Greening Data Centers- Holistic Methodology

Anannya Patra, Anirudh B, Arvindh S K, Gautham M K, Irene Komal P, Lalitashree Hedge, Dr. Sheela S V
Department of Information Science and Engineering
BMS College of Engineering
Bangalore, India

Abstract— **With the advancement of technology and mobile computing, the demand for highspeed data and computationally intensive applications has increased which has in turn led to the creation of complex hardware architectures that include high-capacity servers, storage systems, routers, switches, load balancers, firewalls and more. Data Centers are physical facilities that organisations use to house components that help power their critical applications and data. While trying to provide a secure and accelerated delivery environment, these components and their associated cooling systems end up consuming a huge amount of electricity which in turn rises the carbon-footprint of those appliances that eventually contribute to lobal warming. The paper reviews and proposes a few simple solutions that could help limit the energy consumption of data Centers without affecting efficiency. We are providing solutions that 'green-ify' data Centers as well as some challenges in the realization of these greening goals.**

Keywords—Green data-Center; Power usage effectiveness; Smart cities; Green data-Center network; Green data-Center cooling system

I. INTRODUCTION

Data Centers are facilities sheltering many resources for computing, networking and storage. These data Centers assist in solving large scale problems in various fields. Bigger problems that cannot be solved by a single data Center are solved by a group of distributed data Centers working together. Data Centers are the primary enabling architecture for clouds and grids.

In the discharge of these functions, data Centers perform computations and operations that, naturally, demand a significant amount of energy. This amount of energy consumption usually creates a big impact on the environment in terms of CO2 emissions. In fact, Data Centers are said to consume about 2% of global electricity and form 0.3% of the global carbon emissions. This energy is consumed both by the IT components as well as cooling facilities that support and enable the functioning of these IT components. It is imperative for us to minimize the energy consumption and help the environment in the process.

In lieu of the exceeding impact of data Centers on our environment, the paper proposes a holistic approach to make these data Centers greener and more energy efficient. Renewable energy is, by nature, intermittent and has to be coupled with several methodologies to promote energy efficiency while also keeping in mind the ideal performance and reliability of the data Center.

Some of these methodologies include virtualization and consolidation of virtual machines, workload distribution and optimal scheduling. We also review the approaches to green other aspects of data Centers including networks, cooling facilities and other architectural and infrastructural components.

An objective way to understand the efficacy of our methodologies is presented in the different metrics used to measure them. The paper looks to closely compare the different methodologies and their Power-Utilization-Effectiveness scores (PUEs) amongst other things. This helps us more closely understand the standings of these implementations and also the advantages and disadvantages they pose.

II. RELATED WORK

Bilal et al. focuses on the Data Center network architectures and the problems that arise with it. Due to the increase of the usage of fossil fuels and increase in greenhouse gases emission because of the networking devices, there is a need for energy efficient data Center networking devices. The current network techniques are categorized into consolidation, proportional computing and selective correctness. According to this paper, the legacy data Centers network architectures suffer from inefficiency, high cost, poor agility, low scalability. Based on the network traffic routing model, data Centers network architectures are classified into hybrid model, service centric model and switch-centric model and the three models are compared, and their challenges are discussed. The deduction proves that there is no one size fits all model, and the organization must see for itself which fits his requirements. [1]

Much research has highlighted the increasing trend of energy consumption in DC originating from the IT systems such as server, storage, network and from the non-IT system such as cooling system and power system.
Bahari et al. talks about the issues and challenges on the IT systems and tells ways to assist the data Center operators to help them move towards green data Centers and ways to adopt the Green Data Center. There is a need to measure the Data Center efficiency level towards the Green data Center by measuring the amount of work accomplished or information stored in each footprint and the energy consumed. The areas taken into consideration are-Reliability

and Performance, Resource Management, Energy management. Various solutions are provided like installing the systems in a way that the under-utilized resources are properly utilized, proper planning is done by observing the customer's SLA carefully, monitoring the utilization of resources and finding Power Usage Efficiency and adopting the green computing with energy management which aims in increasing the performance, reliability and energy efficiency of the data Centers. [2]

Liu et al. stresses on the complexity of Data Center architecture and takes an IoT approach to green data Centers. Here, an air-conditioning system assisted by cloud technology and IoT is proposed. The air-conditioning system proposed looks at monitoring, management and control of refrigeration devices. The paper also looks at an intelligent temperature control system, using statistics and machine learning. The entire air-conditioning system is managed using a Cloud Management Platform that handles storage, analysis and prediction. A sensor network is also used that considers the temperature and humidity. This is all fully performed with the ZigBee network architecture. Results for experimentation proved to be promising - with the network of smart air-conditioning reducing total power consumption by up to 10% in some conditions. [3]

Baccour et al. presents energy-related problems in data Centers and reviews the state of the art of the research literature on energy efficient architectures and techniques, the challenges facing each approach and the strategies to build a green data Center. According to the paper, the different obstacles that prevent a network from becoming green are the use of brown energy (fossil-based fuel), inefficient management of cooling infrastructure, power hungry equipment and architectures and under-utilized networks. To overcome these obstacles, the solutions proposed were Renewable sources of energy, Power-aware cooling, Energy efficient equipment, Energy efficient interconnections, Power-aware routing algorithms and energy efficient software. [4]

Nor et al. makes a comparison of five existing energy efficient data Center frameworks- Data Center Energy Efficient Frame work(DCEEF), Guidelines for Energy-Efficient Data Centers, The 4 Pillar Framework, EU Code of Conduct on Data Centers, Best Practices Guide for Energy Efficient Data Center Design. The deduction proves that the Green Data Center Alliance (GDCA) has the most complete framework which covers facility design and engineering, information technology (IT), process, governance and finance. The paper also provides guidelines on choosing a green data Center framework considering cost, architecture and ROI. [5]

Jin et al. talks about how applications running on multiple servers can be combined to work on a single server through virtualization and its impacts. Two leading virtualization models are taken into consideration along with the physical machine while measuring the impact, they are Xen and KVM where the physical server is virtualized into multiple virtual machines. The hypervisor used, the virtualized I/O mechanisms, virtualized CPU model and the virtualized network model are mainly taken into consideration. The setup included three servers in which one acted as a physical server and the VMs installed in the other two. The background energy consumption measurement found that all the servers consumed the same power when turned off and plugged into different power supplies. However when turned on, the KVM had a power overhead, but consumed less power than the physical server. It was concluded that without proper design and planning the energy overhead on the virtualized servers will be more. [6]

Toress et. al, presents that after consolidating the servers and the environment can lead to energy and cost saving. Along with virtualizing the servers, recycling through resource transformation using the memory compression technique is being proposed. Also, the resources being given that will go waste are being identified and are reduced. Here an idea of modern management middleware is introduced. Workload description and adding tailored resources have been proposed. It is being concluded that the combined use of memory compression and request discrimination can lead to saving of energy. [7]

Cavdar et. al talks about the green metrics with focus on computing and networking. Dynamic voltage frequency scaling (optimize use of CPUs) and dynamic power management (number of servers that should be active) are the two approaches proposed for energy consumption. The objective in this paper was to reduce the average response time and average power consumption. This paper focused on data intensive jobs which require low computational load but produce heavy data streams directed out of the data Center. Therefore, the main objective was to avoid congestion while minimizing the number of computing servers. The proposed DENS scheduler was evaluated in a GreenCloud simulation environment. Other approaches included cooling and nano data Centers. [8]

Fawaz et. al reviews the Power Usage Metrics - Power Usage Effectiveness and Green Power Usage Effectiveness and draws up a comparison to determine which was the superior metric. Additionally, the paper also analyses some green determinants of data Centers. There is also a demonstration of the impact of the Parasol data Center. When the Power Usage Efficiency and Green Power Usage Efficiency are compared, the latter provided results with more accuracy as it also took into consideration the type of power used. The paper concludes that Green Power Usage Efficiency was a more appropriate metric to measure CO_2 emissions and the overall carbon footprint. [9]

Kiani et. al identifies the drawbacks of existing practices to make data Centers green and addresses the same in the following ways- Workload distribution Center: a set of servers dedicated to balance out the workload to provide more flexibility and Green versus Brown: Separate green energy maximization problems and brown energy minimization problems. [10]

Iturriaga et. al identifies the problem of scheduling high profile computing workflows and attempts to solve the same by dividing it into smaller scheduling subproblems and solving them one by one. Dividing high level scheduling problems in federated data Centers into smaller problems by considering individual data Centers. Lower-level scheduling problem in green powered data Centers solved by an algorithm that schedules tasks of the workflows assigned to each data Center. [11]

Chen et. al presents an algorithm - MinBrown that focuses on optimizing the consumption of energy by favoring consumption of green energy over brown energy. The algorithm is written to run on data Centers which are widely geographically distributed and deal with load with considerable slack - that is, load that is not interactive. It takes into consideration the relationship between green energy and the physical environment including the need for increase in cooling in hotter climates. The algorithm takes into consideration parameters such as available green energy, data Center capacity, network latency and bandwidth, cooling power and equipment and workload structures, computational requirements and deadlines. The study reports the findings of an experiment conducted by simulating distributed data Centers with old workload traces as well as real world traces of solar energy. The static and runtime optimizations in the algorithm together resulted in a 40% reduction in brown energy consumption compared to traditional methods. [12]

A major problem with using renewable energy sources to power data Centers is the unreliable and weather dependent availability of energy. Aksanli et. al addresses these constraints on data Centers by designing a job scheduler that attempts to predict the energy availability - this involves specifying methods to predict both solar and wind energies in a given period of time. Specifically, for wind energy, the paper proposes an estimation mechanism that adapts to seasonal changes and relies on the wind speed and direction. The approach aims to reduce the completion time and increase the throughput of batch jobs while meeting the response time constraints of service jobs by ensuring the quick energy prediction and scheduling. The algorithm was tested on a self-developed 'discrete event-based simulation platform' that scheduled a mix of service and batch jobs for 4.5 days. The methodology led to three times better green energy usage and reduced the number of terminated tasks by almost 8 times while also ensuring that response time for service tasks stayed below the 90th percentile. [13]

Janacek et. al aims to integrate a data Center into an existing smart grid in order to harness the benefits of smart grids including the infrastructure and renewable energy while also acting as a power balancer to the smart grid by cutting power peaks among other things. It also focuses on repurposing the heat generated by the data Centers into generating energy that benefits other consumers. The work is verified through simulations of both the smart grid and the data Center where the data Center acts as an abstracted power consumer of the grid and the interface between the two contains communication protocols to exchange the amount of renewable energy, power storage capacity available as well as the waste heat supply and demand. The Surrogate Data Center Model is created to represent this data Center powered by a smart grid and all the internal power relationships between the two. The model can adjust to power consumption and adapt both the components as well as the entire architecture to optimize the power consumption. Further, the temporal behavior of the data Center is modeled in order to orchestrate tasks to generate the required waste heat. [14]

Weerts et. al talks specifically about cooling and the methodologies and metrics to implement and measure it. The two methodologies mentioned in the paper are- Airside economization and Indirect Evaporative Cooling. The metric used to measure the effectiveness of these methods is Power Utilization Effectiveness (PUE) as covered in [9]. The results of the paper are promising but restricted to the region of case study - as the results are weather dependent and can change quite dynamically with region. [15]

Kim et. al talks specifically about 'free cooling' - the ability to utilize outside air and temperatures to cool data Centers. The paper also mentions the challenges - free cooling can only occur in certain geographical locations and weather is still unreliable. Therefore, in many cases 'hybrid cooling' is used - where free cooling is considered the primary cooling method, but an electric cooling system (for example, air conditioners) is used as a backup. Apart from this, the paper talks about 'VM consolidation' which is the migration of full virtual machines from one server or node to another to better utilize resources. VM consolidation also helps in cooling. The paper looks for an optimization problem to best solve these two factors: The cooling mode (free or electrical) and the VM consolidation placement - where to migrate all the VMs. The results of the paper are also quite promising - stating a 25.7% increase compared to regular free-cooling. [16]

Capozzoli et. al talks about the current methodologies in cooling data Centers and emerging technologies. This includes the broad classification into air-cooled and liquid-cooled systems. Further, the paper talks about 'economizer modes' which investigate 'free cooling' through utilization of the outside air and temperature. The future technologies mentioned in the paper include the possibilities of directly liquid cooled servers and something called the 'micro-channel two-waste flow system' of cooling. Along with this, the paper looks briefly into two other aspects of energy optimization - these include waste heat recovery, which re-utilizes the heat from IT work in other ways, and usage of RES or renewable energy sources for generating the electricity in the first place. The paper ends providing a brief analysis and overview of these different methods. [17]

All the work mentioned above has been consolidated in a table format for clearer understanding.

TABLE I. CONSOLIDATION OF THE PREVIOUS WORK

Paper	**Method/ Algorithm**	**Architecture/ Framework**	**Energy efficiency**	**Cooling**
[1]		Hybrid model, service centric model, switch centric model- Network architectures		
[2]	Virtualization		Check the performance of resources and prevent it from degrading.	
[3]		IOT Based architecture		
[4]	Power aware routing algorithm		Energy efficient equipment	Power-aware cooling
[5]		Comparison of frameworks - Data Center Energy Efficient Framework (DCEEF), Guidelines for Energy-Efficient Data Centers, The 4 Pillar Framework, EU Code of Conduct on Data Centers, Best Practices Guide for Energy Efficient Data Center Design		
[6]	Virtualization			
[7]	Memory compression, Modern management middleware			
[8]			Dynamic voltage frequency scaling, dynamic power management	
[9]	Power usage effectiveness and green power usage effectiveness			
10]	Workload distribution Centers			
[11]	Lower-level scheduling is done by algorithms that schedule task of the workflows assigned to each data Center.			
[12]	MinBrown algorithm			Hotter geographical areas that need more cooling penalized.
Paper	**Method/ Algorithm**	**Architecture/ Framework**	**Energy efficiency**	**Cooling**
[13]	Job scheduler to predict green energy availability			
[14]	Integrate data Centers with smart grids		Repurposes heat generated by data Center to produce energy.	
[15]				

				Airside economization and indirect evaporative cooling
[16]		Power usage effectiveness		'Free cooling' - utilize outside air, temperature to cool data Centers
[17]				Liquid cooled servers and the 'micro-channel two-waste flow system' of cooling.

III. PROPOSED METHODOLOGY

The first step for a green data Center is proper planning and then executing the plan according to the requirements. When an organization takes a step towards becoming green, it gets recognition as being considerate towards environment and customers nowadays prefer buying appliances with less carbon footprint. We have listed below some solutions to decrease the overall power consumption and carbon footprint of a particular data Center.

A. Limiting the Usage of Brown Energy and Switching to Renewable Energy

A large amount of data Centers uses brown energy which is essentially, energy from burning coal and other fossil fuels. This results in the depletion of these exhaustible resources and additionally increases the carbon footprint. Shifting to greener energy sources like solar and wind energy as an alternative for brown energy can significantly decrease the carbon footprint and increase efficiency. Since these are inexhaustible resources, this approach proves to be sustainable. In case these forms of energy are not favorable due to the weather or other environmental factors, hybrid systems can be developed with efficient algorithms to alternate between green energy and brown energy. There are several other scheduling specific algorithms that ensure that unreliable weather conditions don't necessarily result in nonuse of green computing methods. Application of weather and location aware algorithms that account for non-availability of green energy in certain circumstances have been proven to significantly help improve the usage of green energy. This is done by strategically scheduling time insensitive batch jobs in locations that support green energy or during times of day that have better access to green energy. Even though the initial setup cost is higher, they promise green power for the long run.

B. Energy Efficient Servers to Decrease Power Consumption

Another approach of improving energy efficiency in Green data Centers is to use energy efficient servers like Eco Blade Server, which reduces power consumption and maintains such a CPU throttle speed that decreases the rate of processer's voltage and clock. The Power Usage Effectiveness (PUE) of green data Centers while using energy efficient servers falls to approximately 1.2 and lower Power Usage Efficiency refers to high energy efficiency.

C. Energy Efficient Cooling

Since data Centers produce a lot of heat, a large amount of energy is consumed by their cooling and heat dissipation infrastructure. Cooling systems should be modular with smart monitoring capabilities to dissipate heat from those components that are always exceeding heat thresholds instead of all components. The advancements in cooling technology should be embraced. As an example, Sea Water Air Conditioning (SWAC) systems can be leveraged to efficiently cool data Centers while lowering carbon emissions by nearly 80-90% and dramatically cutting down long term costs. This system runs 24 hours a day, 7 days a week, 365 days a year and as an added benefit, the heat that is wicked away by the sea water can be used to heat local homes and offices. Another experiment by Microsoft's Project Natick team showed that data Centers deployed underwater could be reliable and sustainable. Another method would be to use Cooling as a Service.

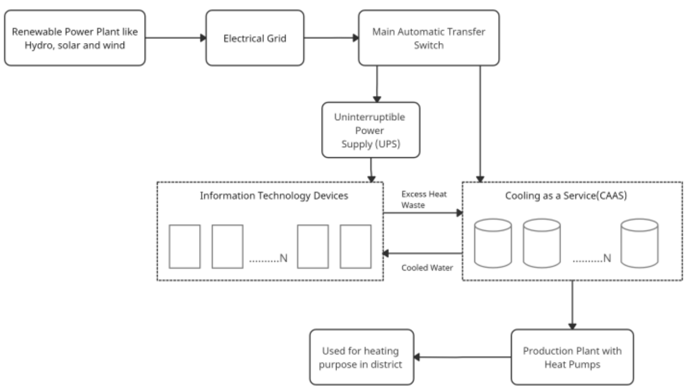

Fig.1. Green Data-Center Architecture

Figure 1 is a prototype of the architecture of a green data Center where we get our electricity from renewable sources like hydro, solar or wind power. This electricity is then transferred to all consumers using power grid and this electricity is used to power up data Centers. On switching on

IT devices, the heat waves exiting from the IT devices are sent to the Cooling-As-A-Service Heat Transfer. It has been known that chilled water is more beneficial and effective for purposes of cooling than air. So, when hot air waves are received from the IT devices, the Cooling-As-A-Service Heat Transfer sends in chilled water. An extra step to be more environmentally sustainable is to send the excess heat to a production plant with heat pumps and this can be supplied to the city for the citizen's use, in geysers for water heating.

On implementing this model, there will be many areas where green house gases emissions can be reduced. Firstly, for the cooling of IT Devices, when we move from traditional air conditioners to the Heat Transfer mechanism, the cooling is more effective with less Power Usage Effectiveness. Secondly, every household requires hot water, instead of using the electricity to power up the geysers, if we take the waste air out of the Cooling-As-A-Service Heat Transfer, make it go through proper treatment and then use for residential heating, it will save power consumption and in turn help the whole state to be more environmentally sustainable.

D. Network Architecture and it's Effects

It should also be noted that data Centers can only become fully green if the underlying devices that they are composed of become green. Most network switches and routers, to guarantee performance, consume large amounts of power and add to the energy woes of a data Center. In this regard, the use of commodity network devices, wireless technology and optical network connections are the way to go.

E. Power Aware Routing

Routing within a network shows potential for efficiency. Implementing power aware algorithms like ElasticTree can help in reducing the number of active devices and links within a network without compromising on performance. The resources that are not being used can then be shut down to reduce power consumption. Network resources can consume a significant amount of power (nearly 70% of peak power) when they are idle or not running at their full potential. In case the network is underutilized, virtualization can be used to run multiple virtual machines on the same physical machine to serve data more efficiently.

IV. CONCLUSION

Data Centers are the driving force behind the IT world and without the IT industry the modern world will not survive. As these data Centers gain more importance the energy they are consuming and the impact they have on the environment cannot be ignored. Hence the green data Centers are not just a forward-thinking approach but has become a necessity. Many enterprises have already opted the alternate energy solutions for their data Centers and taken a step towards making the environment pollution-free. As discussed in this paper, many solutions have been proposed to make the data Centers greener which contribute towards reducing the carbon footprint and increase the efficiency of data Centers. Similarly, an effort is being made to propose a solution in this paper to the same cause of implementing environment-friendly and energy saving practices. The most ideal value for Power Usage Efficiency is 1 and with all these solutions, green data Center services aim at implementing the most feasible solution and achieve this Power Usage Efficiency. One positive point to be considered here is that with all these efforts made to save energy, the data Centers will not lack in performance and energy consumption will also keep the cost low. To conclude, it can be said that with people around the becoming concerned about the environmental impact data Centers are having, green data Center services have made sure to resolve all these concerns without affecting the performance or efficiency of an enterprise or organization.

V. FUTURE WORK

Future work related to this paper includes practical implementation of the solution proposed in this paper to validate how each solution helps in saving energy and making the data Centers green. Working with real time data Centers will help in analyzing the process of implementing all the solutions mentioned and analyzing the outcome of each of them to determine which can be the best possible way to enhance energy saving in data Centers. The categories on which we ca base our judgement would be the Power Consumption, Power Usage Effectiveness, Energy Efficiency and Carbon Footprint of the total data center as a whole.

REFERENCES

[1] Bilal, Kashif, Samee U. Khan, and Albert Y. Zomaya. "Green data Center networks: challenges and opportunities." In 2013 11th International Conference on Frontiers of Information Technology, pp. 229-234. IEEE, 2013.

[2] Bahari, Hazril Izan, and Siti Salbiah Mohamed Shariff. "Review on data Center issues and challenges: Towards the Green Data Center." In 2016 6th IEEE International Conference on Control System, Computing and Engineering (ICCSCE), pp. 129-134. IEEE, 2016.

[3] Liu, Qiang, Yujun Ma, Musaed Alhussein, Yin Zhang, and Limei Peng. "Green data Center with IoT sensing and cloud-assisted smart temperature control system." Computer Networks 101 (2016): 104-112.

[4] Baccour, Emna, Sebti Foufou, Ridha Hamila, and Aiman Erbad. "Green data Center networks: a holistic survey and design guidelines." In 2019 15th International Wireless Communications & Mobile Computing Conference (IWCMC), pp. 1108-1114. IEEE, 2019.

[5] Nor, Norhashimi Bin Mohd, and Mohd Hasan Bin Selamat. "Review on green data Center frameworks." In 2014 4th World Congress on Information and Communication Technologies (WICT 2014), pp. 338-343. IEEE, 2014.

[6] Jin, Yichao, Yonggang Wen, and Qinghua Chen. "Energy efficiency and server virtualization in data Centers: An empirical investigation." In 2012 Proceedings IEEE INFOCOM Workshops, pp. 133-138. IEEE, 2012.

[7] Torres, Jordi, David Carrera, Kevin Hogan, Ricard Gavaldà, Vicenç Beltran, and Nicolás Poggi. "Reducing wasted resources to help achieve green data Centers." In 2008 IEEE International Symposium on Parallel and Distributed Processing, pp. 1-8. IEEE, 2008.

[8] Cavdar, Derya, and Fatih Alagoz. "A survey of research on greening data Centers." In 2012 IEEE Global Communications Conference (GLOBECOM), pp. 3237-3242. IEEE, 2012.

[9] Fawaz, AL-Hazemi, Alaelddin Fuad Yousif Mohammed, Lemi Isaac Yoseke Laku, and Rayan Alanazi. "PUE or GPUE: A Carbon-Aware Metric for Data Centers." In 2019 21st International Conference on Advanced Communication Technology (ICACT), pp. 38-41. IEEE, 2019.

[10] Kiani, Abbas, and Nirwan Ansari. "Toward low-cost workload distribution for integrated green data Centers." IEEE Communications Letters 19, no. 1 (2014): 26-29.

[11] Iturriaga, Santiago, Sergio Nesmachnow, Andrei Tchernykh, and Bernabé Dorronsoro. "Multiobjective workflow scheduling in a federation of heterogeneous-green-powered data Centers." In 2016 16th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGrid), pp. 596-599. IEEE, 2016.

[12] Chen, Changbing, Bingsheng He, and Xueyan Tang. "Green-aware workload scheduling in geographically distributed data Centers." In 4th IEEE international conference on cloud computing technology and science proceedings, pp. 82-89. IEEE, 2012.

[13] Aksanli, Baris, Jagannathan Venkatesh, Liuyi Zhang, and Tajana Rosing. "Utilizing green energy prediction to schedule mixed batch and service jobs in data Centers." In Proceedings of the 4th Workshop on Power-Aware Computing and Systems, pp. 1-5. 2011.

[14] Janacek, Stefan, Gunnar Schomaker, and Wolfgang Nebel. "Data Center smart grid integration considering renewable energies and waste heat usage." In Energy-Efficient Data Centers, pp. 99-109. Springer, Berlin, Heidelberg, 2014.

[15] Weerts, Benjamin A., David Gallaher, and Ron Weaver. "Green data Center cooling: Achieving 90% reduction: Airside economization and unique indirect evaporative cooling." In 2012 IEEE Green Technologies Conference, pp. 1-6. IEEE, 2012.

[16] Kim, Jungsoo, Martino Ruggiero, and David Atienza. "Free cooling-aware dynamic power management for green dataCenters." In 2012 International Conference on High Performance Computing & Simulation (HPCS), pp. 140-146. IEEE, 2012.

[17] Capozzoli, Alfonso, and Giulio Primiceri. "Cooling systems in data Centers: state of art and emerging technologies." Energy Procedia 83 (2015): 484-49.