opportunity for the improvement of energy efficiency through proper scheduling and through the application of low-energy consumption modes to servers, since keeping servers in an idle state is extremely energy-inefficient. Many energy-aware schedulers, which aim to raise server usage, have been proposed in order to free up the maximum amount of machines so that they may put into hibernation [17–19]. In addition to these schedulers, several energy-conservation strategies may be applied in virtualized environments, such as the consolidation and migration of virtual machines [20,21].

Other strategies focus on the reduction of energy consumption in specific scenarios, such as those of distributed file systems [22,23].

The most aggressive approach involves the shut-down of underutilized servers in order to minimize energy consumption. Several shut-down policies have been proposed for grid computing environments in [24]. This strategy is yet to be widely implemented in working data-centers since a natural reticence to worsening QoS is usually present in data-center operators [25].

The innovation of the research presented in this paper involves the utilization of data envelopment analysis (DEA) as a mathematical technique to compare the efficiency regarding the consumption of energy and the performance of various workload scenarios, scheduling models and energy efficiency policies. This efficiency analysis enables data-center operators to make appropriate decisions about the number of machines, the scheduling solution and the shut-down strategy that must be applied so that data-centers run optimally. The final goal is the maximization of the productive efficiency, which is computed as the amount of energy consumed to serve a workload with a determined performance.

The major contributions of this paper can be summarized as follows:

1. Extensive empirical experimentation and analysis of various cloud-computing scenarios with a trustworthy and detailed simulation tool.
2. Impact analysis in terms of the energy consumption and performance of several energy efficiency policies, which shut-down idle machines by means of data envelopment analysis.
3. DEA-conducted analysis of the performance impact and energy consumption of a set of scheduling models for large-scale data-centers.
4. Empirical determination and proposal of corrective actions to achieve optimal efficiency.

The work is organized as follows. In Section 2, the authors introduce the current literature for the utilization of DEA presented for various areas, as well as the DEA model employed in this work. In Section 3, we briefly explain the set of energy efficiency policies that shut down idle servers. The scheduling models considered are explained in Section 4. In Section 5, the tool used for the simulation, the experimental environment, the energy model and DEA inputs/outputs are presented. Natural constant returns to scale (CRS) DEA results are described and analyzed in Section 6. Finally, we summarize this paper and present conclusions in Section 7.

2. Data Envelopment Analysis Model

Data envelopment Analysis (DEA) is a method that analyzes the connections between the outputs and inputs required in a production process in order to establish the efficiency frontiers [26]. This non-parametric technique was first described for the determination of the efficiency of DMUs by [27] and was formally defined by [28]. DEA has been proposed to measure the efficiency in various areas of operations research and management science [29–32]. Moreover, it has been applied to measure the environmental performance by other authors [33–40], who describe the gains of this method in the field of environmental management, which is a matter of undoubted relevance

for the valuation of the sustainable development ability and pathway [41]. A critical feature of DEA for environmental analysis is the inclusion of desirable and undesirable outputs along with its own production variables, which cannot be isolated in an environmental analysis model of these features [42]. In this way, ref. [36] have refined a non-radial and radial model of DEA for environmental measurements. This approach separates the outputs into desirable and undesirable and presents two concepts: natural and managerial disposability. In this work, we employ the DEA radial approach for environmental assessments proposed by [37]. It should be borne in mind that a main feature of this approach is the utilization of DEA-RAM (range-adjusted measure), first proposed by [43] to treat in a unified manner the analysis of managerial and natural disposability.

2.1. Natural Disposability

Natural disposability refers to a DMU that improves its efficiency by decreasing its inputs in order to decrease its undesirable outputs, as well as to increase the desirable outputs.

In Model (1), each j -th DMU $j = 1, \dots, n$, considers inputs $X_j = (x_{1j}, \dots, x_{mj})^T$ for the production of desirable outputs $G_j = (g_{1j}, \dots, g_{sj})^T$ and undesirable outputs $B_j = (b_{1j}, \dots, b_{hj})^T$. Furthermore, $d_i^x, i = 1, \dots, m, d_r^g, r = 1, \dots, s$ and $d_f^b, f = 1, \dots, h$ are all slack variables which are related to inputs, desirable and undesirable outputs, respectively. $\lambda = (\lambda_1, \dots, \lambda_n)^T$ are structural or intensity variables, which are unknown and are used for the connection of the input and output vectors by means of a convex combination. R is the range resolute through the lower and upper limits of inputs, desirable outputs and undesirable outputs, denoted by:

$$\begin{aligned} R_i^x &= (m + s + h)^{-1} (\max\{x_{ij}/j = 1, \dots, n\} - \min\{x_{ij}/j = 1, \dots, n\}) \\ R_r^g &= (m + s + h)^{-1} (\max\{g_{rj}/j = 1, \dots, n\} - \min\{g_{rj}/j = 1, \dots, n\}) \text{ and} \\ R_f^b &= (m + s + h)^{-1} (\max\{b_{fj}/j = 1, \dots, n\} - \min\{b_{fj}/j = 1, \dots, n\}) \end{aligned}$$

The natural efficiency of the k -th policy is computed by the following CRS and radial VRS model (see [37] for a better understanding):

$$\begin{aligned} \max \quad & \zeta + \epsilon \left(\sum_{i=1}^m R_i^x d_i^x + \sum_{r=1}^s R_r^g d_r^g + \sum_{f=1}^h R_f^b d_f^b \right) \\ \text{s.t.} \quad & \sum_{j=1}^n x_{ij} \lambda_j + d_i^x = x_{ik}, i = 1, \dots, m, \\ & \sum_{j=1}^n g_{rj} \lambda_j - d_r^g - \zeta g_{rk} = g_{rk}, r = 1, \dots, s, \\ & \sum_{j=1}^n b_{fj} \lambda_j + d_f^b + \zeta b_{fk} = b_{fk}, f = 1, \dots, h, \\ & d_i^x \geq 0, i = 1, \dots, m, \\ & d_r^g \geq 0, r = 1, \dots, s, \\ & d_f^b \geq 0, f = 1, \dots, h, \\ & \zeta \quad \text{Unrestricted} \end{aligned} \tag{1}$$

where the unrestricted parameter ζ denotes an unknown inefficiency rate expressing the gap between the efficiency frontier and an empirical group of undesirable and desirable outputs. The parameter ϵ takes the value of 0.0001 in this work to minimize the influence of slack variables. If the restriction $\sum_{j=1}^n \lambda_j = 1$ is added to Model (1), then the obtained model is a VRS (Model (1*)).

The first restriction in equation systems ((1), (1*)) explores the values of λ_j to create a composite unit, considering inputs such as: $\sum_{j=1}^n x_{ij} \lambda_j = -d_i^x + x_{ik}, i = 1, \dots, m$. The values of the inputs can be decreased when the positive slack variables d_i^x are present. This may unquestionably vary the given rates, which implies that the system presents some inefficiencies.

In the same way, the second restriction, $\sum_{j=1}^n g_{rj}\lambda_j = d_r^g + \zeta g_{rk} + g_{rk}, r = 1, \dots, s$, indicates that the desirable outputs can be maintained or increased by making an increase of the slack variable d_r^g and a radial expansion ζg_{rk} .

The third restriction, $\sum_{j=1}^n b_{fj}\lambda_j = -d_f^b - \zeta b_{fk} + b_{fk}, f = 1, \dots, h$, shows the decrease of the inputs, and then, we could reduce the undesirable outputs both in their slack variables and radially.

The objective function considers that two origins of inefficiency may be established. A k-policy can be considered efficient when the following two conditions are met: (a) $\zeta = 0$; (b) $d_i^x = 0, d_r^g = 0, d_f^b = 0$. In this case, the k-policy belongs to the efficiency frontier, since it fulfills the constraints present in equation systems ((1), (1*)), and consequently, the objective function takes a value of zero. Otherwise, the value of the objective function for non-efficient policies is greater than zero, due to possible displacements in the slack variables and radial movements.

The natural efficiency is then computed by:

$$\theta^* = 1 - \left[\zeta^* + \epsilon \left(\sum_{i=1}^m R_i^x d_i^{x*} + \sum_{r=1}^s R_r^g d_r^{g*} + \sum_{f=1}^h R_f^b d_f^{b*} \right) \right]$$

The value of this unified efficiency measure ranges between zero and one. If the k-policy is efficient, then the objective function of equation systems ((1), (1*)) is zero, and hence, the efficiency score equals $\theta^* = 1$. Slack variables resulting in the optimality of the models represented in equation systems ((1), (1*)) show the level of inefficiency.

2.2. Managerial Disposability

The managerial efficiency of the k-th policy is evaluated by the following CRS and VRS radial model [37]:

$$\begin{aligned} s.t. \quad & \sum_{j=1}^n x_{ij}\lambda_j - d_i^x = x_{ik}, i = 1, \dots, m, \\ & \sum_{j=1}^n g_{rj}\lambda_j - d_r^g - \zeta g_{rk} = g_{rk}, r = 1, \dots, s, \\ & \sum_{j=1}^n b_{fj}\lambda_j + d_f^b + \zeta b_{fk} = b_{fk}, f = 1, \dots, h, \\ & d_i^x \geq 0, i = 1, \dots, m, \\ & d_r^g \geq 0, r = 1, \dots, s, \\ & d_f^b \geq 0, f = 1, \dots, h, \\ & \zeta \quad \text{Unrestricted} \end{aligned} \tag{2}$$

Similarly, if the restriction $\sum_{j=1}^n \lambda_j = 1$ is added to Model (2), then the obtained model is a VRS (Model (2*)). In this model (2), increasing the inputs is allowed since new technologies that emit less CO₂ emissions to the atmosphere can be used.

By using the VRS models, we can obtain the returns to scale (RTS) and damage to scale (DTS) (see [37] for a better understanding). It is clear that for the natural efficiency, the returns to scale have to be increasing, and for managerial efficiency, the damages to scale have to be decreasing. Otherwise, the technical units are not working well and should correct the imbalances, using the information of the efficient units to which they have to be similar (peers).

3. Energy Policies for Data Centers at a Glance

The following set of energy efficiency policies for shutting down underutilized machines have been developed in this work as an evolution of those presented in [24], which have been adapted to the more complex reality of the cloud-computing paradigm:

- Never: prevents any shut-down process.
- Always: shuts down every server running in an idle state.
- Load: shuts down machines when data-center load pressure fails to reach a given threshold.
- Margin: assures that a determined number of machines are turned on and available before shutting down any machine.

- Random: shuts down machines randomly by means of a Bernoulli distribution with parameter 0.5.
- Exponential: shuts down machines when the probability of one incoming task negatively impacting on the data-center performance is lower than a given threshold. This probability is computed by means of the exponential distribution.
- Gamma: shuts down machines when the probability of incoming tasks oversubscribing to the available resources in a particular time period is lower than a given threshold; this probability is computed by means of the Gamma distribution.

4. Scheduling Models for Data Centers at a Glance

Cluster schedulers constitute a core part of cloud computing systems, since they are responsible for optimal task assignment to computing nodes. Several degrees of parallelism have been added to overcome the limitations present in central monolithic scheduling approaches when complex and heterogeneous systems with a high number of incoming jobs are considered. The following scheduling models are studied in this work:

- Monolithic: A centralized and single scheduler is responsible for scheduling all tasks in the workload in this model [44]. This scheduling approach may be the perfect choice when real-time responses are not required [45,46], since the omniscient algorithm performs high-quality task assignments by considering all restrictions and features of the data-center [47–50] at the cost of longer latency [46]. The scheduling process of a monolithic scheduler, such as that given by Google Borg [51], is illustrated in Figure 1.

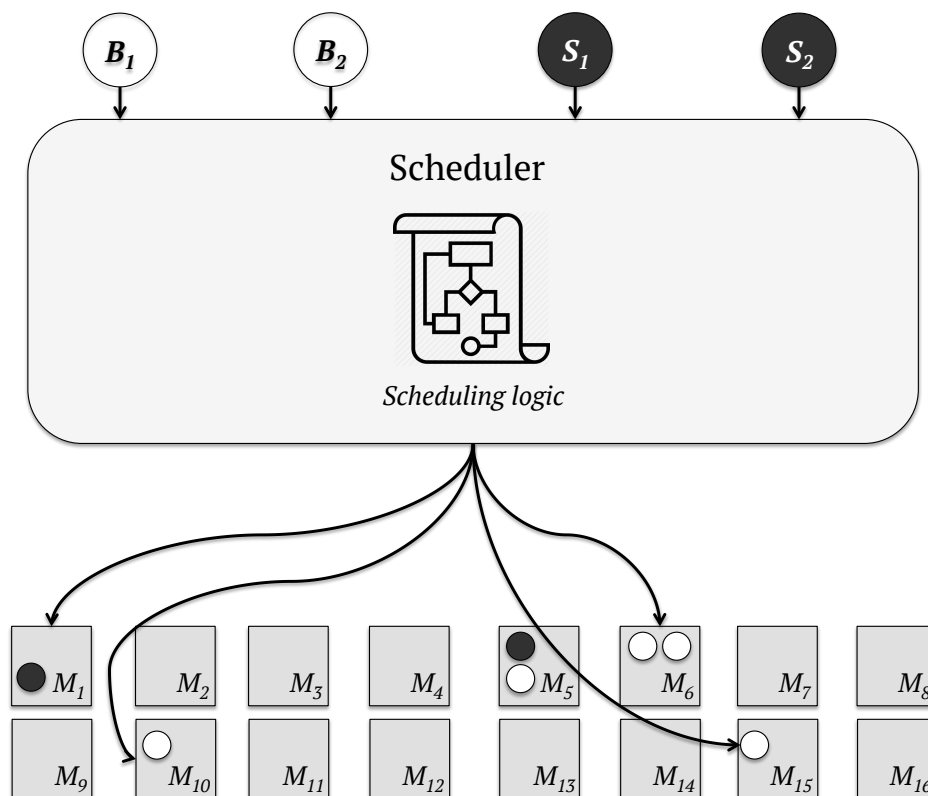


Figure 1. Monolithic scheduler architecture. M, worker node; S, service task; B, batch task [52].

- Two-level: This model achieves a higher level of parallelism by splitting the resource allocation and the task placement: a central manager blocks the whole cluster every time a scheduler makes a decision to offer computing resources to schedulers; and a set of parallel application-level schedulers performs the scheduling logic against the resources offered. This strategy enables the

development of sub-optimal scheduling logic for each application, since the state of the data-center is not shared with the central manager, nor with the application schedulers. The workflow of the Two-level schedulers [53,54] is represented in Figure 2.

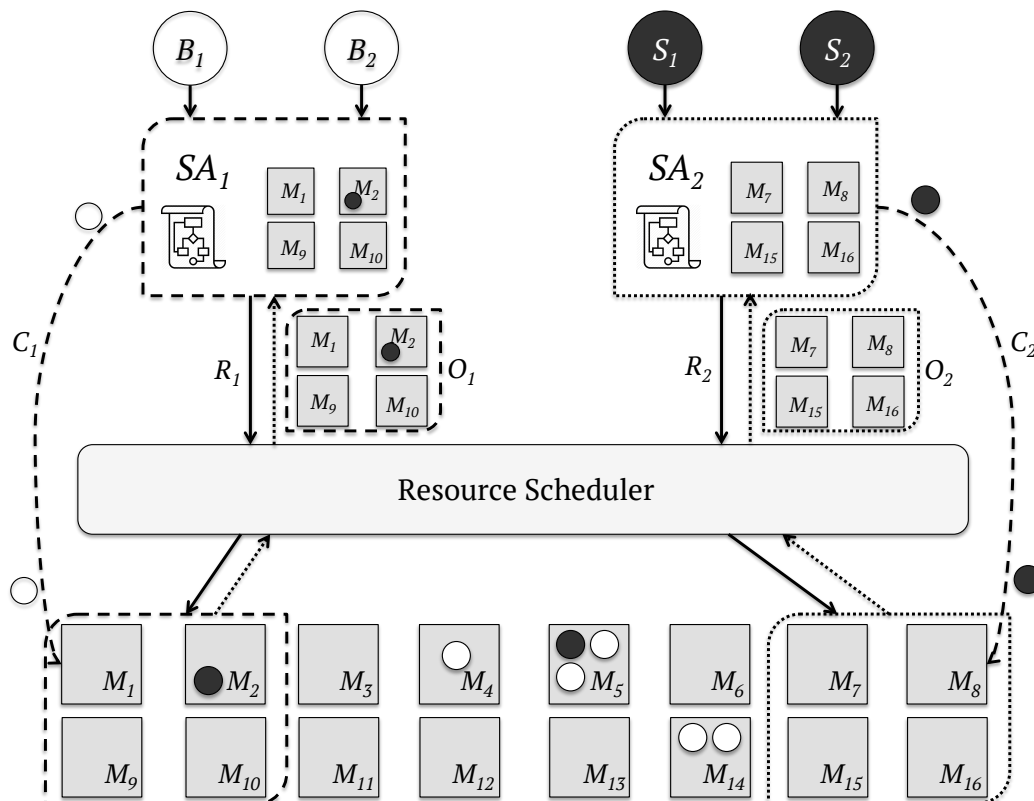


Figure 2. Two-level scheduler architecture. C, commit; O, resource offer; SA-, scheduler agent [52].

- Shared-state schedulers: On the other hand, in shared-state schedulers, such as Omega [55], the state of the data-center is available to all the schedulers. The central manager coordinates all the simultaneous parallel schedulers, which perform the scheduling logic against an out-of-date copy of the state of the data-center. The scheduling decisions are then committed to the central manager, which strives to apply these decisions. The utilization of stale views of the cluster by the schedulers can result in conflicts, since the chosen resources may not longer be available. In such a scenario, the local view of the state of the data-center stored in the scheduler is refreshed before the repetition of the scheduling process. The workflow of the shared-state scheduling model is represented in Figure 3.

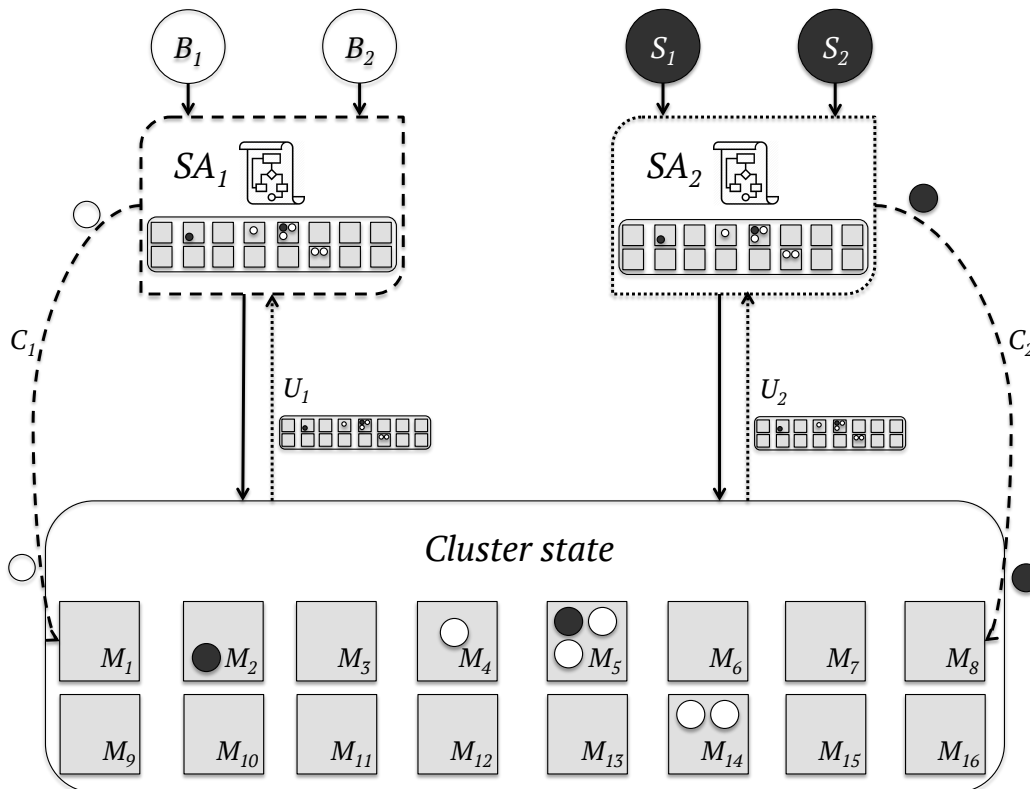


Figure 3. Shared-state scheduler architecture. U, cluster state update [52].

5. Methodology

In these next sections, the experimental environment designed for the implementation of the natural CRS DEA analysis is presented. The workflow followed in this work is shown in Figure 4.

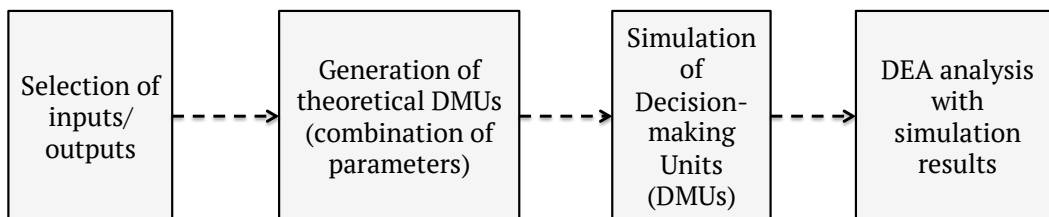


Figure 4. Methodology workflow employed in this work. DEA, data envelopment analysis.

5.1. Simulation Tool

The SCORE simulator [52] is employed in this work, since simulation is the best alternative in scenarios where the implementation of the proposed strategies on real large-scale data-centers remains unfeasible. This simulator provides us with the tools for the development and application of the energy policies described in Section 3 and the scheduling models presented in Section 4 on realistic large-scale cloud computing systems.

5.2. Environment and DMU Definition

Following the trends presented in [56,57], two utilization environments have been simulated in this paper for seven days of operation:

- the low-utilization scenario, which represents highly over-provisioned infrastructures and achieves an average utilization of approximately 30%.
- the high-utilization scenario, which represents facilities of a more efficient nature that use approximately 65% of available resources on average.

These scenarios are applied to three data-center sizes: (a) Small: composed of 1000 computing servers; (b) Medium: composed of 5000 computing servers; and (c) Large: composed of 10,000 computing servers. Each server is equipped with four CPU cores and 8 GB of RAM.

Decision-making units (DMUs) are defined by the following elements: (a) an energy efficiency policy; (b) a scheduling model; and (c) a workload scenario.

5.3. Energy Model

The following states are presented for each resource in the energy model applied in this work: (a) Idle: when the machine is not executing tasks; and (b) Busy: otherwise.

Let t_{idle}^i represent the time the i -th resource is idle, and let t_{busy}^i denote the time during which the machine is computing tasks. In the same way, P_{idle}^i and P_{busy}^i represent the power required for the machines to run in these states, respectively.

The time a machine spends on executing a job may be defined as follows:

$$t_{busy}^{ij} = \max_{t \in Tasks^i} C_t \quad (3)$$

where $Tasks^{ij}$ represents the tasks of the j -th job assigned to M_i and C_t denotes the completion time of the t -th task of the j -th job.

In the same way, the total time a machine is executing tasks and the total time it is in an idle state may be defined as follows:

$$t_{busy}^i = \sum_{j=1}^j t_{busy}^{ij} \quad (4)$$

$$t_{idle}^i = t_{total}^i - t_{busy}^i \quad (5)$$

where t_{total}^i represents the total operation time. Therefore, we can express the energy consumption as follows:

$$\sum_{i=1}^m (P_{busy}^i * t_{busy}^i + P_{idle}^i * t_{idle}^i) \quad (6)$$

The considered power states, transitions and values for the energetic model are shown in Figure 5.

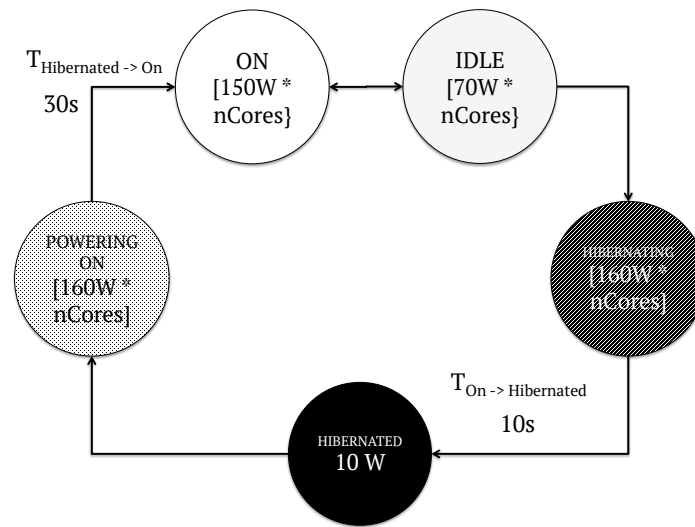


Figure 5. Machine power states [52].

5.4. DEA Inputs and Outputs

The inputs and outputs considered in DEA analysis and representative experimentation values are shown in Tables 1 and 2, respectively. One hundred and eight DMUs were analyzed, which were the result of the combination of all energy policies, scheduling models, data-center sizes and workload types described in Sections 3, 4 and 5.2, respectively. However, for clarity, a subset of the most interesting eighteen DMUs is shown in this paper. Each environment presents the following inputs and outputs:

- Inputs: Two inputs are considered in this work: (a) the number of machines in the data-center (D.C.), as shown in Section 5.2; and (b) the number of shut-down operations performed. These inputs may be reduced or kept equal.
- Outputs: One desirable output and two undesirable outputs are considered in this paper: (a) the time used to perform tasks' operations. The longer the time, the less idle the data-center. This good input can be maximized or kept equal; (b) the energy consumption of the data-center. The lower the energy consumption, the more efficient the data-center. This bad input may be reduced or kept equal; and (c) the average time jobs spend in a queue until they are scheduled. The shorter the time, the more performant the system is. This bad input may be reduced or kept equal.

Table 1. DEA inputs and outputs. Action column arrows mean whether the input/output value may be decreased (down arrow), increased (up arrow) or kept equal.

Parameter	Description	Action
Inputs		
Data-center size	Number of machines in the data-center	↓ ↔
#shut-downs	Number of shut-down operations	↓ ↔
Outputs		
Computation time	Total amount of useful task computation	↑ ↔
Energy consumption	Total data-center energy consumption	↓ ↔
Queue time	Average time until jobs are fully scheduled	↓ ↔

Table 2. Sample from the dataset for DEA analysis. The full dataset showing the results for the 108 DMUs analyzed can be found as the Supplementary Material. Energy policies, scheduling models, data-center sizes and workload types can be found in Sections 3, 4 and 5.2, respectively. D.C., data-center.

DMU		Inputs			Outputs		
Energy Policy	Scheduling Model	Work-Load	D.C. Size	#Shut-Downs	Computing Time (h)	MWh Consumed	Queue Time (ms)
Always	Monolithic	High	1000	37,166	104.42	49.01	90.10
Margin	Mesos	High	1000	13,361	104.26	49.65	1093.00
Gamma	Omega	High	1000	14,252	104.17	49.60	0.10
Always	Mono.	Low	1000	36,404	49.25	23.92	78.30
Exponential	Mesos	Low	1000	19,671	49.63	24.65	1188.70
Load	Omega	Low	1000	32,407	49.34	24.19	1.10
Margin	Mono.	High	5000	6981	99.96	237.09	126.20
Gamma	Mono.	High	5000	9877	99.96	235.92	129.80
Random	Mesos	High	5000	33,589	100.03	234.90	1122.60
Margin	Omega	High	5000	8578	100.26	239.13	0.70
Exponential	Omega	High	5000	11,863	100.26	236.95	1.00
Margin	Omega	Low	5000	15,452	46.70	115.82	0.50
Margin	Mono.	High	10,000	9680	101.56	481.36	325.20
Gamma	Mono.	High	10,000	11,388	101.56	479.36	327.90
Margin	Omega	High	10,000	18,150	101.63	486.11	2.60
Gamma	Omega	High	10,000	18,409	101.63	484.69	2.50
Gamma	Mesos	Low	10,000	29,707	45.83	228.31	1107.60
Random	Omega	Low	10,000	40,772	46.09	233.50	3.80

6. Natural CRS DEA Results

The whole dataset included as an Appendix is analyzed by means of natural CRS and VRS DEA. However, only the most relevant natural CRS DEA results for the most representative DMUs, which are presented in Table 2, are described in this section.

An efficiency analysis depending on the data-center size and on the energy policy is shown in Tables 3 and 4. The following conclusions can be drawn:

- The best efficiency levels are achieved for small data-centers. The data-center size input is predominant in this group of DMUs, since no major differences between energy policies, scheduling frameworks and workload scenarios are present ($\sigma = 0.01$, $\bar{x} = 0.99$).
- Mid-size data-centers should use the margin energy policy and monolithic or Omega schedulers and should avoid all other energy policies and the Mesos scheduler. Moreover, high workload scenarios are also more efficient than low workload scenarios. In addition, the following DMUs achieve a good level of efficiency, but they do not belong to the efficiency frontier: (a) the DMU combining the Gamma energy policy and the monolithic or Omega schedulers; (b) the DMU combining the exponential energy policy and the Omega scheduler.
- No DMU is efficient in large-scale data-centers. However, the following DMUs present good levels of efficiency: (a) the DMUs combining the Gamma, exponential or margin energy policy with the high workload scenario and the monolithic scheduler; and (b) the DMUs combining the Gamma or margin energy policy with the high workload scenario and the Omega scheduler.
- In high-loaded scenarios, the monolithic scheduler presents the lowest deviation regardless of the data-center size ($\sigma = 0.32$).

We can determine that it is always inefficient to operate in a low utilization scenario in medium-sized and large data-centers. Moreover, both the margin and the probabilistic energy policies (Gamma and exponential) perform more efficiently than the rest of the energy policies, as shown in Figure 6. The monolithic scheduler seems to achieve good results even for large-scale data-centers,

while the two-level scheduling approach has a negative impact on data-center performance. However, the trends show that the performance of the monolithic scheduling approach suffers from degradation on larger data-centers and higher workload pressure, and hence, lower efficiency levels are to be expected if larger sizes and higher utilization scenarios are to be considered.

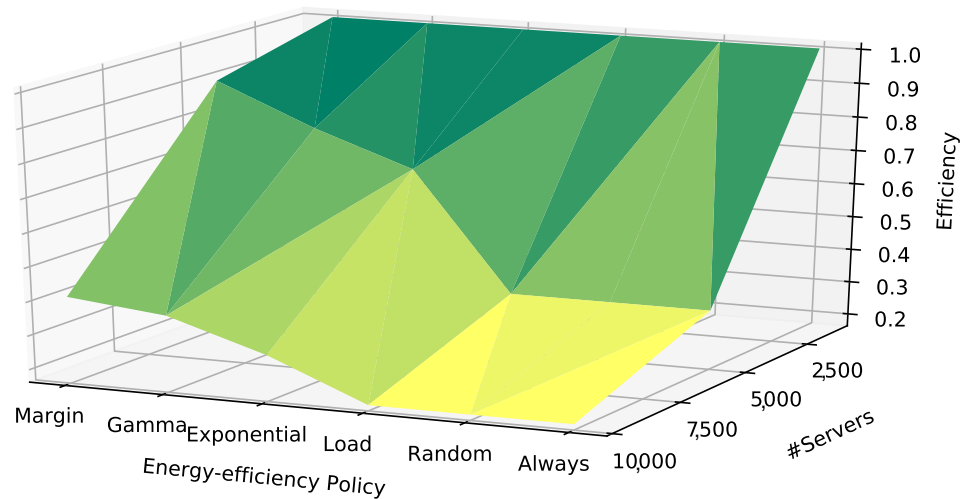


Figure 6. Summary of DEA natural constant returns to scale (CRS) efficiency results for energy efficiency policies.

The actions proposed for the improvement of efficiency of the most relevant DMUs are shown in Table 5.

Table 3. Efficiency analysis for data-center sizes.

Scheduling Model	Workload Scenario	Data-Center Size			Efficiency	
		1000	5000	10,000	σ	\bar{x}
Monolithic	High	1.00	0.60	0.37	0.32	0.66
Monolithic	Low	0.98	0.33	0.18	0.43	0.49
Mesos	High	1.00	0.47	0.18	0.41	0.55
Mesos	Low	0.97	0.32	0.17	0.43	0.49
Omega	High	1.00	0.62	0.27	0.36	0.63
Omega	Low	0.97	0.32	0.17	0.43	0.49
	σ	0.01	0.14	0.08		
	\bar{x}	0.99	0.44	0.23		
					0.40	0.55

Table 4. Efficiency analysis of energy policies.

Energy Policy	Scheduling Model									Efficiency	
	Monolithic			Mesos			Omega			σ	\bar{x}
	1000	5000	10,000	1000	5000	10,000	1000	5000	10,000		
Always	0.99	0.33	0.18	0.99	0.33	0.18	0.99	0.33	0.18	0.37	0.50
Random	0.99	0.33	0.18	0.98	0.32	0.18	0.98	0.33	0.18	0.37	0.50
Load	0.99	0.33	0.18	0.99	0.33	0.18	0.99	0.33	0.18	0.37	0.50
Margin	0.99	0.66	0.42	0.99	0.53	0.18	0.99	0.66	0.30	0.31	0.63
Exp.	0.99	0.54	0.31	0.99	0.40	0.18	0.99	0.58	0.22	0.33	0.58
Gamma	0.99	0.58	0.38	0.98	0.47	0.18	0.99	0.61	0.29	0.31	0.61

6.1. Proposed Corrections for a Sample DMU

DMU #104 is selected to illustrate how corrective actions are proposed by DEA in order to achieve efficiency. This DMU is defined by the combination of the random energy efficiency policy, the Omega scheduling model and a low utilization workload scenario.

DMU #104 presents a natural efficiency of 0.1697. This means it is far from being efficient. The following corrective actions are suggested for it to belong to the efficiency frontier, as shown in Table 6:

- The time the data-center spends on task computation must be increased by 38.28 h (+83%).
- Energy consumption must be reduced by 193.88 MWh (−83%).
- The average time jobs wait in a queue must be reduced by 3.23 s (−83%).
- The number of servers must be reduced by 9190 (−92%).
- Shut-down operations must be reduced by 9680 (−24%).

In addition to these corrective actions, the peers this DMU should emulate are #13, #34 and #18. This means that the workload must be increased, and better energy efficiency policies, such as margin and always, must be used. The full dataset containing all the DMUs and DEA analysis and corrections can be found as Supplementary Material in the Appendix.

Some of the proposed changes involve the switching of the scheduling framework, which is hardly achievable with the current resource manager systems. To implement these corrections, a resource managing system able to dynamically change the scheduling framework during runtime would be necessary. Such a system is an interesting improvement to the current state of the art that the DEA analysis leads us to develop.

Table 5. Resulting proposed corrections following DEA analysis. Peer projections for a DMU indicate which DMU it should emulate. The following actions may be taken for each input and output: ↑ when the parameter must be increased; ↓ if the parameter must be reduced; and ↔ if no further actions are needed to achieve efficiency.

#	DMU			Peer	Corrections				
	Energy Policy	Sched. Model	Work-load	Projec-tions	D.C. Size	#Shut-downs	Comp. Time	Energy Cons.	Queue Time
1	Always	Mono.	High	↔	↔	↔	↔	↔	↔
10	Margin	Mesos	High	4 (88%)	↓	↔	↑	↓	↓
18	Gamma	Omega	High	↔	↔	↔	↔	↔	↔
19	Always	Mono.	Low	↔	↓	↓	↑	↓	↓
29	Exp.	Mesos	Low	23 (56%) 22 (48%)	↓	↓	↑	↓	↓
33	Load	Omega	Low	31 (100%)	↓	↓	↑	↓	↔
40	Margin	Mono.	High	↔	↔	↔	↔	↔	↔
42	Gamma	Mono.	High	6 (59%) 41 (41%)	↓	↔	↑	↓	↓
44	Random	Mesos	High	7 (100%)	↓	↔	↑	↓	↓
52	Margin	Omega	High	↔	↔	↔	↔	↔	↔
53	Exp.	Omega	High	16 (63%) 52 (36%)	↓	↔	↑	↓	↓
70	Margin	Omega	Low	18 (72%)	↓	↓	↑	↓	↓
76	Margin	Mono.	High	6 (55%) 40 (45%)	↓	↔	↑	↓	↓
78	Gamma	Mono.	High	6 (90%)	↓	↔	↑	↓	↓
88	Margin	Omega	High	18 (95%)	↓	↔	↑	↓	↓
90	Gamma	Omega	High	18 (96%)	↓	↔	↑	↓	↓
102	Gamma	Mesos	Low	1 (49%) 22 (38%)	↓	↔	↑	↓	↓
104	Random	Omega	Low	13 (53%) 34 (36%)	↓	↓	↑	↓	↓

7. Conclusions and Policy Implications

In this work, we have confirmed the hypothesis that DEA constitutes a powerful tool for the analysis of technical efficiency in cloud-computing scenarios where large-scale data-centers provide the computational core.

Data envelopment analysis provides cloud-computing operators with the means for the identification of which data-center configuration better suits their requirements, both in terms of performance and energy efficiency.

This methodology allows us to analyze several energy efficiency policies that shut down idle servers, so that their behavior and differences can be compared in various data-center environments. It has been proven that policies based on a security margin and those that use statistical tools to predict the future workload, such as exponential and Gamma, deliver better results than policies based on data-center workload pressure and random strategies.

In addition, it has been empirically shown that even under medium and high workload pressure, in data-centers composed of up to 10,000 machines, monolithic schedulers perform better than other scheduling models, such as the two-level and shared-state approaches.

Table 6. Corrections proposed for DMU #104.

Results for DMU #104					
Natural Efficiency = 0.1697					
Projection Summary:					
	Variable	Original Value	Radial Movement	Slack Movement	Projected Value
Output	Computation (h)	46.09	+83%	0	84.37
Output	MWh consumed	233.50	−83%	0	39.62
Output	Queue time (ms)	3.80	−83%	0	0.6
Input	#Servers	10,000	0	−9190	810
Input	#Shut-downs	40,772	0	−9680	31,092
Listing of Peers:					
	Peer	Lambda Weight			
	#13	53%			
	#34	36%			
	#18	11%			

Finally, cloud-computing infrastructure managers are provided with empirical knowledge of which data-centers are not being used optimally, and hence, they can make decisions regarding the shut-down of machines in order to achieve higher utilization levels of the cloud-computing system as a whole.

As future work related to the limitations of the presented work, we may include:

- The addition of different kind of workload patterns, as well as real workload traces.
- The analysis of other scheduling models, such as distributed and hybrid models.
- The development of a new-generation resource-managing system that could dynamically apply the optimal scheduling framework depending on the environment and workload.
- The analysis of simulation data with other DEA approaches, such as Bayesian and probabilistic models, which could minimize the impact of the noise in current DEA models.

Supplementary Materials: Supplementary Materials are available online at <http://www.mdpi.com/1996-1073/11//2053/s1>.

Author Contributions: D.F.-C. and A.F.-M. conceived of and designed the experiments. D.F.-C. performed the experiments. D.F.-C., A.F.-M. and V.F. analyzed the data. D.F.-C. and F.V. contributed reagents/materials/analysis tools. D.F.-C., A.F.-M. and F.V. wrote the paper.

Funding: This research was funded by VI Plan Propio de Investigación y Transferencia—University of Seville 2018 grant number 2018/00000520.

Acknowledgments: The research is supported by the VI Plan Propio de Investigación y Transferencia (VIPPI), University of Seville.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Koomey, J. *Growth in Data Center Electricity Use 2005 to 2010*; Analytical Press: Piedmont, CA, USA, 1 August 2011.
2. Van Heddeghem, W.; Lambert, S.; Lannoo, B.; Colle, D.; Pickavet, M.; Demeester, P. Trends in worldwide ICT electricity consumption from 2007 to 2012. *Comput. Commun.* **2014**, *50*, 64–76. [[CrossRef](#)]
3. Gómez-López, M.T.; Gasca, R.M.; Pérez-Álvarez, J.M. Decision-Making Support for the Correctness of Input Data at Runtime in Business Processes. *Int. J. Cooper. Inf. Syst.* **2014**, *23*. [[CrossRef](#)]
4. Amirteimoori, A.; Emrouznejad, A. Optimal input/output reduction in production processes. *Decis. Support Syst.* **2012**, *52*, 742–747. [[CrossRef](#)]

5. Chiang, K.; Hwang, S.N. Efficiency measurement for network systems IT impact on firm performance. *Decis. Support Syst.* **2010**, *48*, 437–446.
6. Chang, Y.T.; Zhang, N.; Danao, D.; Zhang, N. Environmental efficiency analysis of transportation system in China: A non-radial DEA approach. *Energy Policy* **2013**, *58*, 277–283. [[CrossRef](#)]
7. Arcos-Vargas, A.; Núñez-Hernández, F.; Villa-Caro, G. A DEA analysis of electricity distribution in Spain: An industrial policy recommendation. *Energy Policy* **2017**, *102*, 583–592. [[CrossRef](#)]
8. Gonzalez-Rodriguez, M.; Velasco-Morente, F.; González-Abril, L. La eficiencia del sistema de protección social español en la reducción de la pobreza. *Papeles de Población* **2010**, *16*, 123–154.
9. Afonso, A.; Schuknecht, L.; Tanzi, V. Public sector efficiency: Evidence for new EU member states and emerging markets. *Appl. Econ.* **2010**, *42*, 2147–2164. [[CrossRef](#)]
10. Jakóbič, A.; Grzonka, D.; Kołodziej, J.; Chis, A.E.; González-Vélez, H. Energy Efficient Scheduling Methods for Computational Grids and Clouds. *J. Telecommun. Inf. Technol.* **2017**, *1*, 56–64.
11. Sharma, R.K.; Bash, C.E.; Patel, C.D.; Friedrich, R.J.; Chase, J.S. Balance of power: Dynamic thermal management for internet data centers. *IEEE Internet Comput.* **2005**, *9*, 42–49. [[CrossRef](#)]
12. El-Sayed, N.; Stefanovici, I.A.; Amvrosiadis, G.; Hwang, A.A.; Schroeder, B. Temperature management in data-centers: Why some (might) like it hot. In Proceedings of the 12th ACM SIGMETRICS/PERFORMANCE Joint International Conference on Measurement and Modeling of Computer Systems, London, UK, 11–15 June 2012; pp. 163–174.
13. Miyoshi, A.; Lefurgy, C.; Van Hensbergen, E.; Rajamony, R.; Rajkumar, R. Critical power slope: Understanding the runtime effects of frequency scaling. In Proceedings of the 16th International Conference on Supercomputing, New York, NY, USA, 22–26 June 2016; pp. 35–44.
14. Fan, X.; Weber, W.D.; Barroso, L.A. Power provisioning for a warehouse-sized computer. In Proceedings of the 34th Annual International Symposium on Computer Architecture, San Diego, CA, USA, 9–13 June 2007; pp. 13–23.
15. Andersen, D.G.; Swanson, S. Rethinking flash in the data-center. *IEEE Micro* **2010**, *30*, 52–54. [[CrossRef](#)]
16. Femal, M.E.; Freeh, V.W. Boosting data-center performance through non-uniform power allocation. In Proceedings of the Second International Conference on Autonomic Computing (ICAC'05), Seattle, WA, USA, 13–16 June 2005, doi:10.1109/ICAC.2005.17.
17. Jakóbič, A.; Grzonka, D.; Kołodziej, J. Security supportive energy aware scheduling and scaling for cloud environments. In Proceedings of the 31st European Conference on Modelling and Simulation (ECMS 2017), Budapest, Hungary, 23–26 May 2017; pp. 583–590.
18. Juarez, F.; Ejarque, J.; Badia, R.M. Dynamic energy-aware scheduling for parallel task-based application in cloud computing. *Future Gener. Comput. Syst.* **2018**, *78*, 257–271. [[CrossRef](#)]
19. Lee, Y.C.; Zomaya, A.Y. Energy efficient utilization of resources in cloud computing systems. *J. Supercomput.* **2012**, *60*, 268–280. [[CrossRef](#)]
20. Sohrabi, S.; Tang, A.; Moser, I.; Aleti, A. Adaptive virtual machine migration mechanism for energy efficiency. In Proceedings of the 5th International Workshop on Green and Sustainable Software, Austin, TX, USA, 14–22 May 2016; pp. 8–14.
21. Beloglazov, A.; Buyya, R. Optimal online deterministic algorithms and adaptive heuristics for energy and performance efficient dynamic consolidation of virtual machines in cloud data-centers. *Concurr. Comp.-Pract. E* **2012**, *24*, 1397–1420. [[CrossRef](#)]
22. Kaushik, R.T.; Bhandarkar, M. Greenhdfs: Towards an energy-conserving, storage-efficient, hybrid hadoop compute cluster. In Proceedings of the 2010 International Conference on Power Aware Computing and Systems, Vancouver, BC, Canada, 3–6 October 2010; pp. 1–9.
23. Luo, X.; Wang, Y.; Zhang, Z.; Wang, H. Superset: A non-uniform replica placement strategy towards high-performance and cost-effective distributed storage service. In Proceedings of the 2013 International Conference on Advanced Cloud and Big Data, Nanjing, China, 13–15 December 2013.
24. Fernández-Montes, A.; Gonzalez-Abril, L.; Ortega, J.A.; Lefèvre, L. Smart scheduling for saving energy in grid computing. *Expert Syst. Appl.* **2012**, *39*, 9443–9450. [[CrossRef](#)]
25. Fernández-Montes, A.; Fernández-Cerero, D.; González-Abril, L.; Álvarez-García, J.A.; Ortega, J.A. Energy wasting at internet data-centers due to fear. *Pattern Recogn. Lett.* **2015**, *67*, 59–65. [[CrossRef](#)]
26. Farrell, M.J. The measurement of productive efficiency. *J. R. Stat. Soc. Ser. A (Gen.)* **1957**, *120*, 253–290. [[CrossRef](#)]

27. Charnes, A.; Cooper, W.W.; Rhodes, E. Measuring the efficiency of decision making units. *Eur. J. Oper. Res.* **1978**, *2*, 429–444. [[CrossRef](#)]
28. Banker, R.D.; Charnes, A.; Cooper, W.W. Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis. *Manag. Sci.* **1984**, *30*, 1078–1092. [[CrossRef](#)]
29. Fernández-Montes, A.; Velasco, F.; Ortega, J. Evaluating decision-making performance in a grid-computing environment using DEA. *Expert Syst. Appl.* **2012**, *39*, 12061–12070. [[CrossRef](#)]
30. Campos, M.; Fernández-Montes, A.; Gavilan, J.; Velasco, F. Public resource usage in health systems: A data envelopment analysis of the efficiency of health systems of autonomous communities in Spain. *Public Health* **2016**, *138*, 33–40. [[CrossRef](#)] [[PubMed](#)]
31. Fernández-Serrano, J.; Berbegal, V.; Velasco, F.; Expósito, A. Efficient entrepreneurial culture: A cross-country analysis of developed countries. *Int. Entrep. Manag. J.* **2017**, *14*, 105–127. [[CrossRef](#)]
32. Exposito, A.; Velasco, F. Municipal solid-waste recycling market and the European 2020 Horizon Strategy: A regional efficiency analysis in Spain. *J. Clean. Prod.* **2018**, *172*, 938–948. [[CrossRef](#)]
33. Scheel, H. Undesirable outputs in efficiency valuations. *Eur. J Oper. Res.* **2001**, *132*, 400–410. [[CrossRef](#)]
34. Färe, R.; Grosskopf, S.; Hernandez-Sancho, F. Environmental performance: An index number approach. *Resour. Energy Econ.* **2004**, *26*, 343–352. [[CrossRef](#)]
35. Zhou, P.; Ang, B.W.; Poh, K.L. Measuring environmental performance under different environmental DEA technologies. *Energy Econ.* **2008**, *30*, 1–14. [[CrossRef](#)]
36. Sueyoshi, T.; Goto, M. Returns to scale and damages to scale on US fossil fuel power plants: Radial and non-radial approaches for DEA environmental assessment. *Energy Econ.* **2012**, *34*, 2240–2259. [[CrossRef](#)]
37. Sueyoshi, T.; Goto, M. DEA radial measurement for environmental assessment: A comparative study between Japanese chemical and pharmaceutical firms. *Appl. Energy* **2014**, *115*, 502–513. [[CrossRef](#)]
38. Halkos, G.E.; Tzeremes, N.G. Measuring the effect of Kyoto protocol agreement on countries' environmental efficiency in CO₂ emissions: An application of conditional full frontiers. *J. Prod. Anal.* **2014**, *41*, 367–382. [[CrossRef](#)]
39. Sanz-Díaz, M.T.; Velasco-Morente, F.; Yñiguez, R.; Díaz-Calleja, E. An analysis of Spain's global and environmental efficiency from a European Union perspective. *Energy Policy* **2017**, *104*, 183–193. [[CrossRef](#)]
40. Vlontzos, G.; Pardalos, P. Assess and prognosticate green house gas emissions from agricultural production of EU countries, by implementing, DEA Window analysis and artificial neural networks. *Renew. Sustain. Energy Rev.* **2017**, *76*, 155–162. [[CrossRef](#)]
41. Yu, S.H.; Gao, Y.; Shiue, Y.C. A Comprehensive Evaluation of Sustainable Development Ability and Pathway for Major Cities in China. *Sustainability* **2017**, *9*, 1483. [[CrossRef](#)]
42. Dios-Palomares, R.; Alcaide, D.; Pérez, J.D.; Bello, M.J.; Prieto, A.; Zúniga, C.A. The Environmental Efficiency using Data Envelopment Analysis: Empirical methods and evidences. In *The stated of the Art for Bioeconomic and Climate Change*; Editorial Universitaria UNAN Leon, Ed.; Red de Bioeconomía y Cambio Climático: Cordoba, Spain, 2017; p. 48.
43. Cooper, W.W.; Park, K.S.; Pastor, J.T. RAM: A range adjusted measure of inefficiency for use with additive models, and relations to other models and measures in DEA. *J. Prod. Anal.* **1999**, *11*, 5–42. [[CrossRef](#)]
44. Delimitrou, C.; Kozyrakis, C. Paragon: QoS-aware scheduling for heterogeneous datacenters. In Proceedings of the Eighteenth International Conference on Architectural Support for Programming Languages and Operating Systems, Houston, TX, USA, 16–20 March 2013; pp. 77–88.
45. Isard, M.; Prabhakaran, V.; Currey, J.; Wieder, U.; Talwar, K.; Goldberg, A. Quincy: Fair scheduling for distributed computing clusters. In Proceedings of the ACM SIGOPS 22nd Symposium on Operating Systems Principles, Big Sky, MT, USA, 11–14 October 2009; pp. 261–276.
46. Delimitrou, C.; Sanchez, D.; Kozyrakis, C. Tarcil: Reconciling scheduling speed and quality in large shared clusters. In Proceedings of the Sixth ACM Symposium on Cloud Computing, Kohala Coast, HI, USA, 27–29 August 2015; pp. 97–110.
47. Grandl, R.; Ananthanarayanan, G.; Kandula, S.; Rao, S.; Akella, A. Multi-resource packing for cluster schedulers. *ACM SIGCOMM Comput. Commun. Rev.* **2015**, *44*, 455–466. [[CrossRef](#)]
48. Zaharia, M.; Borthakur, D.; Sen Sarma, J.; Elmeleegy, K.; Shenker, S.; Stoica, I. Delay scheduling: A simple technique for achieving locality and fairness in cluster scheduling. In Proceedings of the 5th European Conference on Computer systems, Paris, France, 13–16 April 2010; pp. 265–278.

49. Delimitrou, C.; Kozyrakis, C. Quasar: Resource-efficient and QoS-aware cluster management. In Proceedings of the 19th International Conference on Architectural Support for Programming Languages and Operating Systems, Salt Lake City, UT, USA, 1–5 March 2014; pp. 127–144.
50. Zhang, X.; Tune, E.; Hagmann, R.; Jnagal, R.; Gokhale, V.; Wilkes, J. CPI 2: CPU performance isolation for shared compute clusters. In Proceedings of the 8th ACM European Conference on Computer Systems, Prague, The Czech Republic, 15–17 April 2013; pp. 379–391.
51. Verma, A.; Pedrosa, L.; Korupolu, M.; Oppenheimer, D.; Tune, E.; Wilkes, J. Large-scale cluster management at Google with Borg. In Proceedings of the Tenth European Conference on Computer Systems, Bordeaux, France, 21–24 April 2015; p. 18. [[CrossRef](#)]
52. Fernández-Cerero, D.; Fernández-Montes, A.; Jakóbbik, A.; Kołodziej, J.; Toro, M. SCORE: Simulator for cloud optimization of resources and energy consumption. *Simul. Model. Pract. Th.* **2018**, *82*, 160–173. [[CrossRef](#)]
53. Hindman, B.; Konwinski, A.; Zaharia, M.; Ghodsi, A.; Joseph, A.D.; Katz, R.H.; Shenker, S.; Stoica, I. Mesos: A Platform for Fine-Grained Resource Sharing in the Data Center. In Proceedings of the 8th USENIX Conference on Networked Systems Design and Implementation, Boston, MA, USA, 30 March–1 April 2011; pp. 295–308.
54. Vavilapalli, V.K.; Murthy, A.C.; Douglas, C.; Agarwal, S.; Konar, M.; Evans, R.; Graves, T.; Lowe, J.; Shah, H.; Seth, S.; *et al.* Apache hadoop yarn: Yet another resource negotiator. In Proceedings of the 4th Annual Symposium on Cloud Computing, Santa Clara, CA, USA, 1–3 October 2013; p. 5.
55. Schwarzkopf, M.; Konwinski, A.; Abd-El-Malek, M.; Wilkes, J. Omega: Flexible, scalable schedulers for large compute clusters. In Proceedings of the 8th ACM European Conference on Computer Systems, Prague, The Czech Republic, 15–17 April 2013; pp. 351–364.
56. Armbrust, M.; Fox, A.; Griffith, R.; Joseph, A.D.; Katz, R.; Konwinski, A.; Lee, G.; Patterson, D.; Rabkin, A.; Stoica, I.; *et al.* A view of cloud computing. *Commun. ACM* **2010**, *53*, 50–58. [[CrossRef](#)]
57. Ruth, S. Reducing ICT-related carbon emissions: An exemplar for global energy policy? *IETE Tech. Rev.* **2011**, *28*, 207–211. [[CrossRef](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).