

Reducing the operational cost of cloud data centers through renewable energy

Original

Reducing the operational cost of cloud data centers through renewable energy / Lagana, D.; Mastroianni, C.; Meo, M.; Renga, D.. - In: ALGORITHMS. - ISSN 1999-4893. - 11:10(2018), p. 145. [10.3390/a11100145]

Availability:

This version is available at: 11583/2853831 since: 2021-01-28T14:29:34Z

Publisher:

MDPI AG

Published

DOI:10.3390/a11100145

Terms of use:



This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

Article

Reducing the Operational Cost of Cloud Data Centers through Renewable Energy

Demetrio Laganà ¹, Carlo Mastroianni ², Michela Meo ³  and Daniela Renga ^{3,*} 

¹ Eco4Cloud srl, Rende (CS) 87036, Italy; lagana@eco4cloud.com

² ICAR-CNR, Rende (CS) 87036, Italy; carlo.mastroianni@gmail.com

³ Politecnico di Torino, Department of Electronics and Telecommunications, Torino 10129, Italy; michela.meo@polito.it

* Correspondence: daniela.renga@polito.it; Tel.: +39-011-090-4193

Received: 31 July 2018; Accepted: 21 September 2018; Published: 27 September 2018

Abstract: The success of cloud computing services has led to big computing infrastructures that are complex to manage and very costly to operate. In particular, power supply dominates the operational costs of big infrastructures, and several solutions have to be put in place to alleviate these operational costs and make the whole infrastructure more sustainable. In this paper, we investigate the case of a complex infrastructure composed of data centers (DCs) located in different geographical areas in which renewable energy generators are installed, co-located with the data centers, to reduce the amount of energy that must be purchased by the power grid. Since renewable energy generators are intermittent, the load management strategies of the infrastructure have to be adapted to the intermittent nature of the sources. In particular, we consider EcoMultiCloud, a load management strategy already proposed in the literature for multi-objective load management strategies, and we adapt it to the presence of renewable energy sources. Hence, cost reduction is achieved in the load allocation process, when virtual machines (VMs) are assigned to a data center of the considered infrastructure, by considering both energy cost variations and the presence of renewable energy production. Performance is analyzed for a specific infrastructure composed of four data centers. Results show that, despite being intermittent and highly variable, renewable energy can be effectively exploited in geographical data centers when a smart load allocation strategy is implemented. In addition, the results confirm that EcoMultiCloud is very flexible and is suited to the considered scenario.

Keywords: geographical data centers; energy saving; renewable energy

1. Introduction

The number and scale of data centers (DCs) are rapidly increasing, as data centers are the primary means through which companies can satisfy their increasing demand for computing and storage resources, either by building private data centers or by offloading applications and services to external cloud providers. A major issue is that data centers consume a large amount of energy and that consumption is expected to increase at a significant rate.

It is estimated that data center electricity consumption will increase to roughly 140 billion kilowatt-hours annually by 2020, corresponding to about 50 large power plants, with annual carbon emissions of nearly 150 million metric tons. The financial impact for DC management is also huge, since a DC spends between 30% and 50% of its operational expenditure in electricity: the expected figure for the sector in 2020 is \$13 billion per year of electricity bills (updated information can be found on the web portal of the U.S. National Resources Defense Council, <http://www.nrdc.org/energy/data-center-efficiency-assessment.asp>).

The efficient utilization of resources in the data centers is therefore essential to reduce costs, energy consumption, carbon emissions and also to ensure that the quality of service experienced by users is adequate and adherent to the stipulated service level agreements. Efficiency is essential not only in single data centers, but also in geographically-distributed data centers, whose adoption is rapidly increasing. Major cloud service providers, such as Amazon, Google and Microsoft, are deploying distributed data centers to match the increasing demand for resilient and low-latency cloud services, or to interconnect heterogeneous data centers owned by different companies, in the so-called “inter-cloud” scenario.

In this scenario, the dynamic allocation and migration of workload among data centers can help to reduce costs, moving the workload where the energy is less expensive/cleaner and/or cooling costs are lower: the cloud provider has the option of choosing the destination site based on different criteria upon the reception of the user request. Specifically, the increasing adoption of renewable energy plants is a great opportunity for a more efficient management of distributed data centers. Each data center can get its electricity from different electricity providers or can adopt on-site renewable energy sources, which provide green energy such as solar and wind [1]. Moving applications and services to data centers that are equipped with renewable energy sources (RES) can lead to several benefits both for the data center provider, which can reduce the costs of acquiring grid energy, and for the society in general, thanks to a more intense exploitation of green energy and the reduction of carbon emissions.

Unfortunately, workload assignment and migration in a distributed environment involve very complex decision processes due the time-variability of electricity cost, the workload variability both within single sites and across the whole infrastructure and, when the adoption of RES is possible, the intermittent nature of green energy generation. In a geographically-distributed scenario, the peak hours of renewable energy generation can be different in each data center, due to the variability of meteorological conditions and the different time zones. This generates the need for moving the workload to the sites where and when the green energy is currently available. Indeed, while energy storage units can be appropriately used to defer the utilization of green energy for some time, this postponement comes at a significant cost related to the charging and discharging of batteries. The immediate usage of green energy is the most convenient option, but can be exploited only if the infrastructure is able to support the workload redistribution and if efficient algorithms are designed to drive these workload shifts. In the paper, we prove that it is possible to effectively use green energy to reduce operational cost with a smart workload distribution.

In this paper, we present a novel approach for the efficient exploitation of RES in a distributed data center scenario, by adapting and refining a workload management strategy already proposed in the literature, i.e., EcoMultiCloud [2]. EcoMultiCloud includes a hierarchical architecture for the management of geographically-distributed data centers and a set of algorithms that drive the assignment and migrations of virtual machines (VMs) on the basis of the technical and business objectives defined by the management. EcoMultiCloud is composed of two layers: at the lower layer, each site adopts its own strategy to distribute and consolidate the workload internally. At the upper layer, a set of algorithms—shared by all the sites—are used to evaluate the behavior of single sites and distribute the workload among them. This architecture offers several benefits, among which: (i) scalability, because the bigger problem of workload allocation is decomposed into smaller intra-data center and inter-data center problems; (ii) autonomy of single data centers, since each data center can adopt its own algorithm for internal allocation; (iii) flexibility, since the algorithms can be easily adapted in accordance with the desired objectives.

In [2], the objectives that drove the workload redistribution were the load balancing among the sites and the minimization of costs; in that paper, the flexibility and feasibility of EcoMultiCloud were also verified with the support of analytical models. However, the availability of RES was not considered. In this paper, thanks to the flexibility characteristics mentioned above, we adapt the EcoMultiCloud algorithm to exploit the presence of renewable energy plants specifically. The EcoMultiCloud algorithm proves to be suited to the new, more challenging scenario, in which RES is

introduced. EcoMultiCloud accomplishes the objective to use renewable energy when and where it is generated; energy produced in excess with respect to what is consumed can be stored for future use. We analyzed a scenario with four data centers, located in various geographical areas and with different patterns of renewable energy generation. Performance results mainly concern the reduction in the use of grid energy and the cost savings that derive from this reduction. Results were derived when varying the size of photovoltaic panels and the size of batteries, which allowed us to determine the appropriate values of both parameters. Moreover, we evaluated the advantages deriving from the use of the novel strategy with respect to a strategy that does not exploit migrations and to a strategy that moves the workload randomly from the most costly data center to the other data centers. The main contribution of the paper is two-fold. On the one hand, the paper shows that, while it is intermittent and highly variable, renewable energy can be effectively used in geographically-distributed data centers to reduce operational costs due to the infrastructure power supply, given that a smart and dynamic workload management is adopted. On the other hand, the paper proves that EcoMultiCloud, used in the past only in scenarios with traditional power supply, is suited for the purpose and can be easily adapted to consider green energy in the scenario.

The paper is organized as follows. Section 2 describes some related work in the fields of geographical workload distribution; Section 3 summarizes the EcoMultiCloud architecture; Section 4 describes the EcoMultiCloud algorithm for workload redistribution that aims to maximize the use of green energy when and where it is produced; Section 5 illustrates the performance results obtained for a specific scenario including four data centers located in North America and Europe; finally, Section 6 concludes the paper.

2. Related Work

Many successful efforts have been made to increase the physical efficiency of data centers; for example, for its components devoted to cooling and power distribution, and this is confirmed by the general decrease of the PUE (power usage effectiveness index), the ratio between the overall power entering the data center and the power needed for the IT infrastructure. However, much remains to be done in terms of the computational efficiency: for example, on average, only a fraction of CPU capacity of servers—between 15% and 30%—is actually exploited, and this leads to huge inefficiencies due to the lack of proportionality between resource usage and energy consumption [3]. Improvements in this field are related to a more efficient management of the workload and a better use of the opportunities offered by virtualization. The efforts may be categorized into two big fields: workload consolidation within a single data center and efficient workload management in geographical infrastructures that include several remote data centers.

Workload consolidation is a powerful means to improve IT efficiency and reduce power consumption within a data center [4–7]. In [8], the authors presented a multi-resource scheduling technique to provide a higher degree of consolidation in multi-dimensional computing systems. In [9], a novel model was proposed for allocating virtual elements in a software-defined cloud data center, jointly minimizing the energy consumption deriving from computing, data transmission and VM migrations. The authors in [10] proposed an algorithm to manage the power states of the servers in a cloud data center, in order to minimize the electricity consumption and maintenance costs derived from the power variation on the servers' CPU. The algorithm was deployed taking into account the costs for VMs processing on the servers, for transferring data between the VMs and for migrating the VMs across the servers. Some approaches—e.g., [11,12]—try to forecast the processing load and aim at determining the minimum number of servers that should be switched on to satisfy the demand, so as to reduce energy consumption and maximize data center revenues. However, even a correct setting of this number is only a part of the problem: algorithms are needed to decide how the VMs should be mapped to servers in a dynamic environment and how live migration of VMs can be exploited to unload servers and switch them off when possible, or to avoid SLA violations.

Self-organizing and decentralized algorithms are proposed to improve scalability, since the problem of consolidation is known to be NP-hard. In [13], the data center was modeled as a P2P network, and ant-like agents explored the network and collected information needed to migrate VMs and reduce power consumption. The approach presented in [14] decentralized part of the intelligence to single servers that made decisions based on local information, using probabilistic functions, while a central manager coordinated servers' decisions to consolidate the workload efficiently.

The problem is even more complex in geographically-distributed data centers. Research efforts are focused on two related, but different aspects [15]: the routing of service requests to the most efficient data center, in the so-called assignment phase, and the live migration of portions of the workload when conditions change and some data centers become preferable in terms of electricity costs, emission factors or more renewable power generation.

Several studies explored the opportunity of energy cost-saving by routing jobs when/where the electricity prices were lower [16,17]. Some prior studies assumed that the electricity price variations and/or job arrivals follow certain stationary (although possibly unknown) distributions [18–20]. Rao et al. [21] tackled the problem taking into account the spatial and time diversity in dynamic electricity markets. They attempted to minimize overall costs for multiple data centers located in different energy market regions. Shao et al. in [22] studied the effect of transmission delay introduced by the routing of service requests and related data across DCs. The authors in [19] proposed a solution in which the power cost could be reduced under delay-tolerant workloads. By exploiting temporal and spatial variations of both workload and electricity prices, they provided a power cost-delay trade-off, which was exploited to minimize power expenses at the cost of service delay. The considered target applications that can generate delay-tolerant workloads are based on MapReduce programming, including searching, social networking and data analytics.

Liu et al. in [16] proposed a geographical load balancing (GLB) approach to route general Internet service-requests to data centers located in various geographical regions, by computing the optimal number of active servers at each data center. In [23], Yu et al. proposed a GLB algorithm to minimize energy cost and control the risks at the same time, as they modeled the uncertainties of price and workload as risk constraints. In [24], Luo et al. exploited the temporal and spatial diversities of energy price to trade service delay for energy cost. The authors proposed a novel spatio-temporal load balancing approach to minimize energy cost for distributed DCs. The algorithms presented in [25,26] tackled the problem considering the user's point of view and aimed to choose the most convenient data center to which the user should consign a service or VM.

The authors of [1] investigated different parameters that affect energy and carbon cost for a cloud provider with geo-distributed data center sites. The overhead energy of a data center, which is responsible for a portion of the energy consumption as large as half its value in old sites, is considered by defining a model for PUE as a function of data centers' IT load and outside temperature. The devised algorithms aimed to improve access to renewable energy sources, since data center sites that can get their power from renewable sources help the provider decrease its dependency on the electricity drawn from off-site grids, which are costly and less clean.

Inter-DC VM migration is a more novel research topic, as virtualization infrastructures have offered such features only recently: for example, the vSphere 6.0 release of VMware includes new long-distance live migration capabilities, which will enable VM migrations across remote virtual switches and data centers. While opportunities opened by long-distance migrations are big, the issues involved are also extremely complex: among them, determining whether the benefits of workload migrations overcome the drawbacks, from which site and to which site to migrate, what specific portion of the workload should be migrated, how to reassign the migrating workload in the target site, etc.

Some significant efforts have been done in this area. The electricity price variation, both across time and location, was exploited to reduce overall costs using different strategies. The Stratus approach [27] exploited Voronoi partitions to determine to which data center requests should be routed or migrated. Ren et al. [28] used an online scheduling algorithm based on Lyapunov optimization techniques. In [29],

Kayaaslan et al. proposed an optimization framework based on the observation that energy prices and query workloads showed high spatio-temporal variation for throughput-intensive applications like web search engines. The optimization framework was based on a workload shifting algorithm considering both electricity prices, to reduce the energy cost, and workload of data centers at the time of shifting, to reduce response time. Le et al. considered VM placement in the cloud for high performance applications [30]. The authors proposed VM migration policies across multiple data centers in reaction to variable power pricing. In order to adapt to the dynamic availability of renewable energy, the authors in [31] argued for either pausing VM executions or migrating VMs between sites based on local and remote energy availability.

Some recent works that coped with the real-time energy-efficient distribution of load in a geographical scenario were included in a special issue devoted to green and energy-efficient cloud computing [32]. The authors of [33] focused on the energy-efficient execution of scientific workflows that needed to be deployed across multiple data centers due to their large-scale characteristics. The optimal allocation of virtual machines must consider that workflow tasks have dependencies and communication constraints, which make them differ significantly from unrelated tasks. An energy consumption model was devised, and a corresponding energy-aware resource allocation algorithm was proposed for virtual machine scheduling. The energy-efficient management of geographically-distributed data centers was also the subject of [34]. The authors focused on the significant impact of “geotemporal input”, i.e., the time- and location-dependent factors that may impact energy consumption. Among such factors, they considered real-time electricity pricing enabled by the deregulated electricity market, the cooling efforts needed at different sites and different times and the availability of renewable energy. The scheduling of the VMs was tackled through a two-stage approach, which combined best-effort global optimization, driven by genetic algorithms, with deterministic local optimization for constraint satisfaction. Giacobbe et al. [35] used the idea of migrating VMs in a federated cloud environment to reduce costs and carbon footprint. They took advantage of dynamic electricity pricing to migrate the VMs to the data center with the lowest energy cost and move the virtual machines from a high carbon footprint source to data centers with the best access to renewable energy.

Most proposed approaches aimed to solve the problem as a whole, in a centralized fashion, undergoing the risk of originating three main issues, as discussed in the introductory section: poor scalability due to the size of the problem and the heterogeneity of involved business objectives, poor ability to adapt to changing conditions (e.g., changes in amount of workload, electricity price or carbon taxes) and lack of autonomy of single data centers. To cope with these issues efficiently, we believe that it is necessary to decentralize part of the intelligence and distribute the decisions points, while still exploiting the centralized architecture and functionalities offered by virtualization infrastructures in single data centers. This naturally leads to a hierarchical infrastructure, in which single data centers manage the local workload autonomously, but communicate with each other to route and migrate VMs among them. A self-organizing hierarchical architecture was proposed in [36], but so far, it has been limited to the management of a single data center. A recent study [37] proposed a hierarchical approach that combined inter-DC and intra-DC request routing. The VM scheduling problem was decomposed and solved at single data centers and was able to combine different objectives, e.g., minimize electricity cost, carbon taxes and bandwidth cost. While the work certainly deserves attention, it only solved the routing problem and did not exploit the opportunity of dynamic workload migration, nor does the approach seem to be easily extensible in that direction.

In [2], the authors presented EcoMultiCloud, a two-level architecture that is able to assign and redistribute the workload on the sites of a geographically-distributed infrastructure. The EcoMultiCloud algorithms for workload distribution were designed to achieve a better load balancing and a reduction of costs, but the same algorithms can be customized for a variety of technical and business goals. In this paper, the algorithms are adapted to exploit the availability of renewable energy in some or all the data center sites.

3. The System

This section is devoted to the description of the system architecture. As previously mentioned, the multi-site load management strategy considered in this paper is EcoMultiCloud [2]. EcoMultiCloud consists of a two-layer architecture in which the upper layer is used to exchange information among the different DCs so as to properly drive the distribution of VMs among the sites, while the lower layer is used to allocate the workload within single DCs.

At the lower layer, the DCs can use any load management strategy. In this paper, the DCs use the decentralized/self-organizing approach presented in [14] for the consolidation of the workload in a DC. The single DC solution dynamically consolidates VMs to the minimum number of servers and allows the remaining servers to enter low consuming sleep modes. This approach has been proven to be very efficient for the energy consumption reduction of individual data centers. At the lower layer, key decisions regarding the local data center are delegated to the servers, which autonomously decide whether or not to accommodate a VM or trigger a VM migration.

At the upper layer, global strategies are implemented by making decisions about workload allocation and inter-DC VM migrations. The decisions are made by combining some general information about single DCs. The hierarchical architecture, organized in two independent layers that exchange some (limited) information, allows upper layer algorithms to be modified independently on the lower layer algorithms. In this way, different strategies can be adopted locally at the DCs. Conversely, improvements of single sites can be implemented without any explicit notification or involvement of the upper layer, i.e., of other DCs, provided that information exchange between layers is maintained.

The reference scenario is depicted in Figure 1, which shows the upper and lower layers for two interconnected DCs. At each DC, a data center manager (DCM) runs the algorithms of the upper layer. The DCM integrates the information coming from the lower layer and uses it to implement the functionalities of the upper layer. The DCM communicates with the local manager (LM) and acquires detailed knowledge about the current state of the local DC, for example regarding the usage of host resources and the state of running VMs. Then, the DCM extracts relevant high level information about the state of the DC and transmits this high level information to all the other DCMs (upper layer). The algorithms at the upper layer combine the collected information and make decisions about the distribution of the workload among the DCs. The assignment algorithm is used to decide to which DC a new VM should be assigned. Once the VM is delivered to the target site, the LM runs the lower layer algorithms to assign the VM to a specific host.

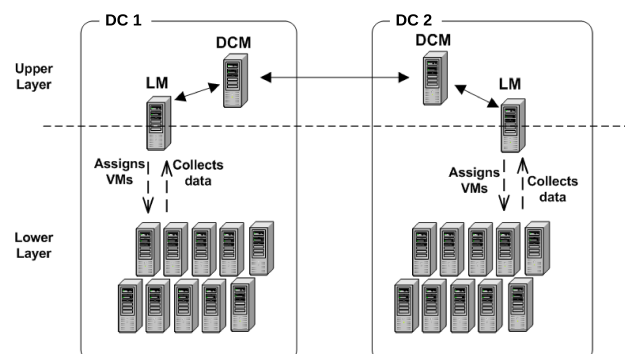


Figure 1. The EcoMultiCloud approach: the hierarchical two-level architecture in the case of two interconnected data centers. DCM, data center manager; LM, local manager.

As shown in Figure 2, the scenario considered in this paper includes four DCs. The DCMs execute the upper layer algorithms by exchanging information among themselves. While the approach requires information exchanges that scale quadratically with the number of DCs, the number of interconnected sites is expected to be small. Moreover, the amount of information to be distributed is tiny, i.e., of the order of a few packets, since it contains only a general description of the status of the DC: only pieces of information such as the load, the PUE and the electricity price need to be exchanged. The periodicity of the information exchange is of the order of an update per one or a few minutes, meaning that less than 1 Kbps of information exchange is needed. Two basic algorithms are executed at each DCM: (i) the assignment algorithm that determines the appropriate target DC for each new VM; (ii) the migration algorithm that periodically evaluates whether the current load distribution is appropriate, decides whether an amount of workload should be migrated and, if so, determines from which source site to which target site.

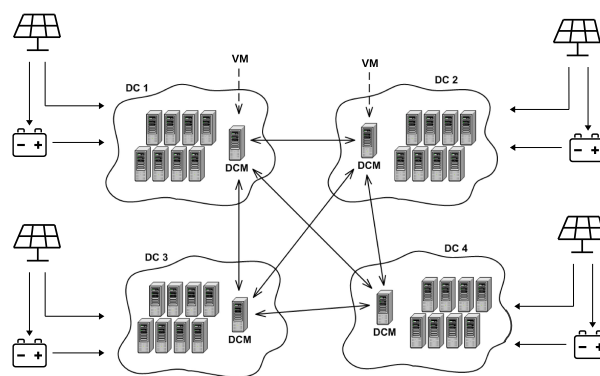


Figure 2. The considered scenario: four DCs equipped with PV panels and energy storage units use EcoMultiCloud for coordinating workload management among the DCs.

The DCs are equipped with renewable energy (RE) generators, such as PV panels or wind turbines, that contribute to reducing electricity costs by generating some green energy. Since they depend on meteorological phenomena such as solar radiation or wind speed, RE generators are intermittent, and their production is time variable and not easy to predict. For this reason, to better exploit what is produced, the DCs can also be equipped with energy storage units. Several losses occur when energy is stored, both during charging and discharging processes. Hence, it is convenient to use RE when it is generated and store only what is produced in excess with respect to what is consumed. Stored energy can be used when the RE production is smaller than the consumption and the electricity price from the traditional power grid is high.

Summarizing, in the considered scenario, the adaptive workload management, EcoMultiCloud, has to take into account a number of variables that influence the cost: the price of brown electricity taken from the power grid, the efficiency of the site (PUE), the amount of produced RE, the amount of stored energy.

4. Workload Assignment and Migration

As mentioned in the previous section, a key responsibility of the DCM is to analyze detailed data about the local data center and summarize relevant information that is then transmitted to remote DCMs and used for the assignment and redistribution of the workload. EcoMultiCloud is very flexible and can take into account several pieces of high level information. This information changes based on the general objectives of the multi-site infrastructure. Possible objectives are: the reduction of consumed energy, the reduction of carbon emissions, the quality of service in terms of the amount of

resources devoted to each VM, the load balance among the sites, the reduction of data transmitted among DCs, and so on.

All the above-mentioned goals are important, yet different data centers may focus on different aspects, depending on the specific operating conditions and on the priorities prescribed by the management. It is up to the company's management to specify the objectives and their relative weights. In this paper, the focus being the benefits of RE generation in multi-site DC operation and the associated reduction of operational costs, the algorithm uses information such as the amount of produced RE, the price of electricity taken from the power grid and the efficiency of the DCs.

In EcoMultiCloud, a key role in the distribution of the workload among the DCs is played by the assignment function. The assignment function is sort of a general key performance indicator that is associated in any moment with each DC. The assignment function balances and weights the chosen business goals, combining the DC information that was mentioned above. The numerical value that the assignment function associates with a DC represents the cost to run some workload in that DC: low values correspond to low overall cost of the DC. The strategy, then, is to assign a VM to the DC with the lowest value of the function. In its general formulation, the assignment function f_{assign}^i for each DC i , is defined as follows:

$$f_{assign}^i = \sum_{i=1}^M \alpha_i \frac{F_i}{F_{max}} \quad \sum_{i=1}^M \alpha_i = 1 \quad (1)$$

where M is the number of performance indicators and the coefficients α_i are the weights that decide the balance among the various targets, and they are decided by the system administrator based on the strategic decisions on how the system should work. The terms F_i are the performance indicators that represent the various targets, for example carbon emissions, overall utilization and energy cost. In order to combine indicators that have different units of measure and are not directly comparable, each indicator is normalized with respect to the maximum values communicated by DCs. After computing the value of f_{assign} for each DC, the VM is assigned to the data center having the lowest value. Once consigned to the target DC, the VM is locally managed and allocated to a physical host using the local assignment algorithm. Clearly, the choice above can also be combined with other rules that include QoS constraints (e.g., it can be requested that the distance/delay between the DC and the user is smaller than a given target) or decisions based on the kind of VM (e.g., interactive versus non-interactive VMs).

Abbreviations reports the notation used throughout the paper.

In this paper, the main objective is to evaluate the impact of the introduction of RE generation into the operation of geo-distributed DCs. Hence, the assignment function is modified and adapted in the following way:

$$f_{assign}^i = (C_i - E_i) P_i \quad (2)$$

where C_i is the consumed energy estimated for the additional workload, E_i is the green produced energy that is still available and P_i is the price of energy from the power grid. Notice some main differences with respect to the traditional EcoMultiCloud formulation. First of all, since only one key target is identified, no normalization is needed. Second, the assignment function can assume also negative values. Indeed, when the consumption is lower than the green production, $C_i < E_i$, there is an extra production of energy; the cost of allocating some workload to the DC is null. Conversely, when the consumption is higher, $C_i > E_i$, the green energy production is not sufficient for all the assigned workload; some energy has to be purchased from the power grid at the price P_i , turning into the cost $(C_i - E_i) P_i$.

Despite these differences between the traditional EcoMultiCloud formulation in (1) and the adapted version (2), the decision on the VM allocation is still taken considering the minimum among the values f_{assign}^i . For positive values, this means the minimum cost; for negative values, this corresponds to taking the DC with the largest cost saving.

The energy consumption term, C_i , is computed taking also into account the efficiency of the DC in terms of PUE. Denoting by $C^{(c)}$ the consumption for a VM that has to be assigned in terms of computation and storage, the consumption in DC i is given by:

$$C_i = PUE_i \cdot C^{(c)} \quad (3)$$

Although the value of PUE may vary with the load, in this paper, for simplicity, we assume that the PUE value of each DC is known and constant, but no modification of the algorithms is required if the PUE changes with the load, as long as it is known. The consumption of a VM, $C^{(c)}$, is assumed to be the same on all the servers of any data center.

For the coordination of the upper layer, each DCM transmits to the other DCMs the following vector of values, which corresponds to the state of the DC:

$$s_i = \{C_i, E_i, P_i\} \quad (4)$$

Figure 3 reports the pseudo-code used by a DCM to choose the target data center, among the N_{DC} data centers of the system, for a VM originated locally. First, the DCM requests the values of C_i , E_i and P_i for all the remote data centers; the amount of exchanged data is very low, but it can be refreshed periodically to reduce the amount of information exchange when the flow of VMs is large. Then, it computes the assign function according to (2) for any data center that has some spare capacity, i.e., for which the utilization of the bottleneck resource has not exceeded a given threshold U_{T_i} . Finally, the VM is assigned to the DC that has the lowest value of (2). Once assigned to the target DC, the VM is allocated to a physical host using the local assignment algorithm.

```

function AssignmentAlgorithm
  while VM arrives
    for each remote data center  $DC_i$ 
      Request/retrieve values of  $C_i, E_i, P_i$ 
    end for
    for each  $DC_i$  :
      if  $DC_i$  is not full, that is,  $U_i < U_{T_i}$ 
         $f_{assign}^i = (C_i - E_i) P_i$ 
      else
         $f_{assign}^i = \infty$ 
      end if
    end for
     $DC_{target} = DC_j$  such that  $f_{assign}^j = \min\{f_{assign}^i \mid i = 1 \dots N_{DC}\}$ 
    Assign VM to  $DC_{target}$ 
  end while
end function

```

Figure 3. The EcoMultiCloud assignment algorithm, executed by the DCM of each data center.

The assignment algorithm optimizes the distribution of the VMs on the basis of energy cost, taking into account the available RE. The effect of the assignment process is that the values of the f_{assign} for the various DCs tend to converge. This convergence accomplishes the performance maximization objectives. One of the advantage of the approach is that it naturally and smoothly adapts to changing conditions, i.e., variations of energy price or renewable energy production. However, the dynamicity of VM arrivals and departures determines the speed at which the system adapts to variations of energy production, consumption and price. To speed up adaptation and mitigate

temporary inefficiencies in the VM distribution due to changing conditions, inter-DC VM migrations are performed. Migrations redistribute part of the workload so as to adapt to new conditions.

The migration algorithm is triggered when the values of the f_{assign} function of two DCs differ by more than a predetermined threshold. The frequency at which this condition is evaluated should depend on the dynamism of the specific scenario and on the cost of migrations in terms of delay; i.e., on the frequency at which the price of energy and RE generation vary and on the typical lifetime of VMs, as well as on the bandwidth that is available for migrations. When such an imbalance is detected, VMs are migrated from the data center having the highest value of f_{assign} to the data center with the minimum value, until the values reenter within the tolerance range. The frequency of migrations is limited by the bandwidth between the source and target data centers. This bandwidth may correspond to the physical bandwidth of inter-DC connections or may be a portion of the physical bandwidth reserved by data center administrators for this purpose. In some cases, a few migrations might be needed between two DCs simply to balance some load fluctuations that make the f_{assign} functions of the DCs differ more than desired. These events typically require only a few migrations. In other cases, a batch of migrations is instead necessary to compensate abrupt changes of the f_{assign} function in a DC, for example due to some electricity price variations. When this happens, the process as described above translates into a sequence of VM migrations between pairs of DCs until a new balance among all the f_{assign} functions is reached. When the abrupt change makes a DC become the best performing DC, multiple VM migration requests will be made by the other DCs; to avoid congestion on the access links of the receiving DC, the VM migration process can be easily coordinated by the DCM of the receiving DC.

5. Results

This section is devoted to the evaluation of the benefits of the introduction of RE in a complex multi-site cloud computing infrastructure and of the effectiveness of EcoMultiCloud as an approach to maximize the performance of the system.

The scenario under analysis is the same as in [38,39], with four interconnected DCs and values of the PUE as reported in Table 1; time zones are also indicated with respect to UTC, assuming that the DC locations are, respectively, California, Ontario (Canada), the U.K. and Germany. Figure 4 reports energy prices in a 24-h interval; again, time is expressed in UTC. Energy prices are taken or extrapolated from the following websites:

- California: www.pge.com/tariffs/IndustrialCurrent.xls
- Ontario: <https://www.hydroone.com/rates-and-billing/rates-and-charges/electricity-pricing-and-costs>
- U.K.: www.westernpower.co.uk/docs/system-charges/LC14-v6-5-SWEB-Final-13-14.aspx
- Germany: www.iwr-institut.de/en/press/background-informations

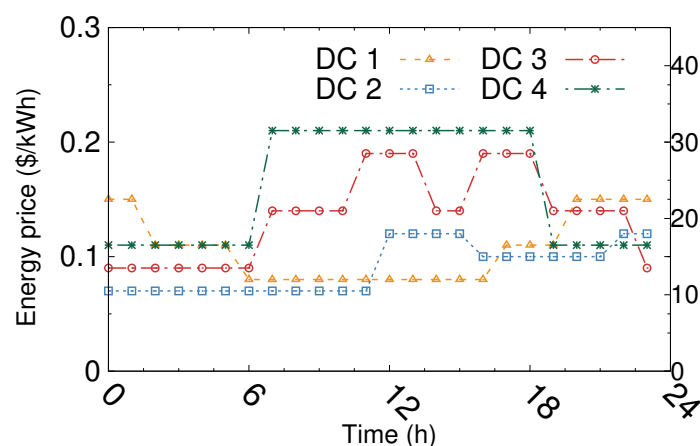


Figure 4. Energy price, expressed as \$/kWh, for the 4 DCs, versus time of the day.

To simplify the analysis, it is assumed that the prices repeat periodically for a few days. For the computation of the term C_i , the PUE of DC i is considered as in Table 1.

Table 1. Power usage effectiveness (PUE) and local time of the four DCs in the examined scenario.

Data Center	PUE	Local Time
DC 1	1.56	UTC-8
DC 2	1.7	UTC-5
DC 3	1.9	UTC
DC 4	2.1	UTC+1

Each DC was equipped with a set of PV panels, made up by multiple modules enabling the conversion of solar radiation into electricity with an efficiency that depends on the adopted PV technology. The nominal capacity of a PV module corresponds to the maximum output power obtained by the PV device under standardized environmental conditions (including a light intensity of 1000 W/m^2 and a temperature of $25 \text{ }^\circ\text{C}$) from the conversion of solar radiation into electricity, and it is commonly measured in peak Watts (W_p). The efficiency of currently available commercial modules may achieve 20% [40]. Considering some among the most efficient PV modules built with traditional technologies, i.e., crystalline silicon, and currently available on the market, a nominal capacity of 0.333 kWp was observed for modules with an area of 1.63 m^2 , hence a PV panel surface of about 4.9 m^2 could be assumed per each kWp of PV panel capacity [41]. The PV panel capacity is denoted as S_{PV} . Real RE generation profiles obtained from the tool PVWatts were used in the simulations. This tool allows one to retrieve real traces of RE production per kWp of PV panel capacity during the typical meteorological year in a given location [42].

In order to address the intermittent RE production, each DC was equipped with a variable number of lead-acid batteries for harvesting purposes, this being one of the most common storage technology adopted in PV systems [43]. Storage elements with a capacity of 200 Ah and voltage of 12 V were considered, corresponding to 2.4 kWh of nominal storage capability. In this work, the battery capacity, expressed in kWh, is denoted as B . An average charge efficiency of 85% is commonly assumed for lead-acid batteries [44]; when considering also the discharging process, the total energy efficiency was estimated to be 75% [45], meaning that for each energy unit (1 kWh) drawn from the storage, around 1.33 kWh of RE must have been produced and harvested in the battery bank.

Data about VMs and physical hosts were taken from the logs of a proof of concept performed by the company Eco4Cloud srl (www.eco4cloud.com), a spin-off from the National Research Council of Italy, on the DC of a telecommunication operator. The DC contained 112 servers virtualized with the platform VMware vSphere 5.0. Among the servers, 76 were equipped with processor Xeon 24 cores and 100-GB RAM and 36 with processor Xeon 16 cores and 64-GB RAM. All the servers had network adapters with a bandwidth of 10 Gbps. The servers hosted 2000 VMs, which were assigned a number of virtual cores varying between one and eight and an amount of RAM varying between 1 GB and 16 GB. Several types of VMs are represented in the dataset, featuring differences in their use of hardware resources, i.e., CPU, RAM and disk. The most utilized resource in this scenario was the RAM; therefore, the RAM utilization of DC i was considered when computing the utilization U_i . A constraint imposed by the DC administrators was that the utilization of server resources must not exceed 80%, i.e., $U_T = 0.8$. Servers and VMs were replicated for all the DCs, while the values of PUE and energy price were differentiated as described above. Notice that the DC was quite small (112 servers only); however, we preferred to use this case since it was a real scenario from which we obtained real data. While relatively small, the scenario could well be used to accomplish the paper objectives to prove the benefits of the use of RE in geographically-distributed DCs and the effectiveness of EcoMultiCloud.

The performance was analyzed with an event-based Java simulator that was previously validated with respect to real data for the case of a single DC [14]. At a time $\text{UTC} = 0$, corresponding to

midnight for DC 3, which is located in the U.K., all the VMs were assigned one by one by executing the assignment algorithm described in Section 4.

Results were obtained for a total load $\Lambda = 50\%$. Since the RAM was the bottleneck resource, the overall load Λ of the system was defined as the ratio between the total amount of RAM utilized by the VMs and the RAM capacity of the entire system. Thus, the overall number of VMs was chosen so as to load the whole system to the desired extent. VMs were assumed to be launched at different rates during the day and the night, namely λ_{day} and λ_{night} , and that $\lambda_{day} = 2 \cdot \lambda_{night}$, with an average lifetime of 5 h per VM. When the VM migration algorithm was applied, the f_{assign} value of each DC was checked every hour, and workload migration was triggered when f_{assign} values of two DCs differed by more than 3%.

Simulations were performed over a period of seven days, considering several combinations of PV panel and storage capacity values. The size of the PV panels may be different for each DC. At the beginning of the simulation period, the battery units were assumed to be empty.

5.1. Impact of Renewable Energy on Cost Saving

We start by reporting in Figure 5 the amount of hourly energy production for a few days in the various DCs. First of all, observe the typical intermittent behavior of RE generation, with production during daytime and no production during night. In addition, since the solar irradiance was taken from real traces, there was some difference in the behavior of different days, as can be clearly seen from the peak values. Second, notice that the different time-zones in which the DCs are located created a shift in the typical intermittent pattern. An effective approach for performance maximization should adapt workload allocation among the DCs taking into account the overall patterns of RE generation, moving the load in accordance to where RE is generated and cost results lower.

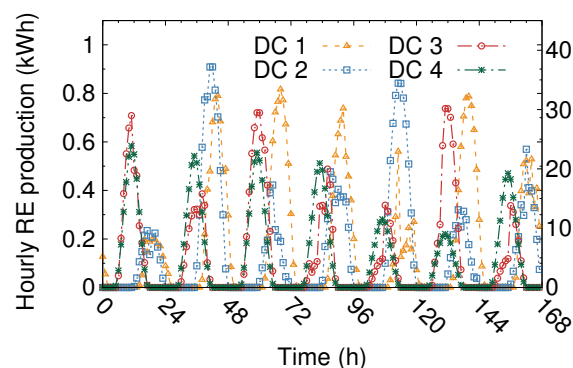


Figure 5. Production of RE in each considered DC location.

Figure 6 reports the hourly energy consumption and the hourly energy cost for each DC in a sample day. As can be observed from Figure 6a, the energy consumption of each DC showed a limited variability among DCs and during the day, with a minimum consumption level of slightly more than 50% the maximum observed hourly energy demand. Conversely, as can be evinced from Figure 6b, the energy costs varied much more over time and from DC to DC, due to energy price fluctuations. In particular, considering the time window from midnight to 6:00, despite similar energy consumption levels in the different DCs, the cost for the energy consumption showed huge variability over different DCs, with the range of variation being as high as 80% the maximum registered hourly energy cost. A proper VM assignment was hence required in order to place energy-demanding VMs on DCs where the energy price was lower.

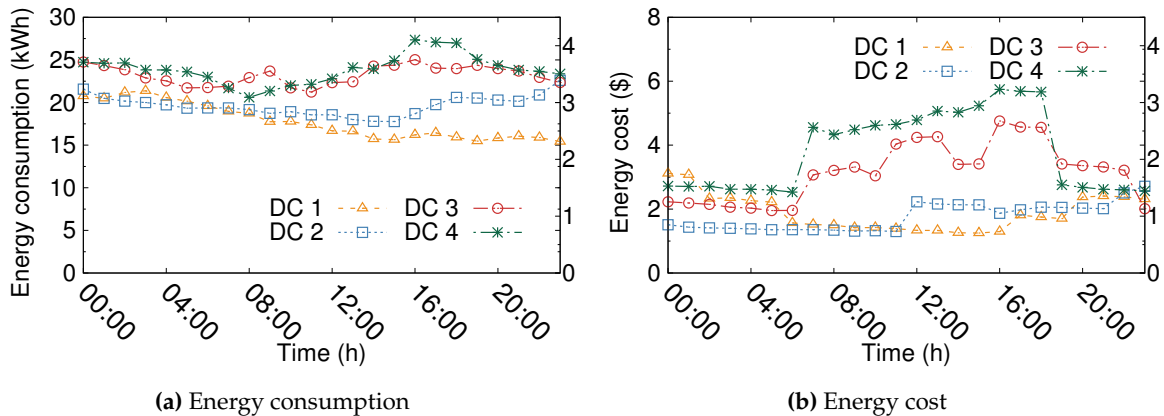


Figure 6. Energy consumption, expressed as kWh, and energy cost, expressed as \$, for the four DCs, versus time of the day.

To understand the benefits of RE generation in the considered scenario, Figure 7 reports the weekly energy cost (Figure 7a) and cost-saving in percentage (Figure 7b) for several cases in which different values of the PV panel size and battery capacity were considered. Clearly, as the PV panel size increased, the amount of produced RE increased, and energy cost decreased. However, the advantages that were large for small RE plants reduced when a condition was reached such that most of the power supply was provided by RE. No additional saving was possible by increasing the PV panel size. Battery capacity had a similar impact. Large values of batteries allowed a better utilization of the produced RE. Once battery capacity was such that all the produced RE was used, no additional advantages could be achieved by using larger values of the battery capacity. Considering that the yearly energy cost was of the order of more than 90,000 \$ in this scenario with four DCs, the cost saving may achieve more than 80,000 \$ per year. Furthermore, since DCs can be composed of thousands of servers rather than just one to a few hundred, this saving might be as high as 5.5 million \$ per year. Since 1 kWp corresponds to 5 m² of PV panels, the case reported in the figure corresponded to quite large areas (around 1000 m² for the 200 kWp case). This is a potential limitation to the adoption of RE, whose feasibility strongly depends on the environment in which the DC is located.

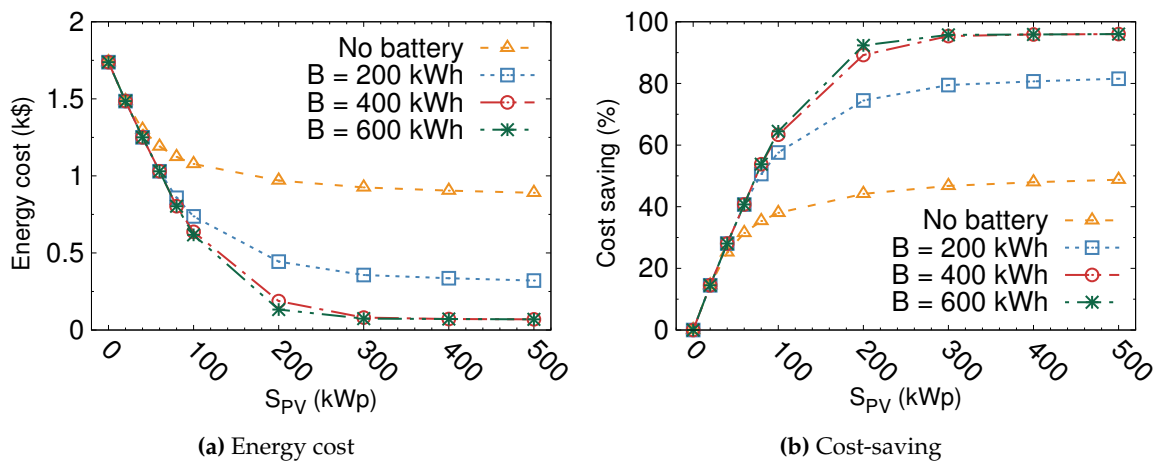


Figure 7. Energy cost and cost-saving. The DCs have the same PV panel size, reported on the x-axis.

5.2. Benefits of Workload Distribution

The workload distribution and its adaptation to RE production can be observed in Figure 8, which reports the number of VMs per DC for the whole simulation, whose duration was one week. The day-night pattern is clearly visible, as well as the temporal shift due to the different time-zones in which the DCs are located. The approach followed pretty closely the RE production patterns, showing a good capability of adaptation to operation conditions that were changing.

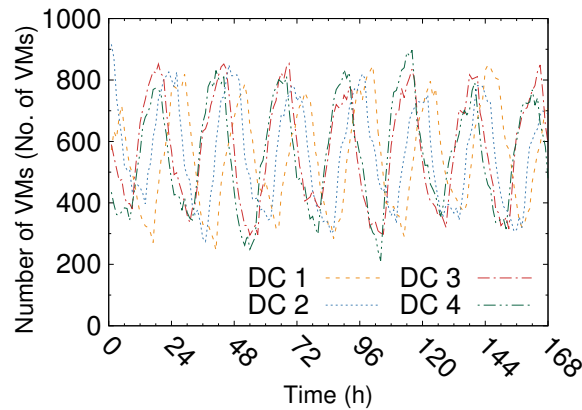


Figure 8. Number of VMs per DC versus time.

In Figure 9 the impact of migrations on the scenario can be assessed by observing the energy cost-saving for different sizes of the PV panels installed in each of the four DCs, either without any storage (Figure 9a) or with battery units providing a storage capacity of 200 kWh and 400 kWh (Figure 9b,c). The lowest curve refers to the case in which no migration was performed (orange curve). For comparison purposes, two other cases were considered as well: workload that was moved from the most costly DC randomly to the other DCs (blue curve) and workload that was moved targeting energy-saving (red curve). Observe from Figure 9a that the possibility to migrate some VMs can improve cost-saving by up to 17%, even without any storage, with the energy-saving migration policy only slightly outperforming the random migration policy. The introduction of a storage of capacity 200 kWh (Figure 9b) provided similar cost-saving percentages under VM migration, but its impact in terms of absolute decrease of the energy bill was higher than in the case without any storage. The same curves are shown in Figure 9c for a battery capacity that is double the value considered in Figure 9b; i.e., the battery is 400 kWh instead of 200 kWh. Saving reached 100%, for large values of the PV panels, meaning that the DCs were powered all the time with RE. In this case, migrations were useless. For intermediate values of the PV panels, migrations helped in improving saving, similarly to the previous case.

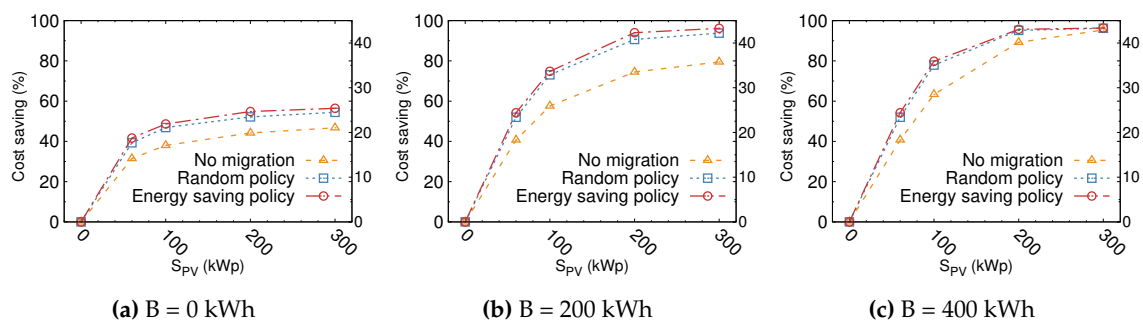


Figure 9. Cost-saving versus PV panel size for different migration policies; battery capacity equal to 0 kWh, 200 kWh and 400 kWh.

5.3. Proper Placement of Photovoltaic Panels

RE generators reduce operational costs, but they require an initial investment. Given the total investment, i.e., given the total size of the PV panels and battery capacity that the company is willing to invest in, the distribution of the generators among the DCs can have a different impact on the overall cost, which depends on a combination of the variables that are involved in the proposed approach: energy price, irradiation, values of PUE of the sites. Figure 10 reports cost-saving (in dark blue) and the amount of energy that was bought from the power grid (in light blue) in two scenarios, both corresponding to a total of 200-kWp PV panels. In the first scenario, the PV panels were installed in the various DCs in proportion to their value of PUE; while the opposite was done in Scenario 2, in

which panels were installed preferentially in those DCs with high efficiency and low PUE. The details of the scenarios are reported in Table 2.

Table 2. Scenarios for the distribution of PV panels.

	DC 1	DC 2	DC 3	DC 4
Scenario 1	20 kWp	40 kWp	60 kWp	80 kWp
Scenario 2	80 kWp	60 kWp	40 kWp	20 kWp
Scenario A	20 kWp	20 kWp	20 kWp	20 kWp
Scenario B	40 kWp	0 kWp	40 kWp	0 kWp
Scenario C	0 kWp	0 kWp	80 kWp	0 kWp

The total amount of energy purchased from the power grid was the same. However, Scenario 1 allowed achieving 25% higher cost-saving, suggesting that it is better to invest in the power supply of those DCs that consume more, so as to get the highest benefit from the green energy that is produced.

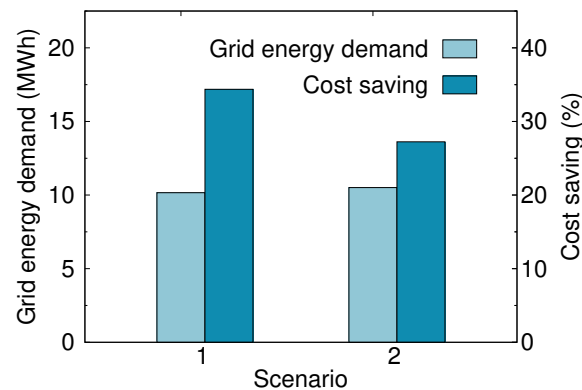


Figure 10. Cost-saving and total grid energy demand for the scenarios 1 and 2 reported in Table 2.

However, price also should be taken into account. To further investigate the PV panel distribution strategies, Figure 11 reports the savings that can be achieved when the PV panels are installed in one DC only, according to what is reported in the x -axis. Observe that cost-saving does not only depend on PUE, but it depends also on the price of electricity, which is, on average, 0.1, 0.09, 0.14 and 0.16 \$/kWh per each of the DCs. The lowest saving of about 7% was obtained by placing the RE generator where the average energy price was the lowest, i.e., 0.09 \$/kWh, with an intermediate value of PUE. Slightly higher cost-savings were obtained by placing the PV panel in DC 1, having a similar average energy price and a slightly lower PUE with respect to DC 1. By placing the PV panels in DC 4, showing the highest average energy price (77% higher than for DC 1) and value of PUE, cost-saving results more than doubled. This behavior is partly due to the variability of energy prices over time and the location-dependent RE production levels. Indeed, the average daily RE production varied depending on the location, being 4.2, 3.7, 3.2 and 3 kWh per kWp of PV panel capacity in each of the different DC locations, respectively.

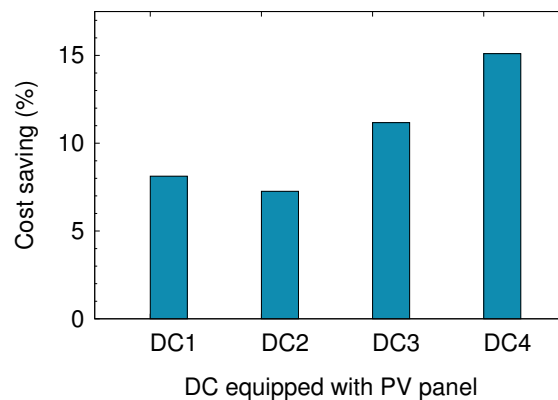


Figure 11. Cost-saving achieved by installing the PV panels in one DC only.

Finally, the effects of PV panel capacity consolidation is analyzed in Figure 12, which shows the cost-saving obtained in three different scenarios having the same total PV panel capacity of 80 kWp. In Scenario A, each DC was equipped with a 20-kWp PV panel; in Scenario B, two PV panels with a capacity of 40 kWp were installed in two DCs; in Scenario C, an 80-kWp PV panel was installed in a single DC. The different scenario configurations are reported in Table 2. For Scenarios B and C, the DCs equipped with PV panels were selected considering the configuration providing the highest cost-saving. Consolidating the RE generation capacity in 50% of the DCs allowed achieving the same cost-saving of almost 15%, which was obtained by distributing the RE generation capacity among all DCs. Conversely, if the same PV panel capacity was located in a single DC, only 10% of the energy bill could be saved. Hence, given the same initial investment for installing a PV panel system with the desired capacity, it was more convenient to select a limited number of DCs to be powered with RE, rather than equipping all DCs with some PV panels.

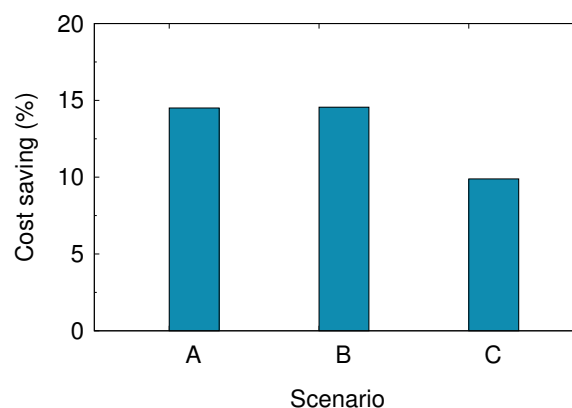


Figure 12. Cost-saving achieved by installing the same total PV panel capacity (80 kWh) according to different configuration scenarios: (A) 20 kWp per DC in four DCs; (B) 40 kWp per DC in two DCs; (C) 80 kWp in a single DC.

5.4. Impact of Assignment Function

The proposed workload assignment policy, which we denote as *min-cost_H policy*, aims at assigning VMs to DCs in which the energy cost is currently the lowest among the various DCs. In order to perform the workload assignment taking into account also the future energy price variations, which are observed during the period corresponding to the average VM lifetime, we defined a slightly different workload assignment policy, denoted as *min-cost_{LT} policy*. We assumed to know in advance the hourly energy prices and that the hourly energy demand was constant during the VM lifetime. Considering that no predictions about actual RE production were made and that the average VM

lifetime was 5 h, we assumed that the hourly RE generation was the same during the entire VM lifetime. The new utility function is the following:

$$f_{assign}^i = \sum_{t=h}^{h'} (C_i - E_i) P_{i,t} \quad (5)$$

where h represents the current time and h' corresponds to $h + lt_A - 1$, with lt_A being the average VM lifetime.

Based on our preliminary results, the performance of both policies in terms of cost-saving looked quite similar under any combination of PV panel and storage size. Only in a few cases, especially when some storage was present, the application of the *min-cost_H* policy may provide some benefit, but still negligible with respect to the *min-cost_{LT}* policy. Higher values of the average VM lifetime may affect the performance of the *min-cost_{LT}* policy, although we do not expect significant variations.

The fact that the *min-cost_H* did not significantly improve the cost-saving might be due, on the one hand, to the variability of the actual VM lifetime duration, whereas our strategy considered the average VM lifetime. On the other hand, it could be due to the slow variations in energy prices. Furthermore, the knowledge of the actual future RE production might impact on the predicted energy cost, hence affecting the assignment process. Future work is required to better study the performance of the latter proposed strategy under varying configuration settings and to investigate whether an accurate prediction of RE generation could play a relevant role in the VM assignment process, allowing further reduction of the energy cost.

6. Conclusions

We consider EcoMultiCloud, a flexible load management tool that implements multi-objective load management strategies, and we adapt it to the presence of renewable energy sources to power data centers. Our work aims at achieving cost reduction in the load allocation process in a multi-data center scenario, where VMs are assigned to a given data center by considering both energy cost variations and the presence of local renewable energy production, in order to reduce the energy bill. Performance is investigated for a specific infrastructure consisting of four data centers. Results show that geographical data centers can significantly benefit from renewable energy sources when a smart load allocation strategy is implemented.

Our results show that more than 60% cost can be saved even with a limited size of PV panels, with energy bill reduction as high as almost 100% under proper dimensioning of the PV panel and battery capacity. Scaling up to realistic scenarios, where data centers are made up of thousands of servers each, this means that the energy cost-saving can be of the order of several million dollars per year. Furthermore, combining the use of RE with a dynamic workload distribution by means of the application of migration policies allows further reduction of the energy bill by up to 17%, even without any storage and regardless of the type of adopted strategy. With large RE generators, the introduction of migration policies becomes less beneficial. Hence, the paper proves the feasibility of the introduction of renewable energy to reduce the operational costs of a complex infrastructure such as cloud data centers with geographically-distributed sites. Renewable energy can be effectively introduced given that a smart workload management is implemented; an approach like EcoMultiCloud is suited for this purpose.

Our results also highlight how the selection of the DCs in which the PV panels are installed is critical and different cost savings can be obtained, depending on the variable energy prices, the local RE production level and the PUE of each DC. In addition, given the same initial investment for a RE generator with a given capacity, although it is not necessary to equip all the DCs with some PV panels, it is more convenient to distribute the installation of PV panels on a fraction of DCs, rather than consolidating the RE generation capacity on a single DC. Future work is required to investigate

whether cost-saving might further benefit from the implementation of load management strategies that consider also the forecast future RE production to perform the VM assignment.

Author Contributions: Conceptualization, methodology, C.M. and M.M.; software, D.L. and D.R.; validation, formal analysis, investigation, resources, data curation, D.L., C.M., M.M. and D.R.; writing—original draft preparation, M.M.; writing—review and editing, D.L., C.M., M.M. and D.R.; visualization, D.R.; supervision, project administration, C.M. and M.M.

Funding: This research received no external funding.

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

Abbreviations

The following abbreviations are used in this manuscript:

N_{DC}	Number of DCs
f_{assign}^i	Assignment function for DC i
$F_i (F_{max})$	Key performance indicator for DC i (max among the DCs)
M	Number of key performance indicators
C_i	Hourly energy consumption for unit of workload in DC i
E_i	Available RE in DC i at a given time slot
P_i	Current hourly price of energy from the power grid in DC i
$P_{i,t}$	Hourly price of energy from the power grid in DC i at time slot t
α_i	Weights of the terms considered in the f_{assign}
PUE_i	Power Usage Effectiveness for DC i
Λ	Total workload
U_{T_i}	Maximum allowed workload in DC i
$\lambda (\lambda_{day}, \lambda_{night})$	Arrival rate of VMs (during daytime and nighttime, respectively)
S_{pV}	PV panel capacity expressed in kWp
B	Battery capacity expressed in kWh

References

1. Khosravi, A.; Andrew, L.L.H.; Buyya, R. Dynamic VM placement method for minimizing energy and carbon cost in geographically distributed cloud data centers. *IEEE Trans. Sustain. Comput.* **2017**, *2*, 183–196, doi:10.1109/TSUSC.2017.2709980. [[CrossRef](#)]
2. Forestiero, A.; Mastroianni, C.; Meo, M.; Papuzzo, G.; Sheikhalishahi, M. Hierarchical approach for efficient workload management in geo-distributed data centers. *IEEE Trans. Green Commun. Netw.* **2017**, *1*, 97–111, doi:10.1109/TGCN.2016.2603586. [[CrossRef](#)]
3. Dayarathna, M.; Wen, Y.; Fan, R. Data center energy consumption modeling: A survey. *IEEE Commun. Surv. Tutor.* **2016**, *18*, 732–794. [[CrossRef](#)]
4. Cardosa, M.; Korupolu, M.R.; Singh, A. Shares and utilities based power consolidation in virtualized server environments. In Proceedings of the 11th IFIP/IEEE Integrated Network Management (IM 2009), Long Island, NY, USA, 1–5 June 2009.
5. Graubner, P.; Schmidt, M.; Freisleben, B. Energy-efficient virtual machine consolidation. *IT Prof.* **2013**, *15*, 28–34, doi:10.1109/MITP.2012.48. [[CrossRef](#)]
6. Schröder, K.; Nebel, W. Behavioral model for cloud aware load and power management. In Proceedings of the 2013 International Workshop on Hot Topics in Cloud Services, Prague, Czech Republic, 21–24 April 2013.
7. Beloglazov, A.; Abawajy, J.; Buyya, R. Energy-aware resource allocation heuristics for efficient management of data centers for Cloud computing. *Future Gener. Comput. Syst.* **2012**, *28*, 755–768. [[CrossRef](#)]
8. Sheikhalishahi, M.; Wallace, R.M.; Grandinetti, L.; Vazquez-Poletti, J.L.; Guerriero, F. A multi-capacity queuing mechanism in multi-dimensional resource scheduling. In *Adaptive Resource Management and Scheduling for Cloud Computing*; Pop, F., Potop-Butucaru, M., Eds.; Lecture Notes in Computer Science; Springer International Publishing: Berlin, Germany, 2014; pp. 9–25.

9. Canali, C.; Chiaraviglio, L.; Lancellotti, R.; Shojafar, M. Joint minimization of the energy costs from computing, data transmission, and migrations in cloud data centers. *IEEE Trans. Green Commun. Netw.* **2018**, *2*, 580–595. [\[CrossRef\]](#)
10. Chiaraviglio, L.; D'Andreagiovanni, F.; Lancellotti, R.; Shojafar, M.; Melazzi, N.B.; Canali, C. An approach to balance maintenance costs and electricity consumption in cloud data centers. *IEEE Trans. Sustain. Comput.* **2018**, doi:10.1109/TSUSC.2018.2838338. [\[CrossRef\]](#)
11. Chen, Y.; Das, A.; Qin, W.; Sivasubramaniam, A.; Wang, Q.; Gautam, N. Managing server energy and operational costs in hosting centers. *SIGMETRICS Perform. Eval. Rev.* **2005**, *33*, 303–314. [\[CrossRef\]](#)
12. Mazzucco, M.; Dyachuk, D.; Deters, R. Maximizing cloud providers' revenues via energy aware allocation Policies. In Proceedings of the 10th IEEE/ACM International Symposium on Cluster Cloud and Grid Computing, Melbourne, Australia, 17–20 May 2010; pp. 131–138.
13. Barbagallo, D.; Di Nitto, E.; Dubois, D.J.; Mirandola, R. A bio-inspired algorithm for energy optimization in a self-organizing data center. In Proceedings of the of the 1st International Conference on Self-Organizing Architectures, Cambridge, UK, 14 September 2009; pp. 127–151.
14. Mastroianni, C.; Meo, M.; Papuzzo, G. Probabilistic consolidation of virtual machines in self-organizing cloud data centers. *Cloud Comput. IEEE Trans.* **2013**, *1*, 215–228. [\[CrossRef\]](#)
15. Kong, F.; Liu, X. A survey on green-energy-aware power management for data centers. *ACM Comput. Surv.* **2014**, *47*, 30:1–30:38. [\[CrossRef\]](#)
16. Liu, Z.; Lin, M.; Wierman, A.; Low, S.; Andrew, L. Greening geographical load balancing. *IEEE/ACM Trans. Netw.* **2015**, *23*, 657–671, doi:10.1109/TNET.2014.2308295. [\[CrossRef\]](#)
17. Qureshi, A.; Weber, R.; Balakrishnan, H.; Guttag, J.; Maggs, B. Cutting the electric bill for internet-scale systems. *SIGCOMM Comput. Commun. Rev.* **2009**, *39*, 123–134. [\[CrossRef\]](#)
18. Xu, D.; Liu, X. Geographic trough filling for internet data centers. In Proceedings of the 31st Annual International Conference on Computer Communications, Orlando, FL, USA, 25–30 March 2012; pp. 2881–2885, doi:10.1109/INFCOM.2012.6195720. [\[CrossRef\]](#)
19. Yao, Y.; Huang, L.; Sharma, A.B.; Golubchik, L.; Neely, M.J. Power cost reduction in distributed data centers: A two-time-scale approach for delay tolerant workloads. *IEEE Trans. Parallel Distrib. Syst.* **2014**, *25*, 200–211, doi:10.1109/TPDS.2012.341. [\[CrossRef\]](#)
20. Guo, Y.; Ding, Z.; Fang, Y.; Wu, D. Cutting down electricity cost in internet data centers by using energy storage. In Proceedings of the 2011 IEEE Global Telecommunications Conference-(GLOBECOM 2011), Houston, TX, USA, 5–9 December 2011, pp. 1–5, doi:10.1109/GLOCOM.2011.6134209. [\[CrossRef\]](#)
21. Rao, L.; Liu, X.; Liu, W. Minimizing electricity cost: Optimization of distributed internet data centers in a multi-electricity-market environment. In Proceedings of the International Conference on Computer Communications, San Diego, CA, USA, 19 March 2010.
22. Shao, H.; Rao, L.; Wang, Z.; Wang, Z.; Liu, X.; Ren, K. Optimal load balancing and energy cost management for internet data centers in deregulated electricity markets. *IEEE Trans. Parallel Distrib. Syst.* **2013**, *99*, 1, doi:10.1109/TPDS.2013.227. [\[CrossRef\]](#)
23. Yu, L.; Jiang, T.; Cao, Y.; Zhang, Q. Risk-constrained operation for internet data centers in deregulated electricity markets. *IEEE Trans. Parallel Distrib. Syst.* **2014**, *25*, 1306–1316. doi:10.1109/TPDS.2013.2297095. [\[CrossRef\]](#)
24. Luo, J.; Rao, L.; Liu, X. Data center energy cost minimization: A spatio-temporal scheduling approach. In Proceedings of the International Conference on Computer Communications, Nassau, Bahamas, 30 July–2 August 2013; pp. 340–344. doi:10.1109/INFCOM.2013.6566791. [\[CrossRef\]](#)
25. Li, W.; Svärd, P.; Tordsson, J.; Elmroth, E. Cost-optimal cloud service placement under dynamic pricing schemes. In Proceedings of the 6th IEEE/ACM International Conference on Utility and Cloud Computing, Dresden, Germany, 9 December 2013; pp. 187–194.
26. Goiri, I.; Le, K.; Guitart, J.; Torres, J.; Bianchini, R. Intelligent placement of data centers for internet services. In Proceedings of the 31st International Conference on Distributed Computing Systems (ICDCS), Minneapolis, MN, USA, 21–24 June 2011; pp. 131–142, doi:10.1109/ICDCS.2011.19. [\[CrossRef\]](#)
27. Doyle, J.; Shorten, R.; O'Mahony, D. Stratus: Load balancing the cloud for carbon emissions control. *Cloud Comput. IEEE Trans.* **2013**, *1*, 116–128. doi:10.1109/TCC.2013.4. [\[CrossRef\]](#)

28. Ren, S.; He, Y.; Xu, F. Provably-efficient job scheduling for energy and fairness in geographically distributed data centers. In Proceedings of the 32nd International Conference on Distributed Computing Systems (ICDCS), Macau, China, 18–21 June 2012; pp. 22–31, doi:10.1109/ICDCS.2012.77. [[CrossRef](#)]
29. Kayaaslan, E.; Cambazoglu, B.B.; Blanco, R.; Junqueira, F.P.; Aykanat, C. Energy-price-driven query processing in multi-center web search engines. In Proceedings of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval, Beijing, China, 24–28 July 2011; pp. 983–992.
30. Le, K.; Bianchini, R.; Zhang, J.; Jaluria, Y.; Meng, J.; Nguyen, T.D. Reducing electricity cost through virtual machine placement in high performance computing clouds. In Proceedings of 2011 International Conference for High Performance Computing, Networking, Storage and Analysis, Seattle, WA, USA, 12–18 November 2011; pp. 22:1–22:12.
31. Akoush, S.; Sohan, R.; Rice, A.; Moore, A.W.; Hopper, A. Free lunch: Exploiting renewable energy for computing. In Proceedings of the 13th USENIX Conference on Hot Topics in Operating Systems, Napa, CA, USA, 9–11 May 2011; p. 17.
32. Bianchini, R.; Khan, S.U.; Mastroianni, C. Guest editors' introduction: Special issue on green and energy-efficient cloud computing: Part I. *IEEE Trans. Cloud Comput.* **2016**, *4*, 119–121. doi:10.1109/TCC.2015.2506298. [[CrossRef](#)]
33. Xu, X.; Dou, W.; Zhang, X.; Chen, J. EnReal: An energy-aware resource allocation method for scientific workflow executions in cloud environment. *IEEE Trans. Cloud Comput.* **2016**, *4*, 166–179. doi:10.1109/TCC.2015.2453966. [[CrossRef](#)]
34. Lucanin, D.; Brandic, I. Pervasive cloud controller for geotemporal inputs. *IEEE Trans. Cloud Comput.* **2016**, *4*, 180–195, doi:10.1109/TCC.2015.2464794. [[CrossRef](#)]
35. Giacobbe, M.; Celesti, A.; Fazio, M.; Villari, M.; Puliafito, A. Evaluating a cloud federation ecosystem to reduce carbon footprint by moving computational resources. In Proceedings of the 2015 IEEE Symposium on Computers and Communication, Larnaca, Cyprus, 6–9 July 2015; pp. 99–104. doi:10.1109/ISCC.2015.7405500. [[CrossRef](#)]
36. Feller, E.; Rilling, L.; Morin, C. Snooze: A scalable and autonomic virtual machine management framework for private clouds. In Proceedings of the 12th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing, Washington, DC, USA, 13–16 May 2012; pp. 482–489.
37. Xiang, X.; Lin, C.; Chen, F.; Chen, X. Greening geo-distributed data centers by joint optimization of request routing and virtual machine scheduling. In Proceedings of the 7th IEEE/ACM International Conference on Utility and Cloud Computing, London, UK, 8–11 December 2014.
38. Forestiero, A.; Mastroianni, C.; Meo, M.; Papuzzo, G.; Sheikhalishahi, M. Hierarchical Approach for Green Workload Management in Distributed Data Centers. In *Proceedings of the 20th International European Conference on Parallel and Distributed Computing, Porto, Portugal, 25–29 August 2014*; Springer: Porto, Portugal, 2014; Volume 8805, LNCS, pp. 323–334.
39. Khosravi, A.; Garg, S.; Buyya, R. Energy and carbon-efficient placement of virtual machines in distributed cloud data centers. In *Euro-Par 2013 Parallel Processing*; Wolf, F., Mohr, B., Mey, D., Eds.; Springer: Berlin, Germany, 2013; Volume 8097.
40. Morgera, A.F.; Lughi, V. Frontiers of photovoltaic technology: A review. In Proceedings of the 5th International Conference on Clean Electrical Power, Taormina, Italy, 16–18 June 2015; pp. 115–121.
41. Renga, D.; Meo, M. Modeling renewable energy production for base stations power supply. In Proceedings of the International Conference on Smart Grid Communications, Sydney, Australia, 6–9 November 2016; pp. 716–722.
42. Dobos, A.P. *PVWatts Version 5 Manual*; Technical Report NREL/TP-6A20-62641; National Renewable Energy Laboratory: Denver West Parkway Golden, CO, USA, 2014, doi:10.2172/1158421.
43. Hua, S.; Zhou, Q.; Kong, D.; Ma, J. Application of valve-regulated lead-acid batteries for storage of solar electricity in stand-alone photovoltaic systems in the northwest areas of China. *J. Power Sources* **2006**, *158*, 1178–1185. [[CrossRef](#)]

44. Stevens, J.W.; Corey, G.P. A study of lead-acid battery efficiency near top-of-charge and the impact on PV system design. In Proceedings of the Conference Record of the Twenty Fifth IEEE Photovoltaic Specialists Conference, Washington, DC, USA, 13–17 May 1996; pp. 1485–1488.
45. Gharavi, H.; Ghafurian, R. IEEE 1013-2007—IEEE Recommended Practice for Sizing Lead-Acid Batteries for Stand-Alone Photovoltaic (PV) Systems. In Proceedings of the IEEE, Detroit, MI, USA, 24–28 July 2011; Volume 99, pp. 917–921.



© 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/>).