AMERICAN UNIVERSITY OF BEIRUT

GREEN DATA CENTER DESIGN WITH AUTONOMIC POWER-AWARE WORKLOAD ALLOCATION

by
LOUMA AHMAD CHADDAD

A thesis
submitted in partial fulfillment of the requirements
for the degree of Master of Engineering
to the Department of Electrical and Computer Engineering
of the Faculty of Engineering and Architecture
at the American University of Beirut

Beirut, Lebanon
September 2014

AMERICAN UNIVERSITY OF BEIRUT

GREEN DATA CENTER DESIGN WITH AUTONOMIC POWER-AWARE WORKLOAD ALLOCATION

by
LOUMA AHMAD CHADDAD

Approved by:

Ali Chehab

Dr. Ali Chehab, Associate Professor, Electrical and Computer Engineering — Advisor

Dr. Ayman Kayssi, Professor, Electrical and Computer Engineering — Member of Committee

Dr. Imad Elhajj, Associate Professor, Electrical and Computer Engineering — Member of Committee

Date of thesis/dissertation defense: September 9, 2014

AMERICAN UNIVERSITY OF BEIRUT

THESIS, DISSERTATION, PROJECT RELEASE FORM

Student Name: Chaddad Louma Ahmad

● Master's Thesis ○ Master's Project ○ Doctoral Dissertation

☒ I authorize the American University of Beirut to: (a) reproduce hard or electronic copies of my thesis, dissertation, or project; (b) include such copies in the archives and digital repositories of the University; and (c) make freely available such copies to third parties for research or educational purposes.

☐ I authorize the American University of Beirut, **three years after the date of submitting my thesis, dissertation, or project,** to: (a) reproduce hard or electronic copies of it; (b) include such copies in the archives and digital repositories of the University; and (c) make freely available such copies to third parties for research or educational purposes.

Louma September 30, 2014

Signature Date

This form is signed when submitting the thesis, dissertation, or project to the University Libraries

ACKNOWLEDGEMENTS

My first of many thanks goes to God for the blessings he gave me to complete this thesis.

My profound gratitude and recognition are addressed to my advisors, Dr. Ayman Kayssi, Dr. Ali Chehab, and Dr. Imad Elhajj for their invaluable guidance and advice during the elaboration of this project. Your support and patience are greatly appreciated.

I am heartily thankful to my lovely family; my father Ahmad, my mother Elham, and my brothers Ali, Mouhamad and Amer for providing me the best support and encouragement during my whole life. You made me become the person I am.

Special thanks for all my friends especially Layale Constantine, Lama Shaer, Farah Moety, Farah Saab, Salam Doumiati, Nadine El Hajj, Rami Awwad, Sara Abdhallah, Khodor Hamandi, and Joseph Loutfi.

Finally, the biggest thank you goes to my husband Hassan Chehade for all the support and tolerance he provided me all throughout. Thank you for being the source of my strength and the shoulder that I can always count on.

AN ABSTRACT OF THE THESIS OF

Louma Ahmad Chaddad for Master of Engineering
Major: Electrical and Computer Engineering

Title: Data Center Design with Autonomic Power-Aware Workload Allocation

Cloud Computing has revolutionized the technology world and businesses by enabling on-demand provisioning of computing resources allowing users to store, process data and run applications remotely. This entailed the deployment of large-scale data centers containing thousands of computing nodes. However, the energy consumed by Cloud Data Centers today is huge resulting in overwhelming electricity bills and carbon dioxide footprints. In 2010, data centers consumed around 1.5% of the worldwide electricity and are likely to consume further.

This thesis presents a novel approach to reduce energy consumption in a data center. In particular we present a mathematical model that represents the energy dissipation optimization problem. We analytically formulate the server selection problem and the supply air temperature as a Non Linear Programming (NLP), and propose an algorithm to solve it dynamically. A simulation study on SimWare, using real workload traces, shows considerable savings for different data center sizes and utilization rates as compared to three other classic algorithms.

The results prove that the proposed algorithm is efficient in handling the energy-performance trade-off. Moreover, they demonstrate that our algorithm provides significant energy savings and maintains a relatively homogenous and stable thermal state at the different rack units in the data center.

CONTENTS

ILLUSTRATIONS

TABLES

CHAPTER 1

INTRODUCTION

In recent years, Internet traffic has increased dramatically especially after the huge expansion in requests for online marketplaces and social networking websites. Data centers constitute the heart of web ecosystem and are the backbone of online services that are now essential for any business. This large spread of internet-based computing services has increased the density of High Performance Computing (HPC) data centers and entailed the deployment of powerful racks to meet the greater need for storage and computing capacity. Therefore, power feeding of data centers has increased tremendously as well as the complexity of optimizing the energy consumption efficiency of their operation.

Cooling cost in a data center accounts for about 30% of its total energy cost [20] as illustrated in figure 1.1. Current generation of servers consumes around 350 Watts of power at maximum utilization; a regular 42U rack would consume around 8 KW with much of this power released as heat. This energy consumption is expected to escalate in coming years, to 55KW per rack with the use of higher power servers that require 5 KW per chassis [21]. Furthermore, the huge energy consumed by data centers today results in overwhelming electricity bills and carbon dioxide footprints. Data centers in USA today consume 7.4 Billion USD annually according to a U.S. Environmental Protection Agency ENERGY STAR Program report. Figure 1.2 shows that the global energy consumed by data centers has increased by 56% from 2005 to 2010 [36]. Besides, carbon dioxide (CO2) emissions of the information and communications technology (ICT) industry are estimated to be 2% of the worldwide emissions, which is equivalent to the emissions of the aviation industry [37].

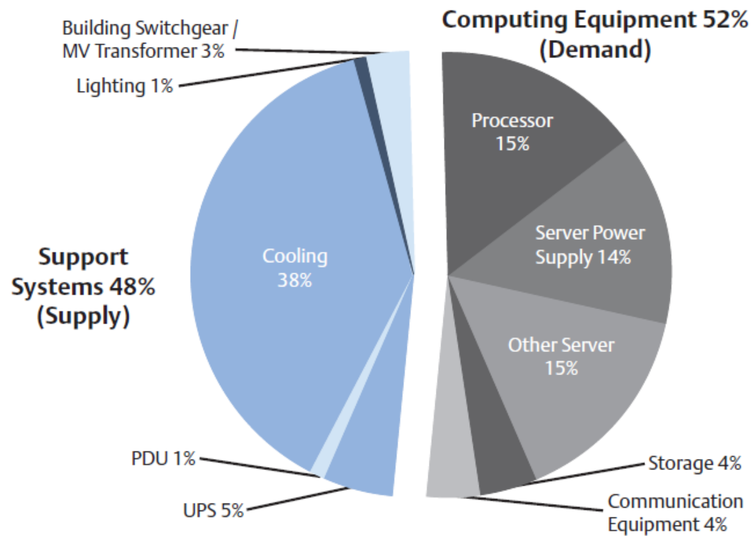

Figure 1.1: Overall data center costs [20]

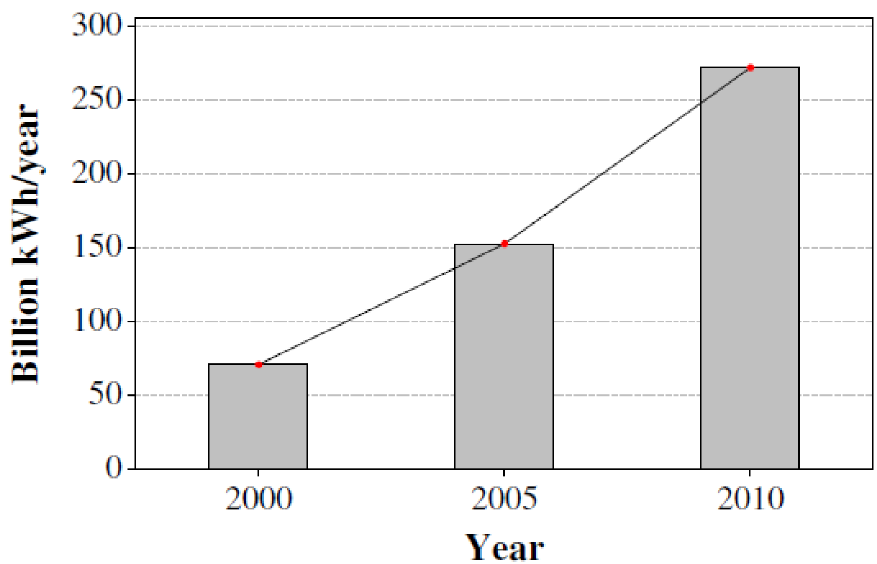

Figure 1.2: The global data center energy consumption 2000-2010 [36]

As estimated by Koomey [38], the energy consumed by data centers will keep on rising quickly except if new efficient resource management solutions are practiced. To address the high-energy use problem in data centers, it is required to eliminate inefficiencies in the way computing resources are used when serving application workloads. One way to achieve this objective is to improve the management algorithms and the resource distribution.

1.1. Why energy-efficiency in data centers?

During the early days of the Internet, data centers provided few services using single-purpose server machines. Traditionally, data centers were storage endpoint from where users can retrieve data using HTTP, FTP and other network protocol requests using a client-server model.

The evolution of computing models and services, specifically the emergence of cloud computing and social media, transformed large-scale data centers from storage media to service providers. They typically serve millions of users internationally. More and more services entailed the deployment of large-scale data centers containing thousands of computing nodes with hundreds of gigabit bandwidth to the Internet, and petabytes of storage. Hence, advance in the design technologies is continually increasing the computing and storage capacities of data centers. A side effect of this capacity growth is a huge increase in the power density and energy consumption of data centers. Thermal management is becoming a critical concern for data centers managers to find efficient greening strategies that would reduce cooling and computing energy at the same time. There are many organizations that are dedicated to promote and develop techniques for improving power efficiency and reducing the energy consumed and as a result reducing CO_2 emission. These include Green Electronics Council, ECO2Clouds, Eco4Cloud, FIT4Green, The Green Grid, International Professional Practice Partnership (IP3), and Climate Savers Computing Initiative (CSCI).

What is needed is the design of a dynamic energy-aware workload assignment algorithm to evenly allocate the computational tasks among the rack units based on their power consumption, and hence heat dissipation.

1.2. Thesis Motivation and Contribution

There are two main motivations behind research into the energy efficiency of data centers, or most fields of the ICT industry in general. The first is cost, since all forms of computing automation consume energy, incurring a cost in the form of the electricity bill. The second is environmental sustainability, since as the amount of computers increases, so does the overall energy consumption attributed to computing. The research area has otherwise been called "Sustainable computing" or "Green ICT" [39].

The focus of this research is dynamic energy-efficiency in data centers. We propose to design and simulate a dynamic power-aware workload allocation algorithm to allocate computational task workloads among servers in cloud-based data centers. Our model takes into consideration the allocation of workload, the heat recirculation in the server room, and the air temperature provided by the Computer Room Air Conditioning (CRAC). The main contribution resides in presenting a precise mathematical formulation of the optimization problem of energy consumption in data centers along with a novel solution algorithm that minimizes the total data center energy consumed. The cooling energy is reduced by selecting an optimum cold supply air temperature value of the CRAC, while the computing energy is reduced by appropriately assigning incoming tasks to optimal servers and set the proper velocity for the CPU fan of each server. The proposed system design is simulated in SimWare [1] using workload traces from experimental datacenters [40].

The research methodology of this thesis comprises some successive steps summarized below:

1. Conduct theoretical study of existing algorithms to obtain insights into the design and theoretical performance estimation of such algorithms.
2. Formulate mathematically the problem of minimizing data center energy costs taking into consideration all the accounting factors of heat inside the data center.
3. Solve the optimization problem using solver packages.
4. Develop a systematic algorithm to efficiently distribute workload among servers.
5. Evaluate the performance of the proposed algorithm in the SimWare simulation toolkit using real-world workload traces collected from experimental datacenters.

1.3. Thesis Organization

The main chapters of this thesis are organized as shown in Figure 1.3. The remainder of the thesis is structured as follows. Chapter 2 describes a survey and classification of energy-efficient computing schemes. Also, it presents the scope of this thesis as well as its positioning in the field. Chapter 3 provides background information and presents the data center power model used. Chapter 4 explains the optimization power problem and our solution algorithm. Chapter 5 presents simulation results of the proposed algorithm in SimWare. It also provides an evaluation of our approach by comparing its results to three classic algorithms. Chapter 6 concludes the thesis by summarizing the key findings and discussing some future directions.

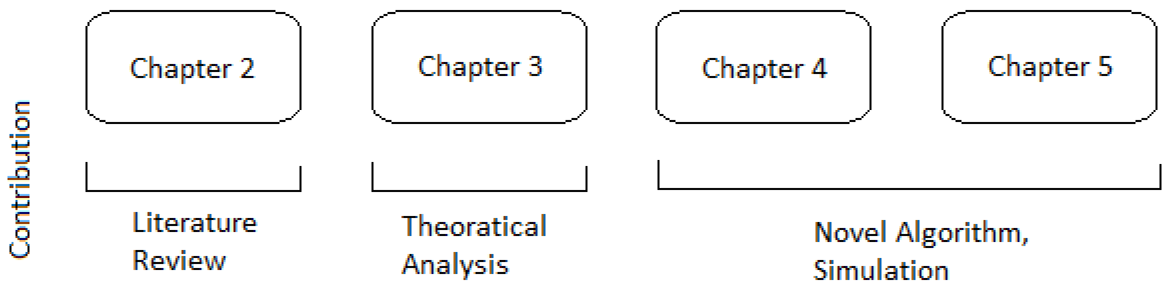

Figure 1.3: The thesis organization

CHAPTER 2

LITERATURE REVIEW

The purpose of this chapter is to give an overview of the most prominent approaches of reducing energy consumption at the data center levels, classify the methods, highlight their main shortcomings, and position our thesis within the research area.

The literature shows a considerable work done in the field of optimizing energy consumption in data centers. It has been tackled in both mechanical and computer science academia. The former focusses on finding best cooling mechanisms and modifying the physical layout of data centers. The latter tries to find efficient algorithms for load distribution on heterogeneous server racks in the data center. Our literature will mention both perspectives but our ultimate scheme will focus on designing the best performance-efficient scheduling algorithm.

2.1. Proposed Solutions

2.1.1. Dynamic Voltage and Frequency Scaling (DVFS)

Dynamic Voltage and Frequency Scaling (DVFS) is one of the first power saving techniques studied and deployed. DVFS adjusts the frequency of the processor in order to reduce power dissipation consumption and therefore heat of CPU. It is based on the fact that power in a chip is proportional to V^2f where V is the voltage and f is the operating frequency. In [5], the scheme implemented in CloudSim predicts, from the resource utilization log, the CPU utilization and adjusts the voltage and frequency of the CPU accordingly to save power. This scheme uses “Linear Predicting Method” (LPM) and “Flat Period Reservation-Reduced Method” (FPRRM) to improve the prediction method based on M/M/1 queue. In [19], several power-aware Virtual

Machine VM distribution schemes based on DVFS are evaluated. Simulation results prove that these schemes can reduce power consumption of data centers.

DVFS decreases the energy dissipation by reducing the supplying voltage and frequency. However, this would result in slowing down the execution time and would raise a performance concern on the latency resulting from using this approach.

2.1.2 VM consolidation

The fact that an idle server consumes around two third of its peak power suggests to minimize the number of running servers by concentrating the workload in the least possible number of computing servers and shutting down unused servers. This is referred to as VM consolidation and it has become a common approach used for energy efficiency. Many commercial cloud management products including Lanamark Suite [6], VMware Capacity Planner [7], IBM WebSphere CloudBurst [8], Novell PlateSpin Recon [9] and CapacityIQ [10] are implementing it. The authors in [11] propose and evaluate a VM consolidation policy that dynamically turns on and off servers based on user demand which is statistically estimated. Beloglazov et al, [12] present an algorithm that dynamically consolidates virtual machines using adaptive utilization thresholds for servers. This approach first selects VMs that have to be migrated based on the Minimization of Migration (MM) policy, then the selected VMs are placed according to a modification of the Best Fit Decreasing (BFD) algorithm. Most of the nodes that became idle after the migration are switched off. In [13], Xiaoli et al. develop a workload-scheduling algorithm aiming to minimize the computing resources. It is an improvement of the classical Bin-Packing algorithm. In addition, the proposed technique avoids migration of virtual machines, which wastes a significant amount of energy. This is realized by setting a threshold. A simulation comparing this algorithm to the Best Fit Algorithm is performed using both C++ and Matlab to prove that the proposed algorithm is more energy efficient. The work in [14] presents an Energy-aware technique that allocates the maximum possible number of VMs on a physical server taking into consideration the application QoS. To realize that, a model of the energy consumed and the time needed for the execution of a standard HPC workload benchmark is developed. It contains all possible combinations of allocating the VM. The algorithm considers the best allocation of VMs based on optimizing the energy consumed and/or the execution time. Hoyer et al [15] assign based on some heuristics VMs to physical machines (PMs) by exceeding the maximum capacity of resources of these PMs because VMs do not use normally all their allocated resources. Their allocation approach takes into consideration both the uncorrelated and the correlated workload.

In [16], the authors develop a provisioning strategy based on prediction using Error Correction Neural Network (ECNN) and Linear Regression as learning algorithms. Kliazovich et al. propose in [17] an approach that minimizes the amount of computing resources and takes into consideration when servers run continuously at peak loads, and as such, they reduce their reliability. In [18], Wang et al. describe their algorithm for VM consolidation when network bandwidth demands are dynamic which makes it difficult to use traditional VM allocation algorithms. The problem is formulated as a Stochastic Bin-Packing problem where bandwidth demands on VMs are modeled as probabilistic allocations.

The main disadvantage of the VM consolidation approach is that when there is a need to turn on a server, a significant amount of energy is consumed. Also, a time delay can occur due to the setup time, which can increase the response time and lead to performance degradation.

2.1.3. Hybrid approaches

Running several physical machines increases the power consumption, since servers consume an important amount of energy even in idle mode. On the other hand, a small number of running servers would require running at higher frequencies to meet the higher percentage of utilization of these servers and therefore the energy consumption would increase. Consequently, [22] and [23] present hybrid approaches that find a balanced tradeoff between the number of running physical machines and the voltage/frequency of the CPU on these servers.

2.1.4. ***Temperature-aware workload scheduling***

Temperature-aware workload scheduling is another approach used in optimizing the efficient cooling of the data center. It consists of scheduling computational workloads in a way to reduce the cooling energy consumed. A main concern in this respect is the possible Heat Recirculation from server outlets to inlets that increases the temperature of inlets and causes hot spots. To eliminate host spots, data center operators control CRAC to supply lower temperature and therefore the cooling cost increases. The authors in [24] present a heuristic technique that minimizes the heat recirculation inside the data center room (minimize-heat-recirculation (MINHR)) and therefore would enhance the cooling efficiency. In [25], the author introduces three task placement algorithms that ensure that the temperature of the outlet is uniform, the dissipated power by a server is minimal, or distribution of workloads among servers is uniform. Temperature-aware workload scheduling proved to be an efficient scheduling approach. Nevertheless, it neglects other important factors contributing to the energy dissipation in data centers other than heat recirculation. Further energy savings could be realized by considering those factors.

2.1.5. ***Other approaches***

Diverse solutions that focus on designing new cooling techniques for a better control of the air conditioning to improve energy consumption are proposed.

An intelligent cooling technique, introduced by Beitelmal et al. in [26], suggest the placement of adaptive vent tiles (AVT) where the openings are adjustable to control the cooling demand of racks. The main challenge of this method is that it requires the synchronization between CRAC units and local AVT for the cooling to be optimal. However, data centers lack the sensors infrastructure on the racks side, and the temperature is only sensed on the return air, which makes this approach less practical. In [27], AVT for local cooling control is introduced. Their method captures the effects of local and zonal cooling mechanisms on the inlet temperatures of the racks containing servers. A predictive controller supports this model and fulfills the temperature needs of the racks in the data center to reduce the cooling energy intake.

The model in [28] presents a hot aisle containment approach to minimize recirculation and limit mixing of hot air coming from the racks outlet with the cold air produced by the CRAC units. This mechanism divides the data center room into a number of overlapping CRAC zones of influence. It uses the predictive controller discussed in [27] to regulate the intensity of cooling generated by the CRAC units in order to adjust the temperature of rack inlets in each zone. VanGeet presents in [29] a number of best practices guidelines to optimize energy-efficiency in data centers. These practices focus on heat exclusion, data center air management, thermal properties of IT equipment, and air conditioning systems. The air inlet temperature of servers must always be kept under a certain threshold to avoid thermal redlining.

The solution considered in [30] proposes the allocation of incoming workload onto servers located in data center locations that are easier to cool than others. Figure 2.1 shows the temperature variation in an HP data center located in Palo Alto. Certain physical areas in the data center are easier to cool than others as shown in Figure 2.2.

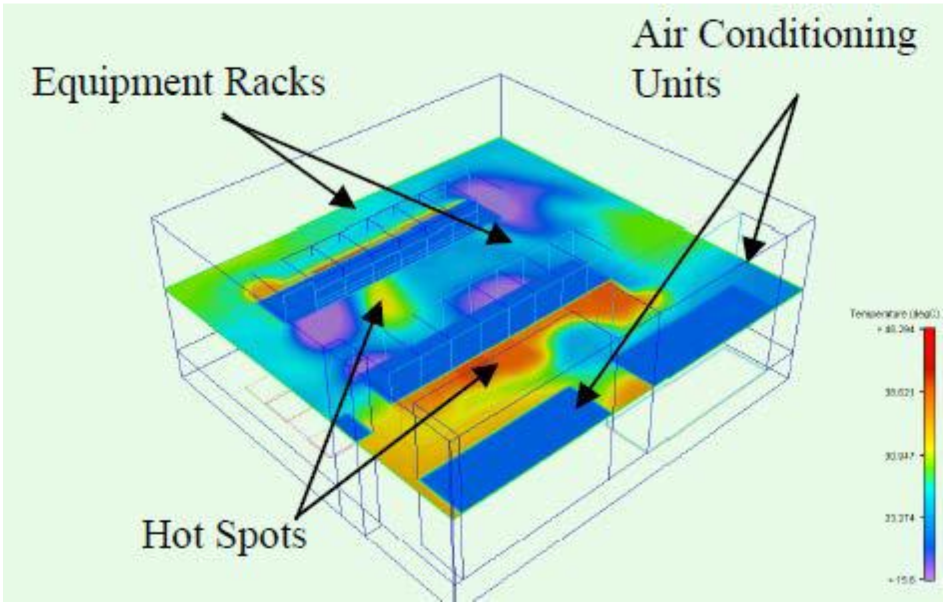

Figure 2.1: Temperature variation in data center [30]

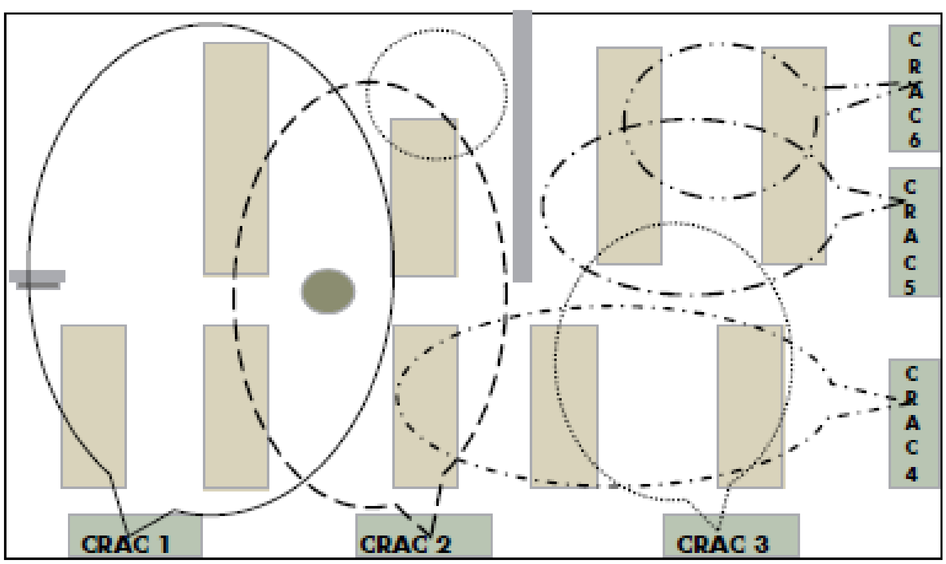

Figure 2.2: CRAC regions of influence [30]

The thermal correlation index (TCI), given by the equation below, is a metric used to define regions of influence in a data center. It calculates the response at the i^{th} rack inlet sensor to a step change in the temperature supplied by the j^{th} CRAC. This metric ranks the different servers physical locations in the data center and thus can be used to develop an algorithm that prefers to place computational jobs on servers that are in better efficient cooling locations inside the data center. The main limitation of this technique is that it is does not support dynamic scheduling tasks and it only copes with batch-processing tasks.

$$TCI_{i,j} = \frac{\Delta T_i}{\Delta T_{crac,j}}$$

Some algorithms introduced in [31, 32] predict the rack inlet temperature relying on intuition-based algorithms, and assign workload tasks to the rack having the lowest inlet temperature. These algorithms have low accuracy since they are intuitive by nature.

Another approach to reduce energy cost in data centers is proposed in [33]. It suggests Content Delivery Networks (CDN) where content is reproduced in each CDN center situated in different geographical places. Incoming traffic is then routed to CDNs where electricity prices are the lowest knowing that electricity varies with the geographical location and time.

2.2. Comparison between our System and Other Solutions

During our review we came across various approaches for reducing data center energy consumption. However, these techniques lack the theoretical formulation of the optimization problem of energy consumption in data centers. Also, there is a lack of a holistic solution that takes into consideration all the factors contributing to the energy dissipation such as heat recirculation, allocation of workload, CPU fan speed control, and temperature of air provided by the CRAC.

This study differs from the prior analytic literature in three aspects. First, we consider dynamic server provisioning instead of relying on batch-processing tasks. Second, we provide an efficient holistic solution. Third, our system minimizes the energy consumption of a data center without having to compromise performance. It presents a complete solution to dynamic energy-efficient task distribution under Quality of Service (QoS) constraints.

We aim that our scheme ultimately results in a homogenous thermal state of racks. This will reduce the energy consumption and therefore would lead to significant cost savings.

CHAPTER 3

PRELIMINARIES

To develop new strategies for power management, it is necessary to develop a model of dynamic power consumption. In this chapter, we provide an overview of a typical data center layout, the cooling system, cold/hot aisles, and the arrangement of servers. Next, we present the power model of the data center and the various metrics to measure its efficiency.

3.1. Data Center Configuration

In modern data centers, computing servers are commonly arranged in rows of racks of blade systems arranged in chassis. Computing servers consume energy according to workload and hardware characteristics and generate heat that raises the room temperature. A typical data center layout is shown in figure 3.1 and consists of alternating “hot aisles” and “cold aisles” that separate rows of equipment racks. Typically, the racks are 42-U size racks. CRAC units are the zonal actuators that stream cooled air to inlet sides of servers via perforated tiles raised in the floor of cold aisles [41]. The heated air is then returned from exhausts of racks to CRAC intakes. This configuration ensures the separation of the cool inlet air from hot exhaust air. However, the recirculation of air in the higher part of the room can cause the infiltration of hot exhaust air into the cold aisles resulting in an increase of inlet temperature and the appearance of “hot spots” whose temperature is larger than average. Hot spots compound the thermal problem and entail a harmful effect on the performance of data centers.

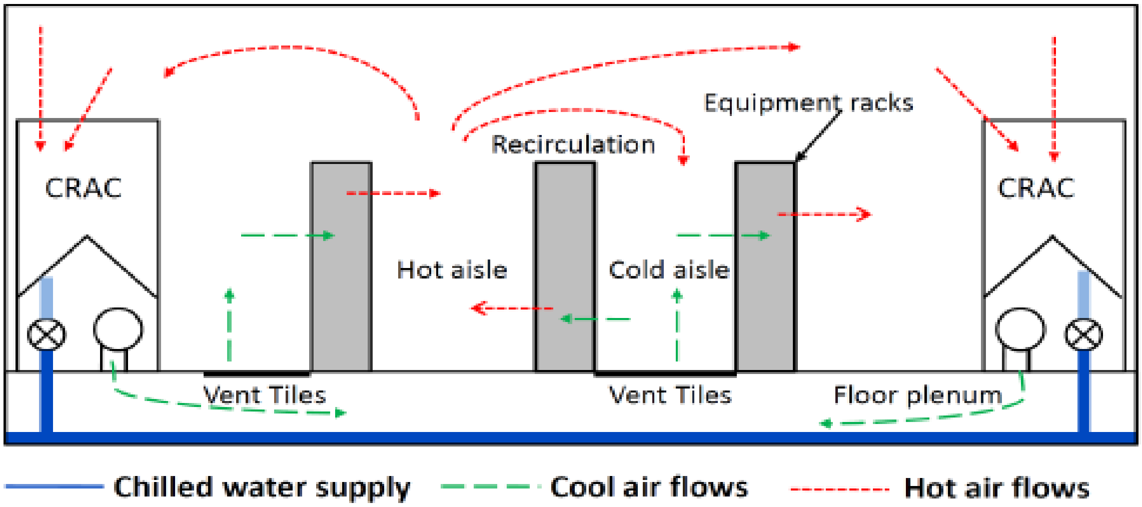

Figure 3.1: Typical data center layout [28]

3.2. Power Model for the Servers

3.2.1 Server Computing Power

Fan et al. [44] formulated the relationship between power consumption and CPU utilization of a server. The proposed model illustrates the fact that power consumed by a server grows almost linearly with the CPU utilization from the value of power consumed when the server is idle up to the power consumed when the server is fully utilized. This relationship can be represented as shown in (3.1) and illustrated in figure 3.2.

$$P\,(u) = P_{idle} + (P_{busy} - P_{idle}) \times u \qquad (3.1)$$

Where P is the estimated power consumption of the server, P_{idle} is the power consumed by the server in the idle state, P_{busy} is the power consumption when the server is fully utilized, and u is the CPU utilization.

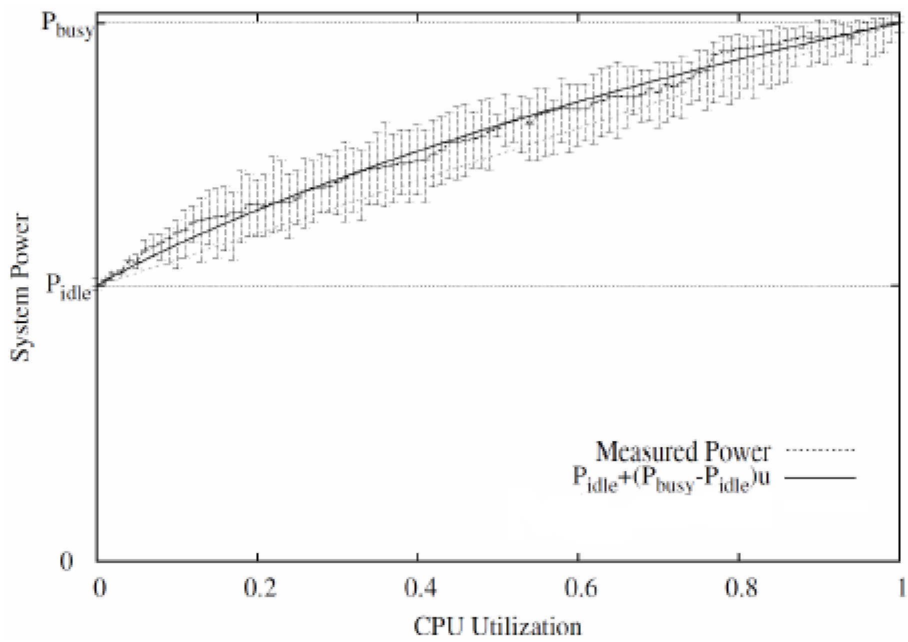

Figure 3.2. Server power consumption depending on the CPU utilization [44]

According to [42] each server consumes half of its peak power in the idle state, therefore the power consumed by the server can be calculated by:

$$P\,(u) = w + w \times u \qquad (3.2)$$

Where w is the power consumption when the server is idle.

3.2.2 Server Fan Power

Individual servers are equipped with a fan to decrease the temperature within the CPU. Many studies have neglected the importance of these fans as an essential element if the thermal system architecture. In fact, CPU fans consume a significant amount of power at a cubical rate related to their speed, making high speeds of the fan expensive. To the best of our knowledge, this study is the first to theoretically calculate the power consumption of CPU fan.

To calculate the power consumption of a fan, we rely on the laws of fan affinity and the law of convective heat transfer [1]. The law of convective heat transfer assumes that the heat transfer is proportional to the amount of air and the difference of temperature between the surrounding air and the cooling object:

$$\text{Heat transfer (watts)} \propto \text{Temperature difference} \times \text{Amount of air} \qquad (3.3)$$

The law of fan affinity assumes that the amount of air is proportional to the speed of the CPU fan denoted by Fan_{rpm}:

$$\text{Amount of air} \propto Fan_{rpm} \qquad (3.4)$$

And that the power dissipation of the fan denoted by Fan_{power} is proportional to the third power of the speed of Fan_{rpm}:

$$Fan_{power} \propto Fan^{3}_{rpm} \qquad (3.5)$$

Equations (3.3) and (3.4) imply that the fan speed is proportional to the ratio of heat transfer and the temperature difference:

$$\frac{\text{Heat transfer (watts)}}{\text{Temperature difference}} \propto \text{Fanrpm} \qquad (3.6)$$

The CPU fan must eliminate the heat generated by the CPU at any time. Therefore, when the fan runs at its highest speed denoted by MaxRPM, it should remove w Watts (the maximum additional CPU power) at the maximum operable temperature $T_{emergency}$. At this operating point, the fan consumes its maximum power denoted by MaxPower. The temperature difference in this case would be calculated as:

$$\text{Temperature difference} = T_d - \text{Temergency} \qquad (3.7)$$

We denote by T_d the die temperature of the processor, T_{inlet} the current inlet temperature of the server at time t, currentCPUpower the power of CPU that needs to be removed at time t, currentRPM the current speed of the fan, u the utilization of the server at time t, FanPower the power consumed by the CPU fan at time t.

The latter situation implies the Temperature difference at time t to be:

$$\text{Temperature difference at } t = T_d - T_{inlet} \qquad (3.8)$$

In this case, we can write based on (3.6):

$$currentRPM = \frac{\frac{currentCPUpower}{T_d - T_{inlet}}}{\frac{w}{T_d - T_{emergency}}} \times MaxRPM$$

$$currentRPM = \frac{\frac{w \times u}{T_d - T_{inlet}}}{\frac{w}{T_d - T_{emergency}}} \times MaxRPM$$

$$currentRPM = \frac{u \times \left(T_d - T_{emergency}\right)}{T_d - T_{inlet}} \times MaxRPM \qquad (3.9)$$

Based on (3.5) we can write:

$$FanPower = MaxPower \times \left(\frac{currentRPM}{MaxRPM}\right)^3$$

Substituting the expression from (3.9) for currentRPM, we obtain:

$$FanPower = MaxPower \times \left(\frac{u \times \left(T_d - T_{emergency} \right)}{T_d - T_{inlet}} \right)^3 \qquad (3.10)$$

3.2.3 ***Server Total Power***

The total power consumption of the server is then given by:

$$P_{server} = P_{idle} + currentCPUpower + FanPower$$

$$P_{server} = w + w \times u + MaxPower \times \left(\frac{currentRPM}{MaxRPM} \right)^3 \qquad (3.11)$$

The total energy cost consumed by n servers during a period T is presented as:

$$E_{IT} = \sum_{i=1}^{n} \left[w + w \times u_i + MaxPower \times \left(\frac{currentRPM_i}{MaxRPM} \right)^3 \right] \times T \qquad (3.12)$$

3.3. Data Center Power Modeling

3.3.1 Power Consumption of the CRAC Unit

The CRAC unit is responsible of the cooling of a data center. The efficiency of CRAC is modeled by its coefficient of performance (CoP), which is defined as the ratio of the heat removed by the CRAC unit (Q) to the total work needed to remove that heat (E). For instance, ratio of two designates that to remove heat at the rate of 1000 W, the work performed by the CRAC is 500 W.

$$CoP = \frac{Q}{E} \qquad (3.13)$$

We will demonstrate in the following the fact that CoP is greater than 1. We define Q_c as the heat drawn out of the CRAC, and Q_h the waste heat dumped into the data center room. CoP in this case is given by:

(3.14)

The ratio Q_h/Q_c is greater than 1 and less than 2. Therefore CoP is greater than 1.

A greater CoP indicates more efficient cooling, needing less work to remove a constant amount of heat. Usually the higher the supplied air temperature the better the CoP. In this thesis, we use the COP model of a typical CRAC unit in a HP utility data center illustrated in figure 3.3 [24].

The COP curve in figure 3.3 is given by:

$$\mathrm{CoP}(T_s) = 0.0068T_s^2 + 0.0008T_s + 0.458 \qquad (3.15)$$

where T_s denotes the cold supply temperature streamed by the CRAC (in degrees C).

The energy cost of the CRAC unit may be quantified as:

$$E_{CRAC} = \frac{E_{IT}}{CoP(T_s)} \text{ [24]} \qquad (3.16)$$

Where E_{IT} is the summation of energy consumption over all servers

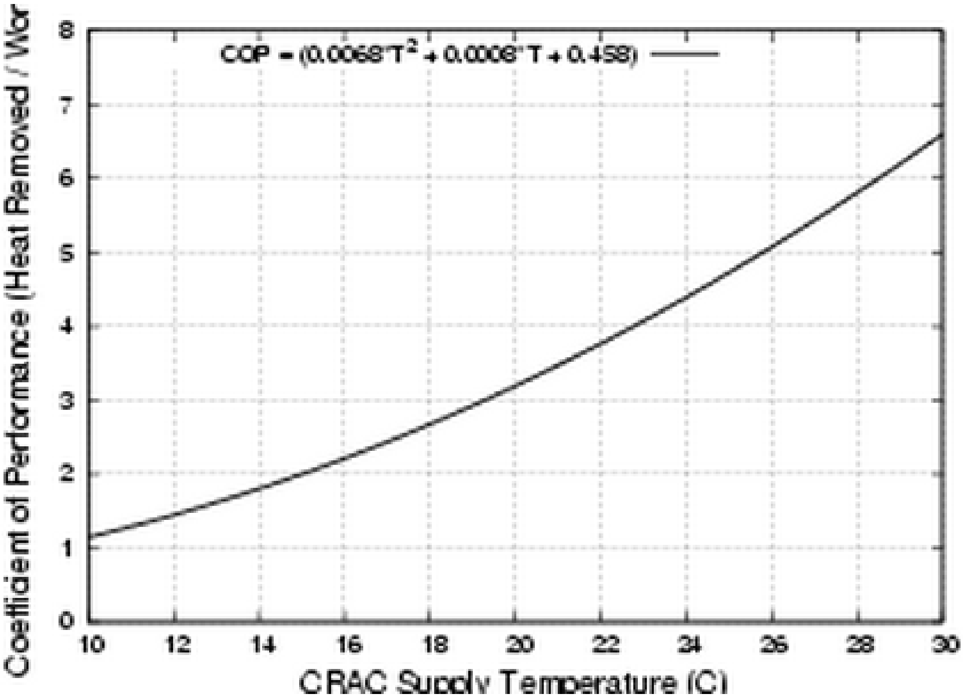

Fig.3.3. COP curve for the CRAC units at the HP labs utility data center [24].

3.3.2 Total Power Consumption

The total energy dissipation in a data center is composed of the energy consumed by the computing elements E_{IT} and the CRAC cooling energy E_{AC} [43].

$$E_{total} = E_{IT} + E_{AC} \qquad (3.17)$$

Substituting the values of E_{IT} and E_{AC} given in (3.12) and (3.15) respectively into (3.16) gives the total energy consumption as follows

$$E_{Total} = \left(1 + \frac{1}{0.0068T_s^2 + 0.0008T_s + 0.458}\right)\sum_{i=1}^{n}\left[w + w \times u_i + MaxPower \times \left(\frac{currentRPM_i}{MaxRPM}\right)^3\right] \times T \qquad (3.18)$$

3.4. Data Center Efficiency Metrics

Researchers have offered different means to measure datacenter efficiency.

Power usage effectiveness [1] is a very common metric introduced by the Green Grid. It is defined as the ratio of total facility power consumed to power brought to IT computing equipment.

$$PUE = \frac{TotalFacilityPower}{ComputingPower} \qquad (3.19)$$

In useful terms, a PUE value of 1 means that all power brought into the data center is being delivered to power computing equipment.

Latency is by definition the delay or wasted time that increases response time of a job beyond the desired one.

Standard Deviation of inlet temperatures is a measure of how inlet temperatures of different servers are spread out. A low standard deviation indicates that the inlet temperatures tend to be close to each other and are distributed in a small range; i.e. the thermal state of servers is homogenous.

CHAPTER 4

SYSTEM DESIGN

4.1. Power consumed in computation affects the inlet temperatures

As computing devices in a data center produce heat when running workloads, the CRAC cooling system must supply cold air to their air inlets to keep their temperature from exceeding the critical emergency temperature; i.e. the highest operational temperature specified by the manufacturer. On the other hand, the inlet temperature of a server in a data center derives from the mixture of cold air supplied from the CRAC and hot air recirculated from the other servers. In other words, the inlet temperature of a server will experience a rise above the supply temperature due to heat from recirculation [25].

According to [44] heat recirculation can be estimated using a heat distribution matrix $D_{n\times n}=\{d_{ij}\}$, where d_{ij} denotes the contribution of server j to the inlet temperature increase of server i in a data center of n servers. This matrix can be generated using Arizona State University's BlueTool. The equation relating the inlet temperature for each server i to the CRAC supply temperature is then formulated as:

$$T_{inlet}=T_s+\mathbf{D}\ \mathbf{P} \qquad (4.1)$$

Where T_{inlet} and T_s are the corresponding inlet temperature and the supply cold air temperature vectors of n elements, respectively and $P = [P_1,\ldots,P_n]$ is the vector of power dissipation of the n servers.

4.2. Data Center Power Minimization

4.2.1 Problem Statement

Our objective is to minimize the total data center energy dissipation given in (3.17) by 1) selecting the optimum value of T_s, 2) determining optimal servers to assign incoming tasks, and 3) setting the proper fan velocity for each server.

We define a binary variable for each server that indicates whether a server is assigned the job or not. Xi denotes the variable for the ith server. Given a data center of n servers, each server with power characteristic w, and with C_{tot} number of CPUs; a task demanding C CPUs; the heat distribution matrix D, $\overrightarrow{T_{in}} = \left[T_{in}^1, T_{in}^2, ..., T_{in}^n\right]$ the vector of inlet temperatures for n servers, the energy optimization problem for serving an incoming task is as follows:

Find Ts, the vector X (of length n), and the vector currentRPM (of length n) every 1 second:

$$
\begin{aligned}
& Minimize\left\{\left(1+\frac{1}{0.0068T_{\sup}^2+0.0008T_{\sup}+0.458}\right)\sum_{i=1}^{n} x_i\left[w+w\times u_i+MaxPower\left(\frac{currentRPM_i}{MaxRPM}\right)^3\right]\right\} \\
& where \\
& u_i=\frac{C}{C_{tot}} \\
& currentRPM_i=\frac{u_i\times\left(T_d-T_{emergency}\right)}{T_d-T_{in}^i}\times MaxRPM \\
& s.t. \\
& T_{\sup}\in\left[T_{low},T_{emergency}\right] \\
& T_{inlet}^i\in\left[T_{low},T_{emergency}\right]; i=1,2,...,n \\
& \min RPM\le\left(currentRPM_i\right)\le\max RPM; i=1,2,...,n \\
& \sum_{i=1}^{n} x_i=1; x_i=0 or 1; i=1,2,...,n \\
& \overrightarrow{T_{in}}\ge\overrightarrow{T_{\sup}}+D_{n\times n}\vec{P}; where \\
& \overrightarrow{T_{in}}=\left[T_{in}^1,T_{in}^2,...,T_{in}^n\right] \\
& \vec{P}=\left[P_1,P_2,...,P_n\right]
\end{aligned}
$$

4.2.2 Solution to the Optimization Problem

In order to solve the mathematical optimization problem, we use IBM ILOG Cplex Optimizer, the well-known solver package for C++. Cplex is fast, robust, and capable of solving complex models and producing precise decisions.

We assume a homogenous data center with w=130W, T_{low}=10°C, $T_{emergency}$=30°C, minRPM=500rpm, MaxRPM=3000rpm, MaxPower=15W. We generate the heat distribution matrix $D_{n\times n}$ is using Arizona State University's BlueTool. The input to BlueTool is a physical description of a data center in the Computer Infrastructure Engineering LAnguage (CIELA). Figure 3.4 illustrates an example of CIELA which is a high-level XML-based specification language to represent the generic layout of the data center. It defines the room architecture consisting of the wall locations, the CRAC position as well as the flow rate, the specification of equipment such as racks, vents, tiles, etc.

The nonlinear programming problem presented in section 4.2.1 is programmed in the CPLEX software as shown in Appendix A. Our algorithm selects an optimum cold supply air temperature value of the CRAC, assigns incoming tasks to optimal servers and sets the proper velocity for the CPU fan of each server.

Figure 4.1. CIELA code example for a homogenous data center

4.3. Other approaches

We present the following classic algorithms to the workload assignment problem, which will be compared with our algorithm. We will refer to our proposed algorithm as "OptimalAlg".

LowTempFirst: This scheme assigns more jobs to servers with low inlet temperatures and fewer jobs to servers with high inlet temperatures [41]. The goal is to reach a uniform temperature distribution inside the data center.

BestPerformance: This algorithm assigns the upcoming jobs to the server with the least utilization [25].

MinHR: It calculates Heat Recirculation Factor (HRF) for each server and assigns jobs according to the ratio of each server's HRF to the sum of all HRFs. This approach allocates fewer jobs to a server that causes higher recirculation [42].

CHAPTER 5

SIMULATION

5.1. Introduction

In this thesis, we use the SimWare toolkit to perform our simulations. SimWare is a holistic Warehouse-Scale data center simulator. It analyses the consumed power by servers, fans, and cooling units. It also considers the effect of heat recirculation and the air travel time from the CRAC to servers. In addition, SimWare can access airflow management strategies and server placement inside the data center [1]. All these features would evaluate our proposed model accurately and effectively. The input for SimWare is a Standard Workload Format (SWF) file where traces are collected from real data center clusters.

5.2. Simulations and Results

In order to test accurately the efficiency and performance of OptimalAlg we consider the effect of two different factors:

1- Data center size

2- Workload utilization rates.

The degree of energy savings will be verified by comparing it to Lowtempfirst, BestPerformance, and minHR.

5.2.1. Setup 1: data center size

We simulate using SimWare three different sizes of warehouse-scale data center as described in Table 5.1. Each server has a 130-W TDP Xeon E7-2850 processor with 10 cores.

Each blade server consumes 130 W in its idle state. According to A1-class server specification for datacenters, the servers' emergency temperature is set to 30°C [45]. We run an SWF file workload from 10 high-performance computing clusters in the Shared Hierarchical Academic Research Computing Network (SHARCNET) in Canada. The log comprises approximately 1.2 million accounting jobs from December 2005 through January 2007.To simplify our simulation, we run 24,112 jobs corresponding to five days running jobs by the cluster.

Table 5.1: Various data center sizes simulation test cases

Test Case	**Data Center Size**	**Number of Servers**
A	Big	500
B	Medium	250
C	small	150

We compare the simulations resulting from OptimalAlg (OA), BestPerformance (BP), LowTempFirst (LTF), and MinHR for the different test cases. Results are reported in Tables 5.2, 5.3, and 5.4 for test cases A, B, and C, respectively.

Table 5.2: SimWare Simulations Results for Test Case A

Scheduling Algorithm	Min_HR	BP	LTF	OA
NUMBER_OF_CHASSIS	50	50	50	50
SERVERS PER	10	10	10	10
CORES PER SERVER	10	10	10	10
FINISHES_AT_DAY	5	5	5	5
Average Supply Temperature (°C)	23.7295	23.721	23.6869	23.816
Peak Power (W)	193823	193823	193823	174914
Average Power Consumption (W)	92541.2	92370	92531.6	90262.4
PUE	1.24854	1.24999	1.24938	1.24256
Average Latency (sec)	75116.4	74159.4	75172.4	72740.1
Standard Deviation of Inlet	2.12	2.11	2.1	2.08

Table 5.3: SimWare Simulations Results for Test Case B

Scheduling Algorithm	Min_HR	BP	LTF	OA
NUMBER_OF_CHASSIS	25	25	25	25
SERVERS PER	10	10	10	10
CORES PER	10	10	10	10
FINISHES_AT_DAY	5	5	5	5
Average Supply Temperature	24.5066	24.3575	24.2884	24.8165
Peak Power (W)	93237.8	93237.8	93237.8	83899
Average Power Consumption	51187.3	50417.8	51057.1	48461.4
PUE	1.24109	1.2413	1.24103	1.23193
Average Latency (sec)	135364	135195	135368	132472
Standard Deviation of Inlet	1.87	1.87	1.88	1.81

Table 5.4: SimWare Simulations Results for Test Case C

Scheduling Algorithm	Min_HR	BP	LTF	OA
NUMBER_OF_CHASSIS	15	15	15	15
SERVERS PER CHASSIS	10	10	10	10
CORES PER SERVER	10	10	10	10
FINISHES_AT_DAY	5	5	5	5
Average Supply Temperature (°C)	24.996	24.7849	24.6189	25.002
Peak Power (W)	54604.4	54604.4	54604.4	49070.7
Average Power Consumption (W)	34342.5	33165.1	34221.4	31412.7
PUE	1.23358	1.23233	1.23348	1.22292
Average Latency (sec)	226828	225067	226833	219628
Standard Deviation of Inlet T°	1.65	1.64	1.68	1.56

In order to accurately analyze the simulations, we calculate the percentage savings of OptimalAlg in comparison to BestPerformance, LowTempFirst, and MinHR in terms of peak power, average power consumption, and latency. Results are summarized in Tables 5.5, 5.6, and 5.7 for test cases A, B, and C, respectively.

Table 5.5: Percentage of savings of OA in Test Case A

Scheduling Algorithm	Min_HR	BP	LTF	OA
NUMBER_OF_CHASSIS	50	50	50	50
SERVERS PER	10	10	10	10
CORES PER SERVER	10	10	10	10
FINISHES_AT_DAY	5	5	5	5
Peak Power (W)	9.75%	9.80%	9.80%	174914
Average Power Consumption (W)	3.76%	3.40%	3.65%	90262.4
Average Latency (sec)	3.26%	2.10%	3.40%	72740.1

Table 5.6: Percentage of savings of OA in Test Case B

Scheduling Algorithm	Min_HR	BP	LTF	OA
NUMBER_OF_CHASSIS	25	25	25	25
SERVERS PER CHASSIS	10	10	10	10
CORES PER SERVER	10	10	10	10
FINISHES_AT_DAY	5	5	5	5
Peak Power (W)	11.10%	11.10%	11.10%	83899
Average Power Consumption (W)	6.82%	5.10%	6.28%	48461.4
Average Latency (sec)	2.18%	2.10%	2.28%	132472

Table 5.7: Percentage of savings of OA in Test Case C

Scheduling Algorithm	Min_HR	BP	LTF	OA
NUMBER_OF_CHASSIS	50	50	50	50
SERVERS PER CHASSIS	10	10	10	10
CORES PER SERVER	10	10	10	10
FINISHES_AT_DAY	5	5	5	5
Peak Power (W)	10.13%	10.13%	10.13%	49070.7
Average Power Consumption (W)	9.73%	6.48%	9.40%	31412.7
Average Latency (sec)	3.27%	2.41%	3.27%	219628

As we can see, the results confirm the hypothesis that OptimalAlg is able to significantly reduce energy consumption costs in a cloud data center. In fact, OptimalAlg simulated in a small data center of 150 servers is found to achieve an average of 6.48%, 9.4%, and 9.73% energy savings in comparison to BestPerformance, Low_temp_first, and MinHR repectively. We consider 0.12 USD to be the cost of 1 kWh. Therefore, OptimalAlg would achieve 320,964.41 USD, 465,596.52 USD, and 481,941.93 USD energy cost savings over BestPerformance, Low_temp_first, and MinHR respectively, while the total energy cost of the data center is 4,953,154.53 USD within a time frame of one year. In a medium size data center of 250 servers, approximately 649,518.45 USD, 799,799.19 USD, and 868,571.73 USD energy cost savings are realized in comparison to BestPerformance, Low_temp_first, and MinHR respectively when the total energy cost is 12,735,655.92 USD. Also, the simulation of a big data center of 500 servers proves energy cost savings of 1,613,025.19 USD, 1,731,629.98 USD, and 1,783,816.09 USD over BestPerformance, Low_temp_first, and MinHR respectively, while the total energy cost of the data center is 47,441,917.44 USD.

On the other hand, OptimalAlg leads to an average of 10% improvement in peak power in comparison to other algorithms in the different test cases. Lowering peak power of servers is important because it would extend the data center hardware life.

The curves in figure 5.1 illustrate the average supply cold temperature for different algorithms, in different data center sizes. It can be noticed that OptimalAlg consistently provides the highest supply cold temperature and therefore it has the smallest cooling energy of CRAC among all algorithms. This is due to the fact that CPU fans play a big role in our algorithm in eliminating a considerable amount of heat. CPU fans do not consume a lot of energy compared to the cooling energy consumed by CRAC. Our algorithm selects CPU fan speed optimally to reduce cooling energy and ultimately the total energy consumed by the data center.

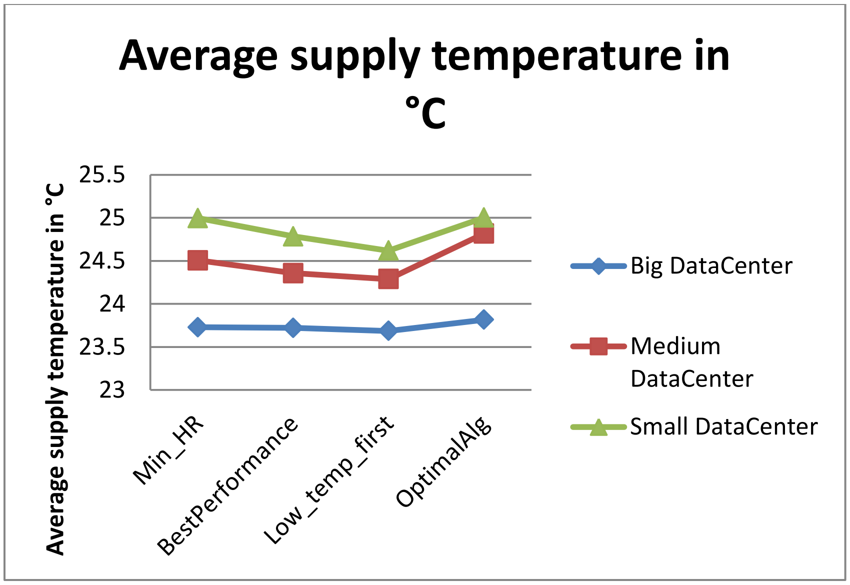

Figure 5.1. Average supply temperature in different data center sizes

PUE for the three different test cases are graphically depicted in Figure 5.2. The results show that OptimalAlg provides the smallest PUE which further verifies our hypothesis that our algorithm is the best performer.

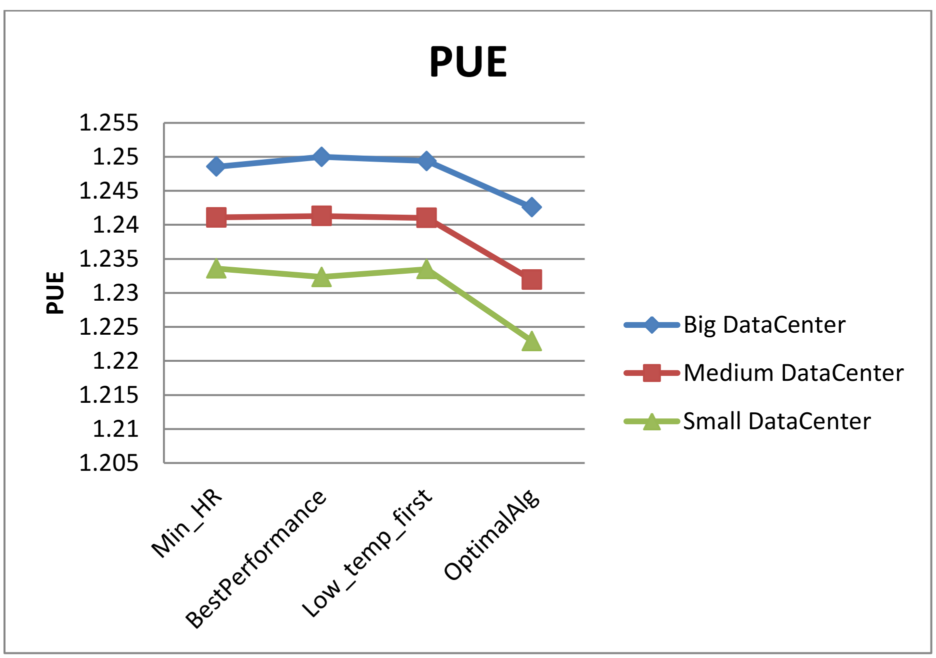

Figure 5.2. PUE in different data center sizes

Moreover, Tables 5.2, 5.3, and 5.4 show that OptimalAlg has the smallest standard deviation of inlet temperature, which means that our algorithm maintains a relatively homogeneous and stable thermal state at the different rack units in the data center. Also, OptimalAlg has the lowest latency in comparison with the three others, which indicates that it provides a better performance impact.

5.2.2. Setup 2: Workload utilization rates

The same simulation described in setup 1 was performed using a warehouse-scale data center containing 500 blade servers in 50 chassis, with 10 servers per chassis. We ran two different SWF files from different physical clusters as shown in Table 5.8 to simulate a medium and a small average utilization data center. The number of tasks differs to vary the data center utilization. However, we could not simulate high data center utilization rate since the input workload traces for SimWare are taken from real clusters and normally the average utilization of a data center does not exceed 30%.

Case	Cluster	Average Utilization	Number of jobs	Time Frame
D	Shared Hierarchical Academic Research Computing Network (SHARCNET) in Canada	25.12%	$1.2\text{x}10^6$	December 2005 January 2007
E	NASA Ames Research Center	2.53%	42,264	October 1993 December 2007

Table 5.8: Various data center utilization rates test cases

SimWare simulation results for test cases D and E are reported in Tables (5.9) and (5.10) respectively.

Table 5.9: SimWare Simulations Results for Test Case D

Scheduling Algorithm	Min_HR	BP	LTF	OA
NUMBER_OF_CHASSIS	50	50	50	50
SERVERS PER CHASSIS	10	10	10	10
CORES PER SERVER	10	10	10	10
FINISHES_AT_DAY	365	365	365	365
Average Supply Temperature (°C)	23.0796	22.9941	22.5371	23.4165
Peak Power (W)	193823	193823	193823	174914
Average Power Consumption (W)	105910	103333	105879	101625
PUE	1.24958	1.25265	1.26398	1.24411
Average Latency (sec)	41141.4	41050.2	41145.6	40985.1
Standard Deviation of Inlet T°	2.34	2.36	2.55	2.2

Table 5.10: SimWare Simulations Results for Test Case E

Scheduling Algorithm	Min_HR	BP	LTF	OA
NUMBER_OF_CHASSIS	50	50	50	50
SERVERS PER CHASSIS	10	10	10	10
CORES PER SERVER	10	10	10	10
FINISHES_AT_DAY	365	365	365	365
Average Supply Temperature (°C)	24.4375	24.420	24.429	24.430
Peak Power (W)	125071	124961	125062	124766
Average Power Consumption (W)	80646.7	80428.	80560.4	80277.
PUE	1.22048	1.2206	1.2205	1.2203
Average Latency (sec)	773.212	773.21	773.212	773.21
Standard Deviation of Inlet T°	1.87	1.89	1.9	1.87

The percentage of savings of OptimalAlg in comparison to BestPerformance, LowTempFirst, and MinHR in terms of peak power, average power consumption, and latency are summarized in Tables 5.11 and 5.12.

Table 5.11: Percentage of savings of OA in Test Case D

Scheduling Algorithm	Min_HR	BP	LTF	OA
NUMBER_OF_CHASSIS	50	50	50	50
SERVERS PER CHASSIS	10	10	10	10
CORES PER SERVER	10	10	10	10
FINISHES_AT_DAY	365	365	365	365
Peak Power (W)	11.80%	11.80%	11.80%	174914
Average Power Consumption (W)	6.04%	3.60%	6.01%	101625
Average Latency (sec)	2.10%	1.41%	2.10%	40985.1

Table 5.12: Percentage of savings of OA in Test Case E

Scheduling Algorithm	Min_HR	BP	LTF	OA
NUMBER_OF_CHASSIS	50	50	50	50
SERVERS PER CHASSIS	10	10	10	10
CORES PER SERVER	10	10	10	10
FINISHES_AT_DAY	365	365	365	365
Peak Power (W)	1.20%	1.10%	1.20%	124766
Average Power Consumption (W)	1.10%	0.80%	1.10%	80277.8
Average Latency (sec)	0%	0%	0%	773.212

Plots of average supply temperature and PUE are provided in figures 5.3 and 5.4 respectively. The analysis conducted in section 5.2.1 remains true for this setup. OptimalAlg has made significant energy consumption reduction with an improved performance impact. In fact, OptimalAlg is found to achieve the highest average energy savings and lowest peak power in comparison to minHr, BestPerformance, and lowtempfirst as shown in Tables 5.11 and 5.12. Also, in both test cases D and E, our proposed algorithm provides the lowest PUE, latency, and standard deviation, which indicates that OptimalAlg performs better than others in different utilization rates. Furthermore, we noticed that energy savings grow with the increase of utilization. This is reflected by the associated higher percentages in average power consumption in test case D over test case E.

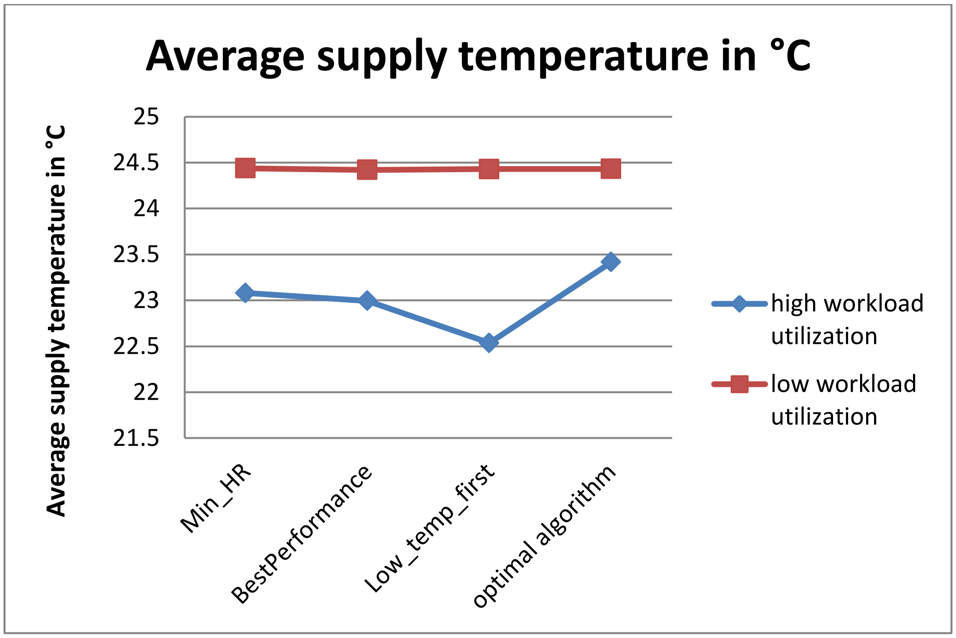

Figure 5.3. Average supply temperature in different average utilization rate

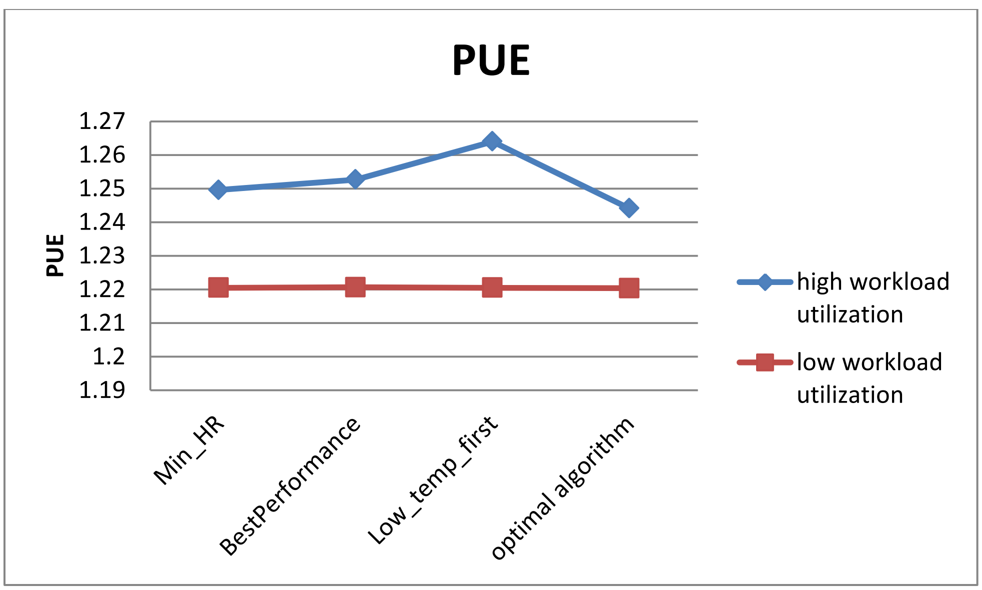

Figure 5.4. PUE in different average utilization rate

5.2.3. Further Data Analysis

Based on the observed simulation results, several conclusions can be made:

1. OptimalAlg significantly outperforms previous allocation algorithms, and produces better energy savings with an improved performance impact.
2. Less energy consumption is achieved using OptimalAlg and consequently less associated carbon dioxide is released into the atmosphere.
3. The performance time of OptimalAlg in SimWare is approximately the same compared to other algorithms. This is due to the fact that CPLEX solver used in the implementation is one of the fastest solvers available.
4. Our algorithm showed the highest supply temperature and consequently the lowest cooling energy. CPU fans were responsible of eliminating the excess heat. Therefore, the assumption that CPU fans play an important role in the thermal management system of a data center is met.

CHAPTER 6

CONCLUSION AND FUTURE WORK

As the technology progresses, the model of bringing computing resources remotely over the Internet will thrive. As a result, cloud data centers are estimated to grow. In this context, data center energy-efficient management is a critical problem with regard to both the huge operational costs and CO_2 discharges.

In this thesis, we presented a novel approach for the implementation of dynamic workload allocation by formulating and solving the optimization problem of minimizing the total data center energy consumption. The proposed algorithm returns the optimal supply cold temperature, task assignment efficiently on different servers, and the corresponding CPU fan velocity. The proposed algorithm resulted in energy savings with no performance compromise under different utilization of the data center. Simulation results showed an average of 6% energy savings for different utilizations and sizes of the data center when compared to three previous schemes. Our proposed algorithm will ensure energy provisioning, performance optimizing, and on-demand workload allocating.

Despite significant contributions of the current thesis in dynamic energy-efficiency data centers, there are some open research challenges that can be studied in order to further advance the area. The current proposed job placement algorithm investigated the mathematical optimization problem of energy consumption in data centers without including virtual machine consolidation. In such data centers, servers run at 10-50% of their capacity most of the time, which leads to extra costs on over-provisioning. Future work may extend the current scheme to include VM consolidation. It is essential to carefully design such an algorithm to deliver close to optimal solutions, as VM consolidation may lead to performance degradation. Another possible enhancement of the job placement algorithm is applying extra constraints on the job placement. Such constraints can be valuable when it is required to allocate jobs on a set of servers. These requirements can be enforced by the users due to security and/or privacy concerns.

.

REFERENCES

[1] S. Yeo, H. Lee. "SimWare: A Holistic Warehouse-Scale Computer Simulator". *Computer,* Vol. 45, Issue: 9, pp. 48- 55, Sept. 2012.

[2] Schneider Electric (2013). *StruxureWareTMData CenterOperation*. [Online]. Available: http://www.apcmedia.com/salestools/RMCR-8N2PLE_R3_EN.pdf.

[3] D. Kliazovich, P. Bouvry, Y. Audzevich, and S. Ullah Khan. "GreenCloud: A Packet-level Simulator of Energy-aware Cloud Computing Data Centers", in *IEEE Global Telecommunications Conference (GLOBECOM 2010),* 2010, pp.1-5.

[4] R. N. Calheiros et al. "CloudSim: a toolkit for modeling and simulation of cloud computing environments and evaluation of resource provisioning algorithms". *Software: Practice and Experience,* Vol. 41, Issue: 1, pp.23- 50, Jan. 2011.

[5] Y. Shi, X. Jiang, and K. Ye. "An Energy-Efficient Scheme for Cloud Resource Provisioning Based on CloudSim", in *2011 IEEE International Conference on Cluster Computing,* 2011, pp. 595- 599.

[6] Lanamark Suite. [Online]. Available: http://www.lanamark.com/.

[7] VMware Inc. VMware Capacity Planner. [Online]. Available: http://www.vmware.com/products/capacityplanner/

[8] IBM WebSphere CloudBurst. [Online]. Available: http://www-01.ibm.com/software/webservers/cloudburst/.

[9] Novell PlateSpin Recon. [Online]. Available: http://www.novell.com/products/recon/.

[10] VMware Inc. VMWare vCenter CapacityIQ. [Online]. Available: http://www.vmware.com/products/vcentercapacityiq/

[11] M. Mazzucco, D. Dyachuky, and R. Deters. "Maximizing Cloud Providers Revenues via Energy Aware Allocation Policies", in *2010 IEEE International Conference on Cloud Computing (CLOUD*), 2010, pp. 131-138.

[12] Anton Beloglazov, Rajkumar Buyya. "Adaptive Threshold-Based Approach for Energy-Efficient Consolidation of Virtual Machines in Cloud Data Centers", *in Proceedings of the 8th International Workshop on Middleware for Grids, Clouds and e-Science*, ACM, 2010.

[13] W. Xiaoli, L. Zhanghui. “An energy-aware VMs placement algorithm in Cloud Computing environment”, in *2012 Second International Conference on Intelligent System Design and Engineering Application (ISDEA),*2012, pp. 627- 630

[14] H. Viswanathan, E.K. Lee, I. Rodero, D. Pompili, M. Parashar, and M. Gamell. “Energy-Aware Application-Centric VM Allocation for HPC Workloads”, in *2011 IEEE International Symposium on Parallel and Distributed Processing Workshops and Phd Forum (IPDPSW),* 2011, pp. 890- 897.

[15] Marko Hoyer, Kiril Schröder, and Wolfgang Nebel. “Statistical static capacity management in virtualized data centers supporting fine grained qos specification”, in *Proceedings of the 1st International Conference on Energy-Efficient Computing and Networking, e-Energy’10,* NY, USA. ACM, 2010, pp. 51–60

[16] Sadeka Islam, Jacky Keung, Kevin Lee, and Anna Liu. “Empirical prediction models for adaptive resource provisioning in the cloud”. in *Future Generation Computer Systems, Elsevier,* Vol. 28, pp. 155–162, 2012.

[17] D. Kliazovich, P. Bouvry, and S. Ullah Khan. “DENS: Data Center Energy-Efficient Network-Aware Scheduling”, in *Green Computing and Communications (GreenCom), 2010 IEEE/ACM Int'l Conference on & Int'l Conference on Cyber, Physical and Social Computing (CPSCom),* 2010, pp. *69- 75.*

[18] M. Wang, X. Meng, and L. Zhang. “Consolidating Virtual Machines with Dynamic Bandwidth Demand in Data Centers”, in *2011 Proceedings IEEE INFOCOM*, 2011, pp. 71- 75.

[19] K. H. Kim, A. Beloglazov, and R. Buyya. “Power-aware Provisioning of Cloud Resources for Real-time Services*”,* in *Proceedings of the 7th International Workshop on Middleware for Grids, Clouds and e-Science,* 2009.

[20] R, Neil. "Calculating total cooling requirements for data centers," in American Power Conversion, white paper 25, 2007.

[21] J. Mouton. Enabling the vision: Leading the architecture of the future, in Keynote speech, Server Blade Summit, 2004.

[22] Y. Chen, A. Das, W. Qin, A. Sivasubramaniam, Q. Wang, and N. Gautam, “Managing server energy and operational costs in hosting centers,” in *SIGMETRICS Perform. Eval. Rev*., vol. 33, no. 1, pp. 303–314, 2005.

[23] E. M. Elnozahy, M. Kistler, and R. Rajamony, “Energy-efficient server clusters,” in *Proceedings of the 2nd Workshop on Power-Aware Computing Systems*, 2002, pp. 179–196.

[24] J. Moore, J. Chase, P. Ranganathan, R. Sharma, “Making Scheduling ‘Cool’: Temperature-Aware Workload Placement in Data Centers”, in *USENIX Annual Technical Conference*, April 2005.

[25] T. Qinghui, S. Gupta, G. Varsamopoulos, "Energy-efficient thermal-aware task scheduling for homogeneous high-performance computing data centers: A cyber-physical approach," in *Parallel and Distributed Systems, IEEE Transactions,* 2008, pp. 1458-1472.

[26] Beitelmal, M. H., Wang, Z., Felix, C., Bash C. E., Hoover, C.McReynolds, A. “Local cooling control of data centers with adaptive vent tiles”. in *the Proceedings of the ASME/Pacific Rim Technical Conference and Exhibition on Packaging and Integration of Electronic and Photonic Systems, MEMS and NEMS*, 2009.

[27] R. Zhou, Z. Wang, C. E. Bash, A. McReynolds, C. Hoover, R. Shih, N. Kumari, and R. K. Sharma, ” A Holistic and Optimal Approach for Data Center Cooling Management”, in *proc. Of the American Control Conference*, San Francisco, USA, June 29 - July 01, 2011.

[28] R. Zhou and Z. Wang , “Modeling And Control for Cooling Management of Data Centers with Hot Aisle Containment”, in *proc. of ASME 2011 International Mechanical Engineering Congress & Exposition IMECE 2011,* Denver, Colorado, USA, Nov. 11-17, 2011.

[29] O. VanGeet, *“Best Practices Guide for Energy-Efficient Data Center Design*”, U.S. Department of Energy Federal Energy Management Program, 2011.

[30] Bash, Cullen, and George Forman, "Cool job allocation: measuring the power savings of placing jobs at cooling-efficient locations in the data center," in *2007 USENIX Annual Technical Conference on Proceedings of the USENIX Annual Technical Conference*. *USENIX Association*, 2007.

[31] J. Moore, J. Chase, and P. Ranganathan, “Weatherman: Automated, online, and predictive thermal mapping and management for data centers,” in *3rd IEEE Int’l Conf. Autonomic Computing*, June 2006.

[32] Q. Tang, S. K. S. Gupta, and G. Varsamopoulos, “Energy-efficient thermal-aware task scheduling for homogeneous high-performance computing data centers: A cyber-physical approach,”in *IEEE TPDS*, vol. 19, no. 11, pp. 1458–1472, 2008.

[33] A. Qureshi, R. Weber, H. Balakrishnan, J. Guttag, and B. Maggs, “Cutting the Electric Bill for Internet-Scale Systems,” in *ACM SIGCOMM*, Barcelona, Spain, August 2009.

[34] Parallel Workloads Archive [Online]. Available: http://www.cs.huji.ac.il/labs/parallel/workload/

[35] U. E. P. Agency, “Data center report to congress,” U.S. Environmental Protection Agency, Tech. Rep., 2007.

[36] “Growth in data center electricity use 2005 to 2010,” Analytics Press, Tech. Rep., 2011.

[37] Gartner, Inc., “Gartner estimates ICT industry accounts for 2 percent of global CO2 emissions,” 2007, (accessed on 17/01/2013).

[38] J. G. Koomey, “Estimating total power consumption by servers in the US and the world,” Lawrence Berkeley National Laboratory, Tech. Rep., 2007.

[39] P. Brenner, D. Thain, and D. Latimer, “Grid Heating Clusters: Transforming Cooling Constraints into Thermal Benefits”, 2009.

[40] www.cs.huji.ac.il/labs/parallel/workload/swf.html. [Online]. Available: www.cs.huji.ac.il/labs/parallel/workload/swf.html

[41] J. Moore, J. Chase, P. Ranganathan, and R. Sharma, “Making scheduling "cool": temperature-aware workload placement in data centers,” in *ATEC ’05: Proceedings of the annual conference on USENIX Annual Technical Conference*. Berkeley, CA, USA: USENIX Association, 2005, pp. 5–5.

[42] L.A. Barroso and U. Hölzle, “The Case for Energy-Proportional Computing,” *Computer*, Dec. 2007, pp. 33-37.

[43] T. Mukherjee, A. Banerjee, G. Varsamopoulos, S. K. S. Gupta, and S. Rungta, “Spatio-temporal thermal-aware job scheduling to minimize energy consumption in virtualized heterogeneous data centers,” *Computer Networks*, June 2009.

[44] Q. Tang, T. Mukherjee, S. K. S. Gupta, and P. Cayton, “Sensor-based fast thermal evaluation model for energy efficient high-performance datacenters,” in *ICISIP 2006*, Dec 2006.

[45] ASHRAE Technical Committee 9.9, “Thermal Guidelines for Data Processing Environments—Expanded Data Center Classes and Usage Guidance”, white paper, Am. Soc. Heating, Refrigerating and Air-Conditioning Engineers, 2011