VEDA - moVE DAta to balance the grid: research directions and recommendations for exploiting data centers flexibility within the power system

Amin Ziagham Ahwazi
amin.ziagham@gmail.com
The Arctic University of Norway
Tromsø, Norway

Chiara Bordin
The Arctic University of Norway
Tromsø, Norway
chiara.bordin@uit.no

Sambeet Mishra
Tallinn University of Technology
Tallinn, Estonia
sambeets@gmail.com

Phuong Hoai Ha
The Arctic University of Norway
Tromsø, Norway
phuong.hoai.ha@uit.no

Alexander Horsch
The Arctic University of Norway
Tromsø, Norway
alexander.horsch@uit.no

ABSTRACT

This paper aims at discussing visions and research directions to investigate the value of data centers flexibility within sustainable electrical energy systems. While optimizing the energy consumption and task scheduling within data centers located in different time zones and connected at national and international level, it is possible to balance the local power grids, to allow a better penetration of intermittent renewable energy sources, and a more economical way to address peak demand by avoiding or postponing costly investments in network expansion. Challenges and opportunities that behind the exploitation of data centers flexibility within sustainable electrical energy systems will be discussed. An interdisciplinary approach to tackle these kind of problems will be proposed, and visions for a novel framework called VEDA (moVE DAta to balance the grid) will be outlined.

CCS CONCEPTS

• **Mathematics of computing → Mathematical analysis**; **Mathematical optimization**; **Continuous optimization**; **Stochastic control and optimization**;

KEYWORDS

Data centers, Mathematical optimization, Scheduling, Power systems, Energy markets, Demand Response

ACM Reference Format:
Amin Ziagham Ahwazi, Chiara Bordin, Sambeet Mishra, Phuong Hoai Ha, and Alexander Horsch. 2021. VEDA - moVE DAta to balance the grid: research directions and recommendations for exploiting data centers flexibility within the power system. In *2021 6th International Conference on Systems, Control and Communications (ICSCC) (ICSCC 2021), October 15–17, 2021, Chongqing, China.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3510362.3510365

ICSCC 2021, October 15–17, 2021, Chongqing, China

ACM ISBN 978-1-4503-8900-6/21/10…$15.00
https://doi.org/10.1145/3510362.3510365

1 INTRODUCTION

Data centers (DC) contain thousands and maybe tens of thousands of servers that are interconnected in different structures and architectures to meet the growing computing and storage requests. Undoubtedly, data centers play an important role in our society today. Various services such as entertainment services, big data analysis, medicine, cyber-physical systems, scientific and engineering computing, e-commerce, etc. are all somehow connected to data centers and used for their services [17]. With the expansion of the global economy and the pervasiveness of the Internet, the trend of using data center services has also increased significantly. The computation and storage resources are added to the data centers annually in order to meet the needs. Data centers have been replicated in different places with different time zones in order to increase the quality of services (QoS). This increase in usage and adding extra resources has led to different challenges and problems. It also affects various aspects such as increasing energy consumption, and consequently rising data center costs, environmental impacts, and increased carbon dioxide emissions. Moreover, it puts pressure on the local power grid (especially at peak hours) that may cause fluctuations. Figure 1 shows the percentage of energy consumption in a data center. As can be seen in Figure 1, the highest percentage of energy consumption in a data center is related to the servers power consumption and cooling systems energy consumption, which together account for about 80% of the total energy consumption of a data center [19].

This high percentage indicates the importance of planning and managing energy consumption in data centers. Various solutions have been proposed to minimize the energy consumption of data centers, in order to eliminate the burden on the local power grid and reduce the server cost. Some of these solutions have been in the field of energy management, and others in the field of energy optimization.

2 RELATED WORKS

Energy management solutions refer to managing the energy consumption of server resources. If a resource is unused, it is temporarily deactivated so that no extra energy is consumed. Various

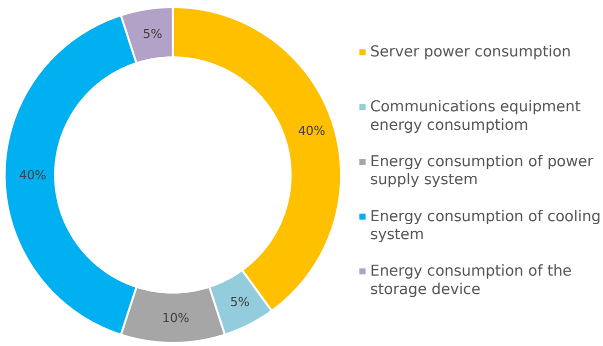

Figure 1: Energy consumption distribution of data centers [19]

studies have been done in this field such as [3, 9, 10]. PowerNap[15] proposes two fixed states: “Nap” and “Active”. When the workload is low, resources are taken to “Nap” state, otherwise the “Active” mode is switched. This way it is possible to accommodate workload variations. KnightShift[26] is a heterogenous server architecture that has active low power nodes (Knight). When the workload is low, KnightShift mechanism puts all servers into deep sleep, except Knight nodes and low-load tasks being served by Knight nodes. SleepScale[13] is a framework that uses different stages of sleep. On the other hand, to avoid wasting time and energy, it has included several other intermediate stages between the deep sleep phase and the active phase. This way, when the server is idle, it will gradually and step by step go into a deep sleep. Another mechanism used to manage the energy consumption in data centers is the Dynamic Voltage and Frequency Scaling (DVFS) method. If there is no workload on the CPU, its frequency is reduced to consume less power. In Pegasus[14], the request latency is monitored and calculated periodically. If a deadline violation happens, it increases the CPU frequency to the maximum, so that it can meet the objectives. But if the latency is less than 65% of the given time budget, the CPU frequency will be reduced.

On the other hand, in energy optimization solutions, all efforts are made in the field of optimal use of data center resources, such as CPU, Memory, Network equipment, etc. In the data center, even at peak loads, only 20-30% of the server’s CPU power is utilized[13]. Using servers with this low utilization means wasting energy. One of the most common methods used to enhance server resources utilization in data centers, is task scheduling. Using some efficient algorithms, tasks are distributed among server resources, so that all resources have the same load. This will both increase resource efficiency and guarantee the quality of services. Various work has been done on scheduling tasks in data centers, especially cloud computing [11, 18, 23, 25]. Some of them focus on optimal resource management and fair distribution of tasks between different resources. Others have benefited from the transfer of tasks to other data centers in different geo-locations, in addition to the local schedule. Chen et al.[6] schedule cloud computing tasks on data centers equipped with on-site Renewable Energy Sources (RESs) to maximize the amount of green energy used. Chen et al.[7] investigate the feasibility of using data centers to offer Demand Response (DR) via server consolidation and dynamic server power capping. In[28] Greenware was proposed as a dispatch system that coordinates the workload assignment of incoming requests among a set of available data centers. Mohsenian-Rad et al.[16] proposed a linear programming approach to reduce the maximum power flow in the power grid, by scheduling the interactive type workload to geo-distributed data centers.

2.1 Paper objectives and key contributions

The proposed paper aims at providing an introductory overview of potential solutions and techno-economic approaches, to investigate the value of data centers flexibility within sustainable electrical energy systems. While optimizing the energy consumption and task scheduling within data centers located in different time zones and connected at national and international level, it is possible to balance the local power grids, to allow a better penetration of intermittent renewable energy sources, and a more economical way to address peak demand by avoiding or postponing costly investments in network expansion. In light of the most recent literature in the field, the proposed paper aims at illustrating challenges and opportunities that lies behind the exploitation of data centers flexibility within sustainable electrical energy systems. Considering the identified challenges and opportunities, the paper will discuss visions for a preliminary novel design of an integrated, robust, and energy aware framework that, in addition to managing and optimizing the energy consumption in data centers, can also balance the pressure on the local power grid. An interdisciplinary approach will be introduced, outlining the main disciplines to be involved in such a study. Different strategies will be discussed, to reduce the data centers energy consumption and exploit data centers flexibility within demand response programs that will balance the power grid, and allow a higher penetration of renewable resources, as well as a better and more stable service for the final energy users. Finally recommendations for future research directions will be drawn, by also discussing the overall potential for academic and societal impact of such solutions.

3 DATA CENTERS FLEXIBILITY WITHIN THE POWER GRID

This section discusses the types of data centers, their applications, operations, and the reasons for the importance of studying them from energy consumption perspective. The data centers flexibility within the power grid can be practical and basically have opportunities in some ways, and naturally it will also be challenging in others. From this point of view, challenges and opportunities of data centers flexibility within the power grid are outlined and discussed.

3.1 Energy consumption within data centers

Based on the available literature, four types of data centers can be identified: Enterprise, Managed services, Colocation, and Cloud data centers[21]. Each of them has its own application and its own energy consumption accordingly. Energy consumption in data centers depends on the type of their application. This means that a high number of requests will naturally lead to higher energy consumption. For instance, entertainment services, due to the greater number of clients, energy consumption in the relevant data centers will be high. Data centers such as the Google cloud, which have a general application (e.g., Email, File sharing, Storage, Video service, etc.), naturally consume more energy than the other type of data centers (e.g., Enterprise).

In general, energy consumption in data centers has grown exponentially over the last two decades. From a global perspective, in 2010, the energy consumption by data centers was estimated to be in the range of 1.1-1.5% of the worldwide energy use and is expected to increase further in the near future[27]. Data centers are the principal electricity consumers in cloud computing, reportedly consuming approximately 70 billion kWh in 2014, equivalent to 1.8% of the US total energy consumption, and are projected to account for approximately more than 73 billion kWh in 2020 and beyond[20]. On the other hand, as the global economy continues to expand, energy consumption and, of course, carbon dioxide emissions will increase in the coming years. CO_2 emissions from ICT industry sources are increasing by about 6% per year, and with such a rapid growth rate, they will account for about 12% of CO_2 emissions worldwide in the years after 2020[19].

From the components perspective, a data center consists of different components (e.g., computing resources, cooling systems, network equipment, and etc.), and each of them have different pattern of energy consumption. Computing resources with cooling systems have the highest energy consumption rates in a data center and, together, they account for about 80% of all energy consumption in data centers[19].

3.2 Opportunities

Despite the fact that data centers consume a lot of energy, they also have opportunities, especially when it comes to provide flexibility within the power grid. Data centers can be very flexible in some ways. Indeed, they can communicate and interconnect with each other, given the capabilities of distribution systems. Hence, there is the ability to transfer portion of the load to another data center if needed. This will reduce some of the energy consumption within the data center. This reduction in energy consumption can benefit both the data center (from cost perspective, quality of service, etc.) and the local power grid. The transition the processing data to other data centers, especially during peak hours, can contribute to balance the local power grid.

From an economic and cost perspective, data centers can transfer data to the other data centers around the world, which have lower energy consumption tariffs or are in off-peak hours. However, this transition can also have challenges, like data sovereignty, that will be addressed in the next section. Another opportunity behind data centers flexibility within the power system, is the ability to balance the local power grid using optimal task scheduling and resource utilization. This refers to the ability to balance energy consumption in data centers to some extent, by designing an optimal task scheduling mechanism. In addition to increasing the quality of services, this balance can also reduce the load on the local power grid. However, designing the optimal task scheduling presents several challenges, that are addressed in the following section.

3.3 Challenges

There are several challenges behind the exploitation of data centers flexibility within the power system, that have not yet been addressed in the scientific literature. A summary is outlined below.

3.3.1 Data center's budget. One of the challenges in data centers that have not been addressed in literature, is the budget that a data center could spend. Energy costs can be a significant part of a data center budget and can greatly affect the power bills. Data centers participating in energy markets can expect savings in power bills, if they reduce their power usage during peak periods. Data centers are also able to increase their total energy consumption by operating more off-peak operations without having to pay more money[1]. Therefore, one way to optimize the data center budget is using optimization models. By using optimization models it is possible to determine the most economical budget for data centers, to get benefits from the participation in energy markets.

3.3.2 Optimal, flexible and holistic task scheduling. As mentioned before, designing task scheduling mechanism and optimal resource utilization in data center is the key challenge. Some scheduling mechanisms reviewed in literature use very limited and strict mechanisms that are not flexible enough to consider task conditions and other policies implications holistically. Indeed, most of the algorithms try to transfer tasks to another data center, without considering data sovereignty, service types, and other criteria. A more holistic approach would allow an improved quality of service.

3.3.3 Data sovereignty. When transferring tasks to another geographical location, an important challenge called "Data Sovereignty" arises. Some data is private and contains certain information such that they cannot be transferred cross-border. Policies may vary from region to region or from country to country. Even though data sovereignty is an important issue in policy adoption for data transferring, it has not been properly addressed in literature. Indeed, data sovereignty should be taken into account in order to protect the users data both at a national and international level.

3.3.4 Participation level. The participation level refers to the collaboration of data centers in order to process part of each other's data. Data centers can cooperate with each other at the national or international level. At the national level, the proximity of energy consumption price, is effective to reduce the load, both on the data center and on the local power grid. But at the level of international participation, energy costs vary. In some locations and within different time zones, the energy cost is lower and in some other locations it is higher. Choosing the right place to transfer data, in terms of geography, time zones, and costs, is key to fully exploit the data centers flexibility.

4 THE VEDA FRAMEWORK

4.1 VEDA - An interdisciplinary approach

The proposed approach is called VEDA that stands for "moVE DAta to balance the grid". It is highly interdisciplinary, touching upon key subjects from power systems, mathematical optimization, energy

Informatics[4], parallel computing, distributed computing, green computing and high performance computing. In the following paragraphs, these key subjects, the relationship between them, and their effects on reducing energy consumption in data centers will be discussed.

4.1.1 Mathematical optimization. Mathematical optimization is the process of maximizing or minimizing an objective function by finding the best available values across a set of inputs[8]. In the VEDA framework, the mathematical optimization contributes to: minimizing the energy consumption of data centers, optimizing the grid balance, optimizing the exchange of jobs between data centers, as well as optimizing the task scheduling within the data centers.

4.1.2 Distributed Computing. Computations and processing within or among data centers components by using message-passing mechanism are considered as distributed computing[22]. How data centers communicate with each other and through what protocols, and how tolerance faults and latency, and other related issues are all included in the VEDA framework as distributed computing concepts.

4.1.3 Parallel Computing. Parallel computing is defined as the process of breaking down larger problems into smaller ones that can be executed simultaneously by multiple processors[2]. One of the advantages of using mathematical optimization in combination with parallel computing techniques, is the possibility to address complex and computationally hard problems. Parallel computing represents therefore a key methodology in the VEDA framework for large scale case studies.

4.1.4 Green Computing. Green computing is a paradigm that allows computing components or resources to work effectively with minimal or no impact on the environment[12]. Green computing is directly related to utilization. This means that an active computing resource that has 30% utilization will have more energy wasted than when it has 80% utilization. In the VEDA framework, using the optimal task scheduling mechanism, an attempt is made to increase the utilization of the computing resources in the data center.

4.1.5 High Performance Computing. High-performance computing (HPC) most generally refers to the use of distributed computing facilities for solving problems that need large computing power. The general profile of HPC applications is constituted by a large collection of compute-intensive tasks that need to be processed in a short period of time. There are supercomputers and clusters that are specifically designed to support HPC applications[5]. The HPC is one of the key concepts of the VEDA framework, since the optimization models and the optimal task scheduling can have a significant impact on the energy efficiency and resources utilization of HPC clusters.

In Figure 2 the interconnection between the different disciplines involved in VEDA framework is shown. Mathematical optimization is closely related to green computing, and both of them are considered as the core of the VEDA framework. Indeed, mathematical optimization contributes to the optimal design, expansion, management and operation of energy and power systems. Mathematical optimization goes hand in hand with green computing. Indeed, it brings green computing into a broader perspective, by including the computers sustainability issues within the power and energy systems optimal design and operations. Then, in order to solve big instances and large scale problems, parallel and distributed computing is necessary. And eventually all of these disciplines will require high performance computing that is associated with data centers and high computational power.

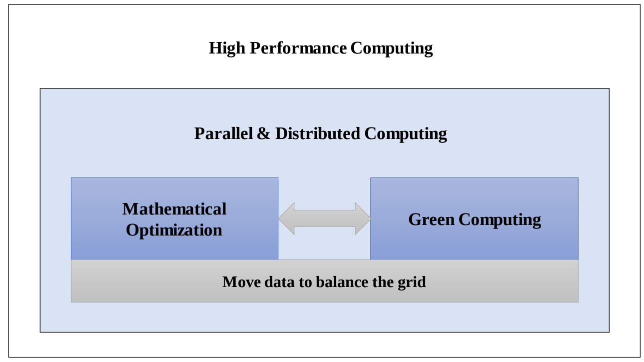

Figure 2: Interconnections between different methodologies and disciplines involved in VEDA framework

4.2 VEDA - Framework conceptualization

As mentioned before, data centers can communicate with each other and thus transfer part of their processing to other locations and other data centers for processing. Figure 3 shows an outlining map of participation and communication among data centers in Norway and other parts of the world. The blue points represent data centers within Norway, and the red points represent other parts of the world.

In Figure 3, the challenge of the level of participation at the national and international levels can be easily understood. Indeed, the blue dots, as a level of national participation, can interact with each other, without considering vital and important criteria such as data sovereignty or data security. However, to communicate with other data centers, at the level of international participation, some factors and criteria (e.g., energy cost and energy market mechanisms in the geographical area of the different data center, location in terms

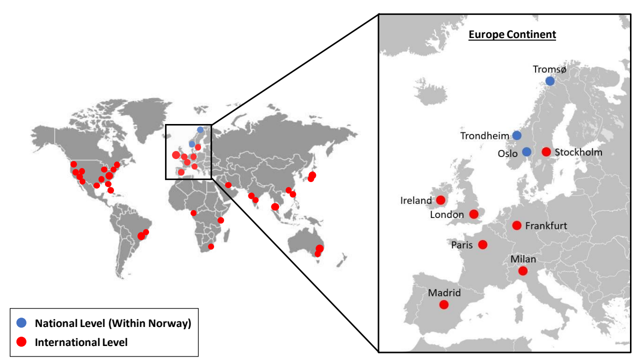

Figure 3: The participation and communication of data centers in Norway and the rest of the world at the national and international level

of distance and communication latency, time-zone out of peak hour, etc.) must be considered. All of these criteria are considered by the optimization model developed within VEDA framework, to select the best geographical location and specific data centers to shift the load when it is needed to balance the grid.

Figure 4 shows the simplified structure of a data center, where the location of the VEDA framework is also shown. From the architectural point of view, the VEDA framework is divided into two main parts: the VEDA Core and the VEDA Controller. The VEDA controller is located inside each rack and is responsible for collecting important input information as well as managing the servers. The most important tasks within the VEDA Controller are: Load Balancer, State Checker, calculation of the energy consumed by the servers, and DVFS Regulator. It also has a mutual communication with the VEDA Core, through which it sends the collected information, as well as receive new policies.

The other part of the VEDA framework is its kernel, which we call the VEDA Core, where the main processes and decisions are made. In a way, it can be said that optimal policies adopting, information processing, and scheduling of tasks are performed in this part. Based on the main architecture of the VEDA framework, the VEDA core is located between the arrival requests and the VEDA Controllers. As can be seen in Figure 4, the VEDA core is made up of several units (Predictor, Energy Consumption Calculator, Policy Manager, Job Departure, etc.). Each of them performs different and important tasks. The Runtime Predictor (RP) works epoch by epoch and is responsible for predicting parameters that will directly affect the overall performance of the system. This refers to parameters such as distribution of arrival requests and utilization prediction. The Energy Consumption Calculator (ECC) is responsible for collecting energy consumption information from various components of the data center such as cooling system, server components, lighting, network equipment, hard disks, etc. In the Logger part, at each epoch, information about the number of incoming requests (or arrival rate), the service time, and other important information specific to the final clients, are logged and stored. This information is used by the Runtime Predictor at the beginning of the next epoch.

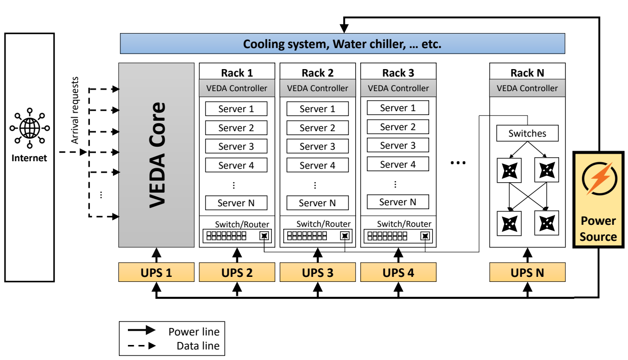

Figure 4: A simplified data center and related components as well as the location of VEDA framework

4.3 VEDA - Value for scientific community and society

The scientific value of the VEDA framework is both methodological and analytical. On the methodological side, the main impact of the VEDA framework is the development of new and innovative mathematical models and algorithms for the optimal management of data centers and for their integration within the power grids and the energy markets. On the analytical side, the VEDA framework can be utilized to perform sensitivity analyses to investigate the value of different data centers demand response strategies for a more renewable and reliable power system, as well as a greener society that moves towards decarbonization objectives. An open source version of the VEDA framework will motivate further research on the use of new techniques to develop novel optimization models and green energy consumption tools.

The VEDA framework has also the potential to strongly contribute to a greener society, and a smarter development of the overall energy and power systems infrastructures. This can impact the residential, commercial, and industrial sectors, enhancing their ability to utilize greener energy, in an optimized way, and with more reliable and stable services. In addition, the proposed VEDA framework has potential to contribute to the United Nations Sustainable Development Goals (SDGs) [24]. From an environmental point of view, reducing energy consumption in data centers can help reduce the production and emission of carbon dioxide and therefore contribute to a decarbonization of the energy system. This is in line with SDG 13: "Take urgent action to combat climate change and its impacts". Moreover, by exploiting the data centers flexibility, it is possible to reduce the fluctuations and the pressure on the local power grid. This can eventually lead to a more reliable energy distribution in the urban areas, prevent possible power outages, and enhance the intermittent renewable energy integration. All of this contributes to the SDG 7: "Ensure access to affordable, reliable, sustainable and modern energy". Finally, the VEDA framework adheres to the principles of data protection of the GDPR implemented by the EU commission. Indeed, data sovereignty is taken into account in order to protect the users data both at a national and international level.

5 CONCLUSIONS

This paper discussed fundamental research directions to exploit the data center flexibility within electrical energy systems. Challenges and opportunities that lie behind the integration of data centers within the power systems have been highlighted. A conceptualization of a framework called VEDA has been proposed. The fundamental visions behind the VEDA framework, suggest research directions that can positively impact not only the scientific community, but also the society as a whole, by meeting sustainable development goals that are currently key issues at national and international level. Moreover, an interdiciplinary approach is recommended as a vital way to address the issues of modern power systems, tackling both energy and ICT perspectives.

ACKNOWLEDGMENTS

This research is part of ARC - Arctic Center for Sustainable Energy, UiT the Arctic University of Norway grant no 310059.

REFERENCES

[1] M. H. Albadi and E. F El-Saadany. 2007. Demand Response in Electricity Markets: An Overview. In *2007 IEEE Power Engineering Society General Meeting*. IEEE press, Tampa, FL, USA, 1–5. https://doi.org/10.1109/PES.2007.385728

[2] George S. Almasi. and Allan Gottlieb. 1989. *Highly Parallel Computing*. Benjamin-Cummings Publishing Co., Inc., San Francisco, USA. 519 pages.

[3] Yair Amir, Baruch Awerbuch, Amnon Barak, S. Borgstrom, and A. Keren. 2000. An opportunity cost approach for job assignment in a scalable computing cluster. *IEEE Transactions on Parallel and Distributed Systems* 11, 7 (2000), 760–768. https://doi.org/10.1109/71.877834

[4] Chiara Bordin, Anne Håkansson, and Sambeet Mishra. 2020. Smart Energy and power systems modelling: an IoT and Cyber-Physical Systems perspective, in the context of Energy Informatics. *Procedia Computer Science* 176 (2020), 2254–2263. https://doi.org/10.1016/j.procs.2020.09.275

[5] Rajkumar Buyya, Christian Vecchiola, and S. Thamarai Selvi. 2013. *Mastering Cloud Computing* (1 ed.). Morgan Kaufmann publishing., USA. 468 pages. https://doi.org/10.1016/C2012-0-06719-1

[6] Changbing Chen, Bingsheng He, and Xueyan Tang. 2012. Green-aware workload scheduling in geographically distributed data centers. In *4th IEEE International Conference on Cloud Computing Technology and Science Proceedings*. IEEE press, Taipei, Taiwan, 82–89. https://doi.org/10.1109/CloudCom.2012.6427545

[7] Hao Chen, Ayse K. Coskun, and Michael C Caramanis. 2013. Real-time power control of data centers for providing Regulation Service. In *52nd IEEE Conference on Decision and Control*. IEEE press, Firenze, Italy, 4314–4321. https://doi.org/10.1109/CDC.2013.6760553

[8] Igor Griva, S Nash, and Ariela Sofer. 2009. *Linear and Nonlinear Optimization* (2nd. ed.). SIAM publishing, New York, NY. 764 pages.

[9] Md E. Haque, Yuxiong He, Sameh Elnikety, Thu D. Nguyen, Ricardo Bianchini, and Kathryn S. McKinley. 2017. Exploiting Heterogeneity for Tail Latency and Energy Efficiency. In *Proceedings of the 50th Annual IEEE/ACM International Symposium on Microarchitecture* (Cambridge, Massachusetts) *(MICRO-50 '17, Vol. 3)*, Reginald N. Smythe and Alexander Noble (Eds.). Association for Computing Machinery, New York, NY, USA, 625–638. https://doi.org/10.1145/3123939.3123956

[10] Harshad Kasture, Davide B. Bartolini, Nathan Beckmann, and Daniel Sanchez. 2015. Rubik: Fast Analytical Power Management for Latency-Critical Systems. In *Proceedings of the 48th International Symposium on Microarchitecture* (Waikiki, Hawaii) *(MICRO-48, Vol. 3)*, Reginald N. Smythe and Alexander Noble (Eds.). Association for Computing Machinery, New York, NY, USA, 598–610. https://doi.org/10.1145/2830772.2830797

[11] Richard Kavanagh, Django Armstrong, Karim Djemame, Davide Sommacampagna, and Lorenzo Blasi. 2016. Towards an Energy-Aware Cloud Architecture for Smart Grids. In *Economics of Grids, Clouds, Systems, and Services*. Springer International Publishing, Cham, 190–204. https://doi.org/10.1007/978-3-319-43177-2_13

[12] Qilin Li and Mingtian Zhou. 2011. The Survey and Future Evolution of Green Computing. In *2011 IEEE/ACM International Conference on Green Computing and Communications*. IEEE press, Chengdu, China, 230–233. https://doi.org/10.1109/GreenCom.2011.47

[13] Yanpei Liu, Stark C. Draper, and Nam Sung Kim. 2014. SleepScale: Runtime joint speed scaling and sleep states management for power efficient data centers. In *2014 ACM/IEEE 41st International Symposium on Computer Architecture (ISCA)*. IEEE press, Minneapolis, MN, USA, 313–324. https://doi.org/10.1109/ISCA.2014.6853235

[14] David Lo, Liqun Cheng, Rama Govindaraju, Luiz André Barroso, and Christos Kozyrakis. 2014. Towards Energy Proportionality for Large-Scale Latency-Critical Workloads. In *Proceeding of the 41st Annual International Symposium on Computer Architecuture* (Minneapolis, Minnesota, USA) *(ISCA '14)*. IEEE Press, USA, 301–312. https://doi.org/10.1109/ISCA.2014.6853237

[15] David Meisner, Brian T. Gold, and Thomas F. Wenisch. 2009. PowerNap: Eliminating Server Idle Power. *SIGARCH Comput. Archit. News* 37, 1 (March 2009), 205–216. https://doi.org/10.1145/2528521.1508269

[16] Amir-Hamed Mohsenian-Rad and Alberto Leon-Garcia. 2010. Coordination of Cloud Computing and Smart Power Grids. In *2010 First IEEE International Conference on Smart Grid Communications*. IEEE press, Gaithersburg, MD, USA, 368–372. https://doi.org/10.1109/SMARTGRID.2010.5622069

[17] Erik Nygren, Ramesh K. Sitaraman, and Jennifer Sun. 2010. The Akamai Network: A Platform for High-Performance Internet Applications. *SIGOPS Oper. Syst. Rev.* 44, 3 (Aug. 2010), 2–19. https://doi.org/10.1145/1842733.1842736

[18] Asfandyar Qureshi, Rick Weber, Hari Balakrishnan, John Guttag, and Bruce Maggs. 2009. Cutting the Electric Bill for Internet-Scale Systems. In *Proceedings of the ACM SIGCOMM 2009 Conference on Data Communication* (Barcelona, Spain) *(SIGCOMM '09)*. Association for Computing Machinery, New York, NY, USA, 123–134. https://doi.org/10.1145/1592568.1592584

[19] Huigui Rong, Haomin Zhang, Sheng Xiao, Canbing Li, and Chunhua Hu. 2016. Optimizing energy consumption for data centers. *Renewable and Sustainable Energy Reviews* 58 (2016), 674–691. https://doi.org/10.1016/j.rser.2015.12.283

[20] Arman Shehabi, Sarah Josephine Smith, Dale A. Sartor, Richard E. Brown, Magnus Herrlin, Jonathan G. Koomey, Eric R. Masanet, Nathaniel Horner, Inês Lima Azevedo, and William Lintner. 2016. *United States Data Center Energy Usage Report*. Technical Report. Ernest Orlando Lawrence Berkeley National Laboratory.

[21] Cisco Systems. 2021. What Is a Data Center? Retrieved May 11, 2021 from https://www.cisco.com/c/en/us/solutions/data-center-virtualization/what-is-a-data-center.html

[22] Andrew S. Tanenbaum and Maarten van Steen. 2017. *Distributed Systems* (3 ed.). CreateSpace Independent Publishing Platform. 596 pages.

[23] Nguyen Tran, Shaolei Ren, Zhu Han, Sung Man Jang, Seung Il Moon, and Choong Seon Hong. 2014. Demand Response of Data Centers: A Real-time Pricing Game Between Utilities in Smart Grid. In *9th International Workshop on Feedback Computing (Feedback Computing 14)*. USENIX Association, Philadelphia, PA. https://www.usenix.org/conference/feedbackcomputing14/workshop-program/presentation/tran

[24] The United Nations (UN). 2016. Sustainable Development Goals. Retrieved May 15, 2021 from https://sdgs.un.org/

[25] Hao Wang, Jianwei Huang, Xiaojun Lin, and Hamed Mohsenian-Rad. 2014. Exploring Smart Grid and Data Center Interactions for Electric Power Load Balancing. *SIGMETRICS Perform. Eval. Rev.* 41, 3 (Jan. 2014), 89–94. https://doi.org/10.1145/2567529.2567556

[26] Daniel Wong and Murali Annavaram. 2012. KnightShift: Scaling the Energy Proportionality Wall through Server-Level Heterogeneity. In *Proceedings of the 2012 45th Annual IEEE/ACM International Symposium on Microarchitecture* (Vancouver, B.C., CANADA) *(MICRO-45)*, Reginald N. Smythe and Alexander Noble (Eds.). IEEE Computer Society, USA, 119–130. https://doi.org/10.1109/MICRO.2012.20

[27] Muhammad Zakarya. 2018. Energy, performance and cost efficient datacenters: A survey. *Renewable and Sustainable Energy Reviews* 94 (2018), 363–385. https://doi.org/10.1016/j.rser.2018.06.005

[28] Yanwei Zhang, Yefu Wang, and Xiaorui Wang. 2011. GreenWare: Greening Cloud-Scale Data Centers to Maximize the Use of Renewable Energy. In *Middleware 2011*, Fabio Kon and Anne-Marie Kermarrec (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 143–164.