Cloud Infrastructure Design Model for Green Smart City: Case Study of Electricity Generating Authority of Thailand

Ririnda Thirasupa[1,2], Chaiyachet Saivichit[2], and Chaodit Aswakul[2]

[1] Communication System Division, Electricity Generating Authority of Thailand (EGAT)
Nonthaburi, Thailand
[2] Wireless Network and Future Internet Research Unit
Department of Electrical Engineering, Faculty of Engineering, Chulalongkorn University, Bangkok, Thailand
Email: ririnda.t@egat.co.th, {chaiyachet.s, chaodit.a @chula.ac.th}

Abstract. "EGAT Eco Plus" is a smart city initiative of Electricity Generating Authority of Thailand (EGAT). To design the cloud infrastructure to support applications in the smart city, we need a modular design model which can dynamically adjust requirements and system parameters. This paper proposes the concept of smart city cloud controller which aims to provide so-called *green* service resource allocation by maximizing green i.e. renewable energy usage for data centers. Based on our formulated linear programming, our proposed SCCC model can provide green service resource allocation for data centers by maximizing the green energy usage. We evaluate the SCCC formulation with the green energy usage per each data center and overall of green energy usage in optimizing green metrics. Afterward, we compare our model with a well-known algorithm for solving the resource allocation problem i.e. round-robin. The model can be applied in the smart city whose data centers have different accessibility to on-site renewable energy sources. Besides, the proposed model can adjust parameters which are design criteria for green computing; the size of tasks, different power sources; and can be used as design guidelines for green computing in smart cities with more than one power source; grid, on-site renewable energy, battery energy storage.

Keywords: Smart City, Smart Grid, Cloud Infrastructure, Design Model, Green Computing, Renewable Energy.

1 Introduction

"EGAT Eco Plus" is a smart city of Electricity Generating Authority of Thailand (EGAT), which participates in Smart Cities – Clean Energy project [1] [2] under the support of Energy Policy and Planning Office (EPPO), Ministry of Energy (MOE) together with Thai Green Building Organization. The project is one of Thailand Energy

Efficiency Development Plan (EEDP) activities which aim to be a driver for the national development plan to Thailand 4.0 and digital transformation for economic and social aspects.

Thailand smart city project intends to develop cities concerning the quality of life and environmental impact under four major principles as follows. Firstly, the cities responsible to provide, manage and distribute city elements including utilities, infrastructures, and amenities effectively while considering strategies to increase the deployment penetration of renewable energy sources. Secondly, the cities should use flexible and applicable Information Communication Technology to support the growing demand of cities. Thirdly, citizens should be able to access to open technology and information inside cities. Lastly, the cities should create a learning society environment which can lead to sustainable innovative cities in the future. As a state enterprise utility, EGAT uses these design principles as a guideline to design EGAT Eco Plus. And to create a smart city, communication infrastructure backbone network together with cloud infrastructure would play a major role to support all applications and services.

In this paper, we present the cloud infrastructure design model with green concept for EGAT smart city. This can be used as a general model that can dynamically adjust design parameter such as data center capacity, data center infrastructure, tasks, on-site renewable source, power source accessibility for data centers during design period in order to help communication and electrical engineer to approach proper infrastructure implementation.

2 Review of Existing Framework and Related Technologies

2.1 Smart City And Cloud Technology

In the literature, international organizations e.g. ITU-T, IEEE, and IEC have given various definitions of smart cities. ITU-T focus group on smart sustainable cities gives the following definition of smart sustainable city [3], “A smart sustainable city is an innovative city that uses information and communication technologies and other means to improve quality of life, efficiency of urban operation and services, and competitiveness, while ensuring that it meets the needs of present and future generations with respect to economic, social and environmental aspects”. Cloud computing, which is a shared pool of computing resources like networks, server, storage and services [4], is one of the technological challenges for cities which can support heterogeneous and complexities of technologies to provide smart end-to-end services and solutions.

For the upcoming smart power grid trend, a smart city initiative can be abstracted as a microgrid which is a power architecture on the distribution system. Besides, smart cities have capabilities to produce and supply electricity from distributed energy resources (DERs) likes on-site renewable energy sources e.g. wind, solar, mini-hydro, biomass, as well as from fossil fuel e.g. diesel generators. The central power control of microgrid is called microgrid controller whose function is to balance electricity demand and supply inside the city by communicating with other energy management systems (EMS) e.g. building energy management system (BEMS) [5,6], advanced metering infrastructure (AMI), and smart meters. The principle is to receive the power consumption demand and monitor the real-time power production, and to make a decision to implement demand response (DR) measures e.g. by shifting load during peak hours to off-peak periods, by flattering load curves and allowing to consume electricity from less expensive power generations. Due to fluctuating behavior of demand in the microgrid, cloud computing is expected to enable the capability to provide computing resources and storages for various type of DR e.g. direct load control, storage technology, and electric vehicles [7].

2.2 Green Computing

In turn, drastic increases in data center demand service lead to the growth of electricity demand. In regard to cloud service provider, their view is also to maximize the beneficial revenue and deliverable quality of services under constraints on the limited resource. Knapsack problem is used as optimization formulation for cloud computing design [8]. Literature [9] proposes an optimization problem to maximize profit by figuring out a VM-to-PM mapping with maximized energy efficiency. The architecture proposed in [10] is energy-efficient resource allocation mechanisms for green cloud computing but not mention about energy sources. In terms of data centers and local renewable energy resources, GreenSlots [13] is a job scheduler for a data center partially powered by solar energy which aims to reduce brown energy consumption and reduce cost by using solar energy during peak load. Also, for studies [11] and [12] that run across multiple data center with renewable energy. Authors in [11] present the conceptual approach of distributed data center model which allocate workload to distributed data centers where green power is available. Their goal is to maximize green energy usage at the data centers, to forecast workload and to plan for purchases of electricity from wholesale electricity market during an off-peak rate and to reduce carbon emission but not mention about resource allocation mechanism. While authors in [12] propose a workload scheduling for green aware and cooling system.

During design phase of smart city, many system parameters and infrastructure need to be estimated and assumed, for example, task requirements, number of data centers, power source accessibility. Our proposed technique focuses on the mathematical model for cloud infrastructure design with green service resource allocation model, which maximize on-site renewable energy for single or multiple data centers in smart city with different energy infrastructure implementation. The main functions are summarized as below:

- it is a resource allocation strategy with green energy concept which could be applied to support applications in smart city,
- it supports various applications' requirement with different task size and amount,
- it can be adapted to smart city with various type of energy sources such as power grid, on-site renewable energy, energy storage, etc.

3 Cloud Infrastructure Design Model for Green Smart City

3.1 Smart City Cloud Controller (SCCC)

The main objective of SCCC is to be a green service resource allocation model for smart applications that want to use cloud services from two available data centers in EGAT headquarter. Each data center consists of physical machines, virtual machines (VMs). In addition, we can adapt the container [14] technology with the data centers [15]. When demand requests arise, SCCC will calculate the computational power required by the demand requests. Likewise, SCCC will calculate the electrical power consumption needed to run virtual servers if those demand requests would be accepted into the data centers. SCCC sends that electrical power requirement to the energy management system, e.g. microgrid controller or BEMS, to try to purchase or reserve first the electricity from on-site renewable energy sources e.g. solar rooftop systems. However, if the renewable energy sources do not have sufficient power generation to service the demands, then SCCC will try next to reserve the electricity from the conventional power grid. To solve the problem of allocating the demands and corresponding electrical power requirements, in this paper, a linear programming (LP) model is formulated. The intention is for SCCC to achieve the optimal green resource allocation by selecting the computational resource options while concerning the availability of on-site renewable power.

3.2 Linear Programming Formulation For SCCC Design

According to the described strategy, we model the green service resource allocation for SCCC with an optimization problem to find the optimal task assignment proportions. Table 1 summarizes the list of symbols used in the presented model.

The proposed objective function in (1) consists of two terms with linear combination weights. The weights with constraint in (6) are assigned of their proper values to ensure that the model will perform resource allocation as intended with the priority of utilizing the green energy sources first. Particularly, we set the value of w close to 1 (e.g. 0.999 in our reported case study), which renders the objective function maximization dominated by only the first term, representing the amount of electricity power usable from green energy sources. Consequently, the model optimally allocates task *i* to use cloud compute resource from data center *j* and consume electricity from available on-site renewable power sources, $k \in PV$. In case that there are exist multiple solutions with the same value of this first term, the model will then consider the effect of the second term i.e. by further allocating optimally remaining tasks *i* to use cloud compute resource from data center *j* and consume electricity from other power sources, $k \in G$. The objective function (1) and constraints (2) - (6) are expressible as:

Maximize:

$$w \cdot \sum_{k \in PV} \sum_{j \in M} \sum_{i \in N} x_{ijk} R_i + (1-w) \cdot \sum_{k \in G} \sum_{j \in M} \sum_{i \in N} x_{ijk} R_i \qquad (1)$$

subject to:

$$\sum_{k\in L}\sum_{i\in N} x_{ijk} R_i \le C_j \quad ,\forall j \in M \tag{2}$$
$$\sum_{j\in M}\sum_{i\in N} x_{ijk} \gamma R_i \le S_k \quad ,\forall k \in L \tag{3}$$
$$\sum_{k\in L}\sum_{j\in M} x_{ijk} \in [0,1] \quad ,\forall i \in N \tag{4}$$
$$0 \le x_{ijk} \le 1 \quad ,\forall i ,\forall j ,\forall k \tag{5}$$
$$0 \le w \le 1 \tag{6}$$

Table 1. Symbol Definition.

Symbol	Definition
T_i	Task *i* that requests for cloud compute resources from data centers
R_i	Size of resource that task *i* requests from data centers with resource unit abstracted as million instructions per second (MIPS), bandwidth unit, storage capacity units, depending on resource type under consideration
C_j	Capacity of cloud *j* with the same resource unit as task's resource requirement
S_k	Available electricity power from power source *k*. (power unit: PU)
x_{ijk}	Proportion of resource as required by task *i* to use available resource at data center *j* and consume electricity from power source *k*
γ	Linear proportionality constant for increasing relationship between cloud compute resource requirement at data center and electricity power required for operating that cloud compute resource
N	Set of tasks to be allocated
M	Set of data centers in smart city
L	Set of energy sources in smart city
G	Set of energy sources from conventional power grid
PV	Set of on-site renewable energy sources e.g. solar rooftop
w	Linear weight constant

Constraints (2) and (3) are responsible for resource allocation in data centers and energy allocation in smart city, respectively. Constraint (2) ensures that the total resources of all the tasks that use cloud compute resource from each data center *j* must not exceed the resource capacity of that data center *j*. And (3) guarantees that the total power requirement by each data center *j* must not exceed the available capacity of each power source *k*.

For task assignment proportion x_{ijk}, constraints (4) and (5) ensure that total resources required by a task are splittable and assigned to multiple data centers. This assumption is easily justifiable in case of storage resources. For compute resource, the assumption can also be justified with the containerization framework. That is, a task e.g. requiring a big data analytics can be composed by relying on multiple small containers that can run simultaneously. The containers can be split and sent to run at multiple data centers, in case that all containers cannot be accommodated by either data center alone. Note also from (4) and (5) that some tasks can be rejected for accessing the data centers when their cloud resources or available supplying energy sources run out of capacities. In practice, if this happens, then alerts would be raised to responsible engineers and planning for upgraded capacities would be in place. The formulated optimization can also be applied to help engineers understand the effect of capacity upgrade scenarios.

4 Experimental Settings

In order to study the performance and effectiveness of the proposed SCCC model, we have conducted three different implementation scenarios. Moreover, in each scenario, two case studies are set up to evaluate the performance of the proposed model. We formulate and solve the optimization problem in MATLAB R2016a.

In this section, we set up three test scenarios with different accessibility levels of renewable energy sources from each data center, as shown in Figs. 1. We assume task T_1, T_2, T_3, and T_4 have the size of required resource R_1, R_2, R_3, and R_4 and request to use cloud compute resources from data centers 1 and 2 with cloud capacity C_1 and C_2 respectively. Data centers 1 and 2 can request to purchase electricity from power sources 1 and 2 with available electricity power of S_1 and S_2 respectively. Moreover, S_2 presents available electricity from grid and S_2 presents available electricity from on-site renewable energy; based on solar rooftop.

Test scenario (a), all data centers have full access to green energy and grid: Both data centers can consume electricity from both power sources.

Test scenario (b), only some data center has full access to green energy and grid: Data center 1 can consume only electricity from power source 1, and data center 2 can consume electricity from both power sources.

Test scenario (c), each data center has access to only green energy or grid: Data center 1 can consume only electricity from power sources 1, likewise, data center 2 can consume only electricity from power sources 2.

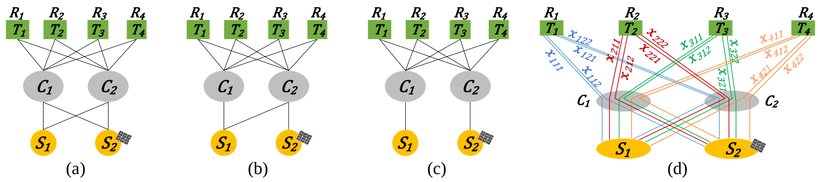

Fig. 1. Preliminary test scenarios (a), (b), and (c) respectively. (d) Task assignment proportion x_{ijk}.

For interpretation convenience, all tasks are considered of the same service type with the same size of resource requirement. Recall that all tasks have no mutual dependencies and resources required by each task can be divided and assigned for uses at multiple data centers. Therefore, the task resource assignment proportion x_{ijk} can possibly be floating numbers. Without loss of generality, γ is normalized to 1 with the proper unit to convert to the required PU per each resource unit of different types.

Consider two case studies: case study 1 to study the impact of increasing the size of the task resources, case study 2 to study the impact of increasing available electricity power capacity from power source 2 i.e. renewable power source. In practice, power source 2 is intermittent energy resource type, thus available electricity from the power source 1 always is much more than that available from power source 2. Table 2 summarizes our parameter settings.

Table 2. Parameter Settings for Case Studies.

Parameter Setup	Case Study 1	Case Study 2
Size of task resource requirement R_1, R_2, R_3, and R_4	Varying from 5 up to 25	10
Data center capacity C_1 and C_2	30	30
Available electricity from power source 1 (grid)	1000	1000
Available electricity from power source 2 (on-site renewable energy)	20	Varying from 0 up to 200

5 Experimental Results and Discussions

In this section, we present the experimental results for SCCC optimization from two case studies in test scenarios (a), (b), and (c). Fig. 2 and Table 3 give the results of case study 1; Fig. 3 and Table 4 give the results of case study 2.

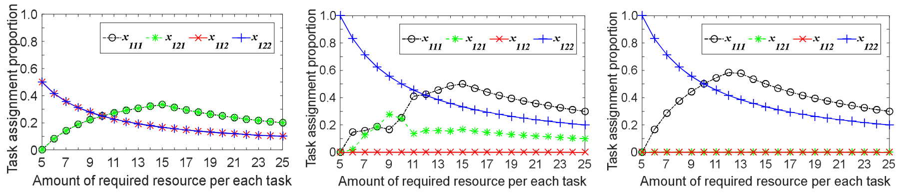

Fig. 2. Task assignment proportion x_{ijk} in case study 1.

Table 3. Case study 1: impact of increasing size of task resources.

Test scenario	Size of task resource requirement	Total size of resources from all tasks	Data Center Utilization (%)		Green Energy Usage				Overall ratio of green energy usage (%)
					Data Center 1		Data Center 2		
			Data Center 1	Data Center 2	Power unit (PU)	Percentage of PU compared to total electricity uses in data center	Power unit (PU)	Percentage of PU compared to total electricity uses in data center	
(a)	5	20	33.3	33.3	10	100	10	100	100
	15	**60**	**100**	**100**	**10**	**33.3**	**10**	**33.33**	**33.33**
	25	100	100	100	10	33.3	10	33.33	33.33
(b)	5	20	0	66.66	0	0	20	100	100
	15	**60**	**100**	**100**	**0**	**0**	**20**	**66.66**	**33.33**
	25	100	100	100	0	0	20	66.66	33.33
(c)	5	20	0	66.66	0	0	20	100	100
	12.5	**50**	**100**	**66.66**	**0**	**0**	**20**	**100**	**40**
	25	100	100	66.66	0	0	20	100	40

From Fig. 2 and Table 3, the result of test scenario (a) shows that, at the size of each task equal to 5 with a total size of 20, both data centers 1 and 2 having the total cloud capacity equal to 60 can accommodate all tasks by allocating to use cloud capacity C_1 and C_2 equally. In additions, for green energy usage, both data centers consume electricity from power source 2 with green energy usage being 100% of total electricity consumption by data centers 1 and 2. The result of test scenario (b) shows that our model allocates tasks to use cloud capacity C_2 at first because only data center 2 can access to green energy. Hence, in the beginning, only data center 2 has 66.7% utilization with green energy usage being 100% of total electricity consumption by data center 2. After no available of green energy for data center 2, the model allocates tasks to use cloud capacity C_1 and C_2 by consuming excessive electricity from power source 1. Lastly, the result of test scenario (c) shows that at the size of each task equal to 5 with a total size of 20, all tasks are assigned to use cloud capacity C_2 at first because only data center 2 can access to green energy. Then, in the beginning, only data center 2 has 66.7% utilization with green energy usage 100% of total electricity consumption by data center 2. However, while increasing the size of each task, the model increases task assignment proportion to only cloud capacity C_1 which consumes electricity from power source 1 without allocating more tasks to use cloud capacity C_2 even there is still available resource at data center 2. The reason is there is no more available electricity for data center 2 to execute more tasks; therefore, total cloud capacity reduces to 50.

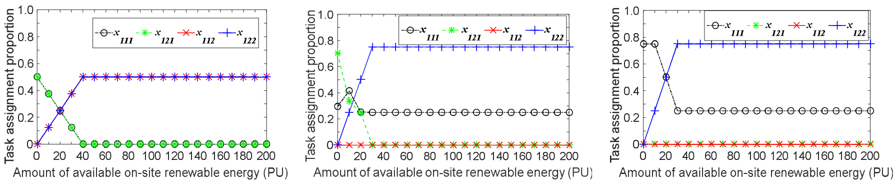

Fig. 3. Task assignment proportion x_{ijk} in case study 2.

Table 4. Case study 2: impact of increasing available electricity power capacity from renewable power source.

Test scenario	Available electricity from power source S_2	Data Center Utilization (%)		Green Energy Usage				Overall ratio of green energy usage (%)
				Data Center 1		Data Center 2		
		Data Center 1	Data Center 2	Power unit (PU)	Percentage of PU compared to total electricity uses in data center	Power unit (PU)	Percentage of PU compared to total electricity uses in data center	
(a)	0	66.66	66.66	0	0	0	0	0
	30	66.66	66.66	15	75	15	75	75
	≥ 40	**66.66**	**66.66**	**20**	**100**	**20**	**100**	**100**
(b)	0	40	93.9	0	0	0	0	0
	30	**33.33**	**100**	**0**	**0**	**30**	**100**	**75**
	≥ 40	33.33	100	0	0	30	100	75
(c)	0	100	0	0	0	0	0	0
	30	**33.33**	**100**	**0**	**0**	**30**	**100**	**75**
	≥ 40	33.3	100	0	0	30	100	75

In case study 2, the results of all test scenarios show that when increasing the available electricity from S_2, the model can assign tasks to use compute resource from both data centers while maximizing the green energy usage. For that reason, refer to parameter setting in Table 2, we compare the results at the total required resource from all tasks equal to 40. In test scenario (a), both data centers can provide green computing service to all tasks with 40 PU of green energy usage in total. While in test scenarios (b) and (c), only data center 2 can provide green computing service with 30 PU of green energy usage in total. Moreover, in terms of data center utilization for test scenarios (b) and (c), data center 2 always has higher utilization than data center 1.

Based on the preliminary test results of two case studies; impact of increasing size of task resources, and impact of increasing available electricity power capacity from renewable power source; with different test scenarios, we can conclude as follows. For both case studies, the results show that the model can adjust parameter; size of task resource and available electricity power capacity from renewable power source; and then maximize green task allocation.

Afterward, we compare our model with a well-known algorithm for solving the resource allocation problem i.e. round-robin. Fig.4 presents how the two strategies perform green service resource allocation when increasing the available electricity power from a renewable power source. We can observe that the SCCC can perform better optimize green service allocation than round-robin in every test scenarios. This happens because round-robin tries to evenly assign tasks to use resources in equal portions without concerning types of power sources. In some circumstance that resource of a data center is not available due to power shortage, round-robin will consider remaining available resource from another data center to handle equal task allocation. Conversely, the SCCC elastically allocates tasks to the data center, which can access to green energy since it is available renewable energy until reaching the maximum required electricity power for data centers to perform the optimal green service resource allocation.

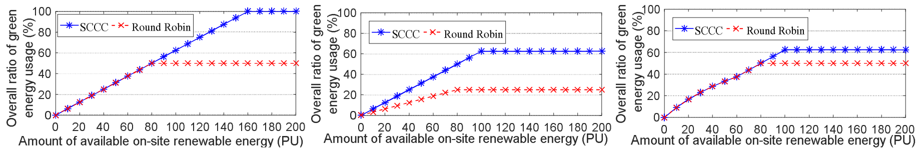

Fig. 4. Comparison of SCCC against round-robin on overall ratio of green energy usage.

In different energy source accessibility scenarios, in case all or only some data centers have full access to green energy and grid, the results show that the model can maximize task allocation to use compute resource from both data centers. However, only test scenario (a) that both data centers can provide green computing. In contrast, if each data center has access to only green energy or grid, the results show that the available cloud capacity C_2 is limited by the available electricity from only power source S_2. For these results it can imply that to design data centers to achieve green computing, we need to concern the accessibility and availability of green energy of each data center.

In a real implementation, scenario (a) and (b) can be seen in any smart city. In fact, scenario (a), all data centers have full access to green energy and grid is for smart cities which DERs are pooled and can supply electricity to any buildings or areas which data center located. Similarly, scenario (b), only some data center has full access to green energy and grid is for smart cities which DERs can supply to some buildings or areas which data center located. In such a case, the cloud provider in a smart city can design and manage data centers to provide green computing and achieve carbon credit target.

6 Conclusion

We propose SCCC model which can provide green service resource allocation for data centers by maximizing the green energy usage. Moreover, the model can be applied to use in the smart city whose data centers have different accessibility to on-site renewable energy sources. It can be a guideline for cloud and energy infrastructure planning to achieve green computing. Besides, the proposed model can adjust parameters which are design criteria for green computing; the size of tasks, different power sources; and can be used as design guidelines for green computing in smart cities with more than one power source; grid, on-site renewable energy, energy storage.

To improve our model, we will develop and abstract more input parameters; type of tasks, compute requirement of each task type, network bandwidth; and also the concept of select on-site renewable energy sources while concerning power loss in cable due to distance between generation and load.

Acknowledgement

The first author is grateful to Electricity Generating Authority of Thailand (EGAT) to support the Master degree research scholarship.

References

1. Ministry of Energy: Thailand Smart Grid Master Plan 2015-2036, (2015).
2. Energy Policy and Planning Office, Ministry of Energy: หลักเกณฑ์การประเมินการออกแบบเมืองอัจฉริยะ [Criteria for Smart Cities-Clean Energy], (2016).
3. International Telecommunication Union, ITU-T Focus Group on Smart Sustainable Cities: An overview of smart sustainable cities and the role of information and communication technologies. ITU-T/FG-SSC Technical Report, (2014).
4. Mell, P. and Grance, T.: The NIST Definition of Cloud Computing - Recommendations of the National Institute of Standards and Technology. National Institute of Standards and Technology, (2011).
5. Inthasut, T. and Aswakul, C.: Development and Reliability Testing of IEEE1888 Gateway for ZigBee Wireless Sensor Network in Chulalongkorn University's Building Energy Management System (Excellent Paper Award). In: Proceedings of 8th International collaboration Symposium on Information, Production and Systems, ISIPS 2014. Fukuoka (2014).
6. Le, D. H. and Pora, W.: Development of smart meter for building energy management system based on the IEEE 1888 standard with Wi-Fi communication. In: Proceedings of ICEIC 2014. Kota Kinabalu (2014).
7. Cao, Z., Lin, J., Wan, C., Song, Y., Zhang, Y. and Wang, X.: Optimal Cloud Computing Resource Allocation for Demand Side Management. In: IEEE Transactions on Smart Grid, vol. 8, pp. 1943-1955. IEEE, (2016).
8. Banerjee, A., Ghosh, S. C., and Banerjee, N.: Pack Your Sack for the Cloud. In: Proceedings of ISEC '12, pp. 157-160. ACM, India (2012).
9. Zhang, X., Wu, T., Chen, M., Wei, T., Zhou, J., Hu, S., and Buyya, R.: Energy-aware virtual machine allocation for cloud with resource reservation. The Journal of Systems and Software 147, 147-161 (2019).
10. Buyya, R., Beloglazov, A. and Abawajy, J.: Energy-Efficient Management of Data Center Resources for Cloud Computing: A Vision, Architectural Elements, and Open Challenges. In: Proceedings of the 2010 International Conference on Parallel and Distributed Processing Techniques and Applications (PDPTA 2010). (2010).
11. Bird, S., Achuthan, A., Maatallah, O.A., Hu, W., Janoyan, K., Kwasinski, A., Matthews, J., Mayhew, D., Owen, J., and Marzocca, P.: Distributed (Green) Data Centers: A New Concept For Energy, Computing, And Telecommunications. Energy for Sustainable Development 19, 83–91 (2014).
12. Chen, C., He, B. and Tang, X.: Green-Aware Workload Scheduling in Geographically Distributed Data Centers. In: 2012 IEEE 4th International Conference on Cloud Computing Technology and Science. IEEE, (2012).
13. Goiri, Í., Haque, M. E., Le, K., Beauchea, R., Nguyen, T. D., Guitart, J. et al.: Matching renewable energy supply and demand in green datacenters. Ad Hoc Networks 25, 520-534 (2015).
14. Roy, G., Washbrook, A., Crooks, D., Qin, G., Skipsey, S. C., Stewart, G. And Britton, D.: Evaluation of containers as a virtualisation alternative for HEP workloads. Journal of Physics: Conference Series 664 022034, 1-8 (2015).
15. Souppaya, M., Morello, J., and Scarfone, K.: Application container security guide. NIST Special Publication 800-190, (2017).