



Sustainable Computing: Informatics and Systems

Volume 40, December 2023, 100903

Operations of data centers with onsite renewables considering greenhouse gas emissions

José Luis Ruiz Duarte ^a Neng Fan ^b

Show more

Share Cite

<https://doi.org/10.1016/j.suscom.2023.100903>

[Get rights and content](#)

Highlights

- Job migration and scheduling for a data center network considering carbon emissions.
- Energy management systems manage the data centers' onsite and grid energy resources.
- Robust optimization is applied to integrate the intermittency of renewables.
- Numerical experiments demonstrate the effectiveness of the proposed methodology.

Abstract

In a highly technology-dependent society, processing large amounts of data has become essential, resulting in a growing number of data centers. The global data center industry accounts for a significant proportion of the world's energy consumption. A partial or total transition of the energy sources that power the operation of data centers to cleaner energy, particularly onsite renewable technologies, will help to achieve the greenhouse gas emission reduction goals established in international agreements. However, integrating such technologies is difficult due to the intermittency and availability of renewable sources. Nevertheless, temporal and geographical flexibility for processing part of the data centers' workload provides an opportunity to efficiently integrate these renewable energy sources. In this paper, we develop an operational scheme for a network of interconnected data centers with a

significant electricity supply from onsite technologies consisting of (i) an optimal job migration and scheduling policy for data centers to match their operation and the availability of renewable energy under an emissions cap-and-trade program; (ii) an energy management system to control the data centers' onsite renewable and grid energy resources; and (iii) a robust optimization approach to account for the intermittency of the renewables and to identify the optimal migration and scheduling policies under the worst-case scenario. A case study is carried out to assess the economic and environmental impacts. The results indicate that the installation of onsite technologies and an optimal migration policy reduce the operational costs and the greenhouse gas emissions of the networks of data centers.

Introduction

The use of information technology (IT) services has become inherent to our daily living as part of belonging to an interconnected society. According to the World Bank [1], the number of internet users grew from 413 million (6.74% of the world's population) in 2000 to 3.4 billion (45.79% of the world's population) in 2016. Also, the internet traffic has increased from 60 Petabytes in 1997 to 1.1 Zettabytes in 2017 [2]. The increasing demand for internet-based data services to acquire, analyze, process, store, retrieve, and disseminate data, as well as the trend of numerous organizations to offload their IT demands via cloud computing, has led to substantial growth in the data center (DC) industry [3], [4], [5]. A DC is a facility that houses computing resources networked together using communication infrastructure for data storage and application hosting [6]. The growing number of DCs worldwide is important due to their resource usage and the services they provide to society.

The energy consumption of the DCs industry has gained attention due to its total energy impact. The global electricity usage of communication technologies accounted for 8%–14% of the world's consumption in 2010, and by 2030 the consumption can lead to 8% in the best-case scenario, 21% in the expected scenario, and up to 51% in the worst-case scenario [7]. The combined energy usage of DCs accounted for 194 TWh in 2010 [8]. Optimizing energy usage in DC elements crucial for their operation, such as CPU computations and servers cooling, has resulted in a deceleration of energy consumption increase compared with the growth of computation power and number of services [9]. Masanet et al. [8] estimate that the global data center energy consumption in 2018 was 205 TWh, around 1% of global electricity consumption; it means that, in the period from 2010 to 2018, the data center compute instances increased by 550%, while the energy consumption grew only 6%. Nevertheless, the authors suggest that these energy efficiency trends may find an inflection point when they reach their physically feasible limits, causing electricity consumption to increase with a slope more aligned to the computation instances trends.

From a business model point of view, energy consumption within a DC accounts for almost half of its operating expenses [10]. Therefore, there is an interest in optimizing the DCs load scheduling to improve their energy efficiency. For extensive coverage of green energy-aware power management for DCs, including an overview of power utilization, work scheduling, virtual machines management, and capacity planning, see [6], [11], [12]. Most data centers provide services for critical interactive applications that run 24 × 7, like internet services, and delay-tolerant batch-style applications, such as scientific or financial jobs. Generally, the latter can be scheduled to be processed anytime as long as they finish before deadlines [13], providing flexibility for load-shifting. The efficiency of a DC can be achieved by balancing the utilization of the servers [14], [15], scheduling batch jobs to periods when the energy price is lower [9], and migrating jobs to other DCs [16]. Different approaches consider individual job performances, quality of service, traffic demands, deadline constraints, and energy consumed by the DC [17], [18], temporal and geographical variation in the price of grid energy, green energy usage [19], [20], [21], heterogeneous resources of the DCs [22], the energy consumption of servers and cooling systems [23], virtual machine migration [24], energy source control, power system monitoring and architectural support for power-aware computing [25], and a holistic vision for the selection and capacity planning the energy sources, including renewables, grid energy,

and storage for a green DC at the construction phase [26]. The optimization approach has been applied to improve energy efficiency in DC services, including mixed-integer programming [27], [28], [29], [30], [31], constrained convex optimization [13], and genetic algorithms [32]. In their work, Pierson et al. [27] present a mixed-integer linear program to allocate the resources of one DC with thermal constraints while assigning the workload to the set of heterogeneous servers of the DC; this project does not include aspects such as networked DCs, emissions analysis, energy supply and management, and exact solutions for larger instances. The research of Haddad et al. [33], [34] presents an optimization model to determine the size and operation of onsite solar and wind renewable energy sources (RES) and energy storage systems to power DCs operations while maintaining the quality of service. The problem under study consists of a negotiation module that coordinates the IT operations and the power management, where the optimization model decides the power supply while the demand is treated as a parameter; aspects such as DCs optimal scheduling, networked DCs, emissions analysis, energy supply from the grid, and uncertainty of the RESs are not considered. Sharma et al. [21] present a system to schedule the workload of a DC, prioritizing their execution during periods with higher renewable availability. The system predicts the energy availability and workload, and a scheduler allocates the jobs to geographically distributed DCs for planning horizon; elements such as RESs uncertainty, greenhouse gas (GHG) emission analysis, and the costs associated with the scheduling are not considered. Zhu et al. [35] propose a matching and resource allocation algorithm for a multi-cloud system with heterogeneous servers, maintaining the quality of the service while balancing the workload among the servers and minimizing the variance of the estimated completion time; other aspects, such as emissions analysis, energy supply and management, and exact solutions are not contemplated in the research. Yuan et al. [36] develop a queueing model to minimize the operational cost of assignment of tasks of a distributed green cloud system, considering the price of grid energy and availability of renewable energy, which is formulated as a constrained optimization model solved by heuristic methods; elements such as onsite green technologies, GHG emission analysis, and uncertainty of the RESs are not considered. In a similar approach, Yuan et al. [37] propose an optimization model to maximize the revenue of job processing while analyzing distributed DCs that are powered by the grid, as well as solar and wind energy sources; elements such as batch-type jobs, GHG emission analysis, and uncertainty of the RESs are not considered. Following the research pipeline, Yuan et al. [38] propose a bi-objective optimization problem to jointly minimize the energy cost while maximizing the revenue obtained from processing jobs by applying an improved multiobjective evolutionary algorithm based on decomposition. Liu et al. [39] study how geographically balancing the loads of DCs can enhance the use of renewable energy while considering demand-side management policies for the price of grid energy. The proposed model minimizes both the energy cost and the delay cost of the operations. Goiri et al. [40] propose a greedy algorithm for a parallel batch job scheduler for a DC powered (mainly) by solar RES and the grid (as a backup). The proposed approach tries to maximize the use of renewable energy and use grid energy during low-price periods; this project does not include aspects such as networked DCs, emissions analysis, and the uncertainty of the RESs. In addition, from a real-world application point of view, Goiri et al. [41] construct a solar-powered DC and an information system for optimal scheduling of its workloads. When considering the different elements discussed in this Section, different research projects have been proposed to increase the energy efficiency of DCs. To maximize the economic benefits of DCs, it is important that their energy management systems (EMS) efficiently control the grid and onsite energy resources by leveraging on the flexibility provided by the type of jobs being processed on the DC [8], [42].

On the other hand, the intense energy consumption of the DCs industry impacts the environment due to GHG emissions, accounting for approximately 98 million metric tons in 2017, corresponding to the combined emissions of Peru and Portugal [2], [43]. The Paris Agreement [44] is the most ambitious climate action carried out by the international community to reduce the world's GHG emissions, and it has been reproduced in region-based specific strategies. For example, the State of California in the U.S. developed the Cap-and-Trade (C&T) Program of the Air Resources Board to reach the emissions reduction goals cost-effectively by setting a cap on the total amount of certain GHG facilities covered by the system can emit [45]. Within the cap, organizations receive or buy emission

allowances, which give the holder the right to emit one ton of carbon dioxide (CO_2) or the equivalent amount of nitrous oxide (N_2O) and perfluorocarbons. If a company reduces its emissions, a tradeable offset credit can be issued equivalently to a metric ton of CO_2e . Then, another entity can use it to meet up to a certain proportion of its compliance obligation. After one year, the company must have enough allowances and offset credits to cover all its emissions; otherwise, fines are imposed. More and more regions have begun implementing C&T mechanisms for emission-intensive industries in addition to the power generation sector [46], [47]; using RESs represents an opportunity for major energy consumers to reduce their carbon footprint.

Powering DCs with RES, such as photovoltaic (PV) and wind turbines, as a means for reducing their carbon footprint is becoming a trend and is aligned with international efforts such as the United Nations Sustainable Development Goals and the European Union and the State of California renewable mix targets [10], [42], [48], [49]. Major companies like Google, Facebook, Microsoft, Apple, and Verizon Communications have ambitious projects to reduce the environmental impact of web hosting by turning their DCs into green data centers (GDCs). A GDC is a type of DC that seeks to reduce its carbon footprint by powering its operation by RESs by engaging in power purchase agreements from third-party owners, by purchasing green energy offsets that account for one unit of brown energy used from green hosting providers, or by installing their own solutions to produce renewable energy to increase their power efficiency [50]. The latter requires integrating the electricity from the onsite and grid sources. There are mainly two options: either by using automatic transfer switches that isolate the electricity produced onsite and switch to the grid when the onsite generation is not enough, or by transforming electricity from the onsite sources into electricity with the same frequency and voltage as the grid before combining them, via a grid-tie inverter system [51], [52]. For example, the Apple's DC, located in Maiden, North Carolina, claims to be totally powered by renewable energy, of which 64% comes from the onsite solar farms [53]. In addition, the energy efficiency of a GDC can be increased by integrating energy storage systems (ESSs) to mitigate the intermittency of the renewables, adapting the workload to match the energy supply [42], and balancing the workload of geographically distributed GDCs according to their renewable energy availability [31]. By the abovementioned elements, powering the GDCs operation with RESs represents an operative challenge, while the required initial investments can be economically restrictive.

The DCs' grid energy demand curve is influenced by demand-side management (DSM) policies implemented by utility operators, either by actively participating in incentive-based programs such as reducing energy consumption during a time window in exchange for an economic stimulus, or by dynamic pricing programs [54], [55]. The latter includes the Time-of-Use (TOU) program that charges higher unitary energy price rates during peak periods. An extension of this program that also involves energy consumption is the Time-and-Level-of-Use (TLOU) [56]. The opportunities presented by DSM policies for diverse types of consumers have been evaluated in residential [57], [58], [59], [60], [61], [62], [63], [64], industrial [65], [66], and grid [67], [68], [69] levels. Hence, efficiently utilizing such policies can lead to a considerable reduction of the DCs energy costs.

This paper explores a new research direction to jointly optimize the operation of interconnected GDCs and their renewable energy resources, enabling the transition from DCs to GDCs cost-effectively. This is achieved by scheduling batch-style jobs within the time window defined by the user (temporal flexibility) and by migrating this type of job to other GDCs of the network to regions with better energy profiles (geographic flexibility). In particular, the operation cost can be minimized, and the use of renewable energy can be enhanced by shifting the job processing to time frames and locations to match a higher renewable availability and lower energy prices. To this end, a model to identify the optimal job migration, process scheduling, and energy management decisions of a network of interconnected GDCs for a planning horizon is developed. The decisions to be made in the model include job migration and process scheduling, operations of onsite energy storage, energy supply from the onsite RES and the grid under DSM policies, and carbon allowances and offset credits management for a GHG emissions C&T program.

The main contributions of this paper consist of (i) a new research direction that combines complex operating and energy management schemes to mitigate the environmental impact of networked GDCs while minimizing the operational costs; (ii) a tool consisting of an optimization model that identifies the migration and scheduling policy of a network of interconnected GDCs considering each center's energy profile, which can be used for short- and long-term assessment, e.g., to evaluate the acquisition of green technologies; (iii) an algorithm that integrates robust optimization and heuristics to identify the optimal scheduling and assignment of batch-style jobs under the worst-case scenario of renewables availability; and (iv) a comprehensive case study to show the effectiveness of the proposed scheme to evaluate the economic and environmental impact under different operating and investment policies. To the best of the authors' knowledge, the presented proposal is the first in the literature that includes the following elements simultaneously: (i) The utilization of exact optimization algorithms to optimize the workload migration and scheduling jointly and its resulting power profile, and the energy management system that satisfies the power demand, minimizing the total cost; (ii) The inclusion of both geographical and temporal flexibility for optimal operation; (iii) The consideration of DSM policies that affect the energy consumed from the grid; (iv) The use of onsite green technologies for power generation and energy storage; (v) The management of GHG emissions due to energy consumption; (vi) The explicit inclusion of uncertainty of the RESs into the solution procedure.

The remainder of the paper is organized as follows. In Section 2, the model for optimal job assignment and processing for GDCs under the C&T for reducing GHG emissions is developed. The solution methodology is presented in Section 3. Section 4 discusses the application of the proposed approach under different scenarios in a case study. Finally, the paper is concluded in Section 5.

Section snippets

Overview of green data centers

The system under study consists of a network of interconnected GDCs, which process both online and batch jobs during a time horizon, onsite renewable technologies, including solar PVs and energy storage to power the GDCs, and connected grid energy (see Fig. 1). It is considered that the energy consumption of the GDCs comes from the utilization of the servers, which are used to process online and batch jobs. Since batch-style jobs can be migrated among GDCs and processed anytime before the ...

Solution approach

The model presented in (15) is a two-stage mixed-integer linear program (MILP) with opposite **max** and **min** operators in the objective function of the second stage. To solve this problem, the two stages are analyzed separately. First, its inner part is reformulated by utilizing the Karush–Kuhn–Tucker conditions to find the dual of the minimization problem and merge the operations of the objective function. Then, this reformulation allows the application of a nested version of the ...

Case study

The model is utilized to find the operation that maximizes the environmental and economic benefits of a network of interconnected GDCs with different configurations to evaluate the economic and environmental impact of using the green energy technologies of ESS and RES and an optimal migration policy, as well as a sensitivity analysis in which the onsite technologies and optimal migration policy are not implemented. The case settings for the network are

presented in Section 4.1. Sections 4.2 ...

Conclusions

The participation of the service sector is necessary to achieve GHG reduction targets. Since the DC industry accounts for an important proportion of energy consumption, transitioning to cleaner energy sources to power DCs operations is a direct approach to reducing their carbon footprint. Onsite installation of direct-owned energy technologies, such as RES and ESS, is an option that major cloud computing companies are exploring. In this paper, a mathematical programming model for optimal job ...

CRediT authorship contribution statement

José Luis Ruiz Duarte: Conceptualization, Methodology, Software, Investigation, Writing – original draft, Visualization. **Neng Fan:** Conceptualization, Validation, Writing – review & editing. ...

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. ...

Acknowledgments

The authors also appreciate the discussion with Dr. Chris Gniady during the preparation of this manuscript. An allocation of computer time from the UA Research Computing High-Performance Computing (HPC) at the University of Arizona, USA is gratefully acknowledged. Vectors utilized in Figs. 1, 2, and 10 were created by Fullvector, Macrovector, Katemangostar/Freepik. ...

Recommended articles

References (85)

BilalK. et al.

[A taxonomy and survey on green data center networks](#)

Future Gener. Comput. Syst. (2014)

HuangS. et al.

[Congestion control in high-speed lossless data center networks: A survey](#)

Future Gener. Comput. Syst. (2018)

IrtezaS.M. et al.

[Efficient load balancing over asymmetric datacenter topologies](#)

Comput. Commun. (2018)

DubeyK. et al.

[A novel multi-objective CR-PSO task scheduling algorithm with deadline constraint in cloud computing](#)

Sustain. Comput.: Inform. Syst. (2021)

SharmaD. et al.

Scheduling computing loads for improved utilization of solar energy

Sustain. Comput.: Inform. Syst. (2021)

JiK. et al.

A joint energy efficiency optimization scheme based on marginal cost and workload prediction in data centers

Sustain. Comput.: Inform. Syst. (2021)

ZhangW. et al.

Dynamic multi-technology production–inventory problem with emissions trading

IIE Trans. (2016)

MahdiF.P. et al.

A holistic review on optimization strategies for combined economic emission dispatch problem

Renew. Sustain. Energ. Rev. (2018)

BehrangradM.

A review of demand side management business models in the electricity market

Renew. Sustain. Energ. Rev. (2015)

FarrokhifarM. et al.

Model predictive control for demand side management in buildings: A survey

Sustain. Cities Soc. (2021)



[View more references](#)

Cited by (11)

Edge-cloud collaboration for low-latency, low-carbon, and cost-efficient operations

2024, Computers and Electrical Engineering

[Show abstract](#) ▾

A two-stage stochastic collaborative planning approach for data centers and distribution network incorporating demand response and multivariate uncertainties

2024, Journal of Cleaner Production

Citation Excerpt :

...Li Duarte and Fan (2023) proposed a power provisioning architecture that is specially designed for DCs to coordinate workload balancing while switching power supply between renewable energy and the grid. Duarte and Fan (2023) proposed an energy management system to control the data centers' onsite renewable and grid energy resources. Gügül, G. N (2023) suggests that the most feasible way to achieve DC zero energy consumption is to combine cooling systems with photovoltaic(PV) generators, which demonstrates the importance of synergies between DCs and renewable energy sources....

[Show abstract](#) ▾

Carbon Emission Modeling for High-Performance Computing-Based AI in New Power Systems with

Large-Scale Renewable Energy Integration ↗

2025, Processes

Forecasting US data center CO₂ emissions using AI models: emissions reduction strategies and policy recommendations ↗

2024, Frontiers in Sustainability

Development of a PV/Battery Micro-Grid for a Data Center in Bangladesh: Resilience and Sustainability Analysis ↗

2023, Sustainability Switzerland

Computational Engineering based Approach on Artificial Intelligence and Machine Learning Driven Robust Data Centre for Safe Management ↗

2023, Journal of Machine and Computing



[View all citing articles on Scopus ↗](#)

[View full text](#)

© 2023 Elsevier Inc. All rights reserved.



All content on this site: Copyright © 2025 Elsevier B.V., its licensors, and contributors. All rights are reserved, including those for text and data mining, AI training, and similar technologies. For all open access content, the relevant licensing terms apply.

