

Thermal-Aware Virtual Machine Allocation for Heterogeneous Cloud Data Centers

Abbas Akbari ^{1,*} , Ahmad Khonsari ¹ and Seyed Mohammad Ghoreyshi ² 

¹ School of Electrical and Computer Engineering, University of Tehran, Tehran 146899-5513, Iran; ak@ipm.ir

² School of Mathematical Sciences, University of Southampton, Southampton SO14 0AB, UK; S.M.Ghoreyshi@soton.ac.uk

* Correspondence: a_akbari@ut.ac.ir

Received: 7 May 2020; Accepted: 1 June 2020; Published: 4 June 2020



Abstract: In recent years, a large and growing body of literature has addressed the energy-efficient resource management problem in data centers. Due to the fact that cooling costs still remain the major portion of the total data center energy cost, thermal-aware resource management techniques have been employed to make additional energy savings. In this paper, we formulate the problem of minimizing the total energy consumption of a heterogeneous data center (MITEC) as a non-linear integer optimization problem. We consider both computing and cooling energy consumption and provide a thermal-aware Virtual Machine (VM) allocation heuristic based on the genetic algorithm. Experimental results show that, using the proposed formulation, up to 30% energy saving is achieved compared to thermal-aware greedy algorithms and power-aware VM allocation heuristics.

Keywords: thermal-aware; energy-efficiency; Virtual Machine; heterogeneous data center; Cloud computing

1. Introduction

Cloud computing has emerged as a new and popular computing paradigm for providing on-demand hosted services over the Internet [1,2]. The computing and storage capacity of data centers have increased rapidly as the demand for Internet-based services has increased [3–6]. The continuous increase of energy consumption and power density is one of the serious side effects of such rapid growth in the data center facilities and imposes a significant cost on Cloud data centers that reduces the marginal profit of Cloud providers and leads to a growth in the rate of carbon dioxide emissions. Thus, energy efficiency has become a major economic and environmental issue in data centers, which has attracted the attention of researchers and industries [7–9].

A considerable amount of literature has been published on energy-efficient computing in data centers [10,11]. One of the most significant parts of the energy consumption in data centers appears to be the energy consumed in the cooling system [12,13]. Recent cooling technologies, such as in-rack, in-row, and container-based cooling, have been proposed to improve data centers' energy efficiency. However, the majority of data centers are still restricted to older cooling technologies, such as chilled-water air cooling [14]. Thus, it seems that there is still room for further improvement in energy efficiency by applying resource management techniques [15,16].

To the best of our knowledge, most works into minimizing cooling energy consumption have proposed optimization algorithms to maximize the temperature of the cooled air supplied by data centers' Computer Room Air Conditioning (CRAC) unit [17]. The maximum allowable temperature is determined by the lowest red-line temperature of the installed servers in a data center. More efficient cooling processes allow higher temperatures for supplied cooled air without exceeding the red-line temperature in each server for a given workload. These temperature-based optimization methods

reduce the recirculating heat in the data center [18]. The recirculation of exiting hot air from devices' air outlets back into their air inlets increases inlet temperatures and consequently decreases the maximum allowable temperature of the supplied cooled air.

Several studies have parametrized heat recirculation and heat flow fields based on the temperature distribution in the data center [19,20]. The most detailed thermal model of a flow field to predict temperature distribution is a Computational Fluid Dynamics (CFD) model [21]. Similarly, thermal maps of data centers can predict the temperature distribution for a given CPU utilization pattern [22]. However, these models are not suitable for on-line resource management due to their high computational costs. Thermal models with lower complexity, such as Supply Heat Index (SHI), Recirculation Heat Index (RHI) [23], and Heat Recirculation Factor (HRF) [23], have been proposed in the literature.

In addition to the above-mentioned traditional techniques, recent advances in virtualization methodology have made it possible for designers to reduce the idle power costs of modern data centers by improving system manageability [24–26]. A Virtual Machine (VM) is a software abstraction of an actual physical computing system which is machine-independent and easily portable [27]. Virtual Machines have many benefits; for example, they keep the users separate from the underlying infrastructure and isolate the workloads that share common facilities. A VM benefits its user by improving performance and scalability and reducing hardware and software requirements, as well as manageability and fault tolerance. Moreover, it improves resource utilization, reduces the energy consumption through the consolidation of multiple workloads on fewer servers, and enables live migration [28–30]. It can also provide load balancing between servers by enabling Virtual Machine migration to eliminate thermal hotspots in data centers.

In this paper, we formulate the problem of thermal-aware Virtual Machine allocation in a heterogeneous data center as a non-linear integer optimization problem. To the best of our knowledge, our proposed formulation is the first to mathematically formulate the problem of how to allocate Virtual Machines in a thermal-aware manner among the computing nodes in order to minimize the total energy consumption of the data center (the formulated problem is named MITEC). We provide a novel heuristic approach to solve the MITEC problem based on a genetic algorithm. Furthermore, we compare the performance of the proposed heuristic with thermal-aware greedy algorithms and power-aware VM allocation heuristics.

The main contributions of this paper are the following:

- Creating a formal definition of optimal thermal-aware VM allocation by considering both computing and cooling energy consumption and providing a novel heuristic based on a genetic algorithm to obtain a near-optimal solution in less computing time;
- Designing a trade-off between the power-aware consolidation techniques and thermal-aware load balancing approaches to obtain higher energy savings in Cloud data centers;
- An extensive simulation-based evaluation and performance analysis of the proposed algorithm.

The remainder of this paper is organized as follows: we address the related work in Section 2. The general physical description of the data center and the power and the total energy consumption models are described in Section 3. Section 4 presents the thermal-aware VM allocation problem formulation, while a genetic heuristic approach is provided in Section 5. Simulation results are provided in Section 6. Finally, Section 7 concludes the paper.

2. Related Work

Recently, there has been a large volume of published studies carried out on energy-efficient resource management in data centers [7–9,11]. One of the most significant discussions aimed at minimizing the total power consumption in data centers is the Dynamic Voltage/Frequency Scaling (DVFS) of servers [31–33]. Pakbaznia et al. [34] formulate the total data center power consumption

and propose an efficient algorithm which reduces the server power by appropriately allocating tasks to servers and choosing the optimal voltage-frequency level for servers. The algorithm also considers the cooling power and chooses the proper supplied cold air temperature value in order to minimize the cooling cost. Dynamically “right-sizing” the data center [35] by turning off inactive servers is another technique which can be applied to achieve more energy savings. In [36], the authors consider the problem of server provisioning and DVFS control at hosting centers. The paper formulates this problem and presents three main techniques based on steady-state queuing analysis, feedback control theory, and a hybrid between the two.

Due to the fact that a considerable reduction in the cooling energy requirements of data centers can be achieved by thermal-aware resource management, several studies have been conducted to explore different techniques in this area. As explained by [5,37], common designs for cooling systems for data centers result in an inefficient and limiting phenomenon called heat recirculation. Heat recirculation is caused by the mixing of the two hot and cold air flows and prevents ideal cooling efficiency. In a study by HP Labs and Duke University, Sharma et al. [23] introduced two scalar and dimensionless parameters, the Supply Heat Index (SHI) and Recirculation Heat Index (RHI), as the metrics that can be used to evaluate the thermal efficiency of a data center by assessing of heat recirculation.

The study presented in [38] includes a detailed thermo-fluid analysis of a typical data center to resolve temperature distribution for a given steady-state power density. This paper uses steady-state and transient CFD techniques and proposes a method describing the provisioning of data center cooling resources based on CFD simulation. Studies prior to [39] explored efficient methods of provisioning cooling resources for data centers [40–42]. However, Moore et al. propose a thermal-aware workload placement algorithm based on the minimization of the cooling energy needed at the facilities level [39]. They first introduce a parameter to assess the heat recirculation phenomenon, called the heat recirculation factor (HRF), which resulted from thermodynamical considerations. Then, they propose Zone-Based Discretization (ZBD) and Minimize-Heat-Recirculation (MinHR) algorithms for workload placement based on the HRF metric to minimize the effects of heat recirculation on cooling inefficiency.

Tang et al. [43] introduce an abstract heat flow model to characterize the heat recirculation phenomenon in data centers. This model can also predict the inlet air temperature of each chassis based on the supplied cooled air temperature and workload distribution faster than CFD simulations and is fairly accurate. The main idea in this work is profiling a matrix, called the cross interference matrix, using the measurement or prediction of the inlet and outlet temperature of each chassis by on-board and ambient sensors or via CFD simulations. In [44], Tang et al. developed a new, low-complexity, linear heat recirculation model based on the cross interference matrix introduced in [43] to minimize the peak inlet temperature within a data center. They presented two heuristic methods, called XInt-GA (genetic algorithm) and XInt-SQP (Sequential Quadratic Programming) algorithms, to solve the optimization problem in a reasonable time.

Further optimization strategies have been developed based on abstract heat recirculation models in recent years. In [45], the authors proposed a Power and Thermal Management (PTM) engine which determines the number and placement of ON servers and adjusts the supplied cold air temperature simultaneously. Using this method, the total power consumed by servers and CRACs has been minimized. Abbasi et al. [46] developed a method for a thermal-aware dynamic server set provisioning and introduced a thermal-aware workload distribution algorithm among the active servers.

3. System Model

In this section, we provide the general physical description of the data center and present the power and the total energy consumption models.

3.1. Layout of Data Center

The data center room generally contains several rows of racks; each rack consists of few chassis in which a number of blade servers are assembled. Current data centers are designed with hot/cold

aisles, as shown in Figure 1. Each row is placed between a hot and a cold aisle. Chilled air for room cooling is supplied by the CRAC unit through the floor tiles. Each server sucks the chilled air into the rack to cool the servers, and the hot air is blown out by the fans into hot aisles formed by the data center flow pattern. Then, the accumulated hot air is extracted by the CRAC vents on the ceiling. The important duty of the CRAC unit is to keep the inlet temperature of each chassis under the red-line temperature (T_{red}), which otherwise would degrade performance or harm the data center facilities.

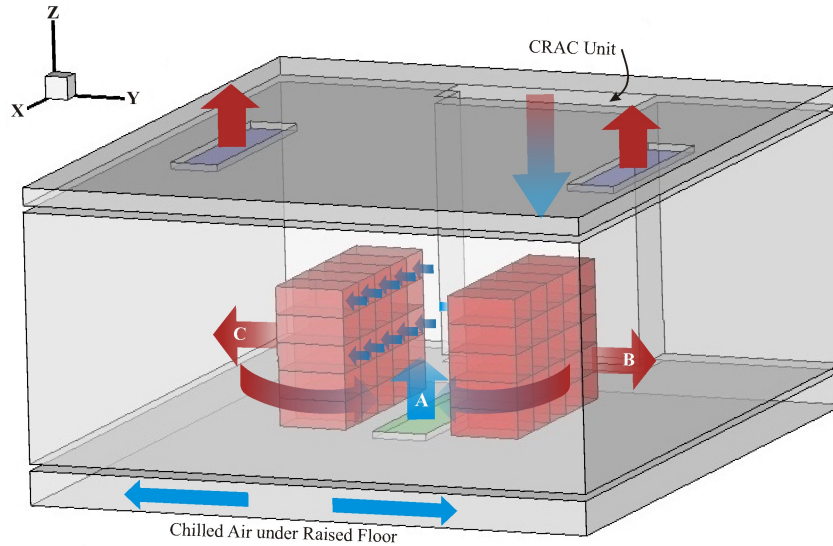


Figure 1. Data center layout.

3.2. Heat Recirculation Phenomena

The air circulation system in a data center provides coolant flow for chassis i at its inlet with temperature T_{in}^i and transfers the exhausted hot air from its outlet with temperature T_{out}^i . After extracting server-generated heat, the CRAC supplies chilled air of temperature T_{sup} via floor tiles. The key optimization idea is to isolate the cold aisles as much as possible, reducing heat leakage into coolant cold aisles. Obviously, such leakage increases the differences between T_{in}^i and T_{sup} and consequently decreases the maximum T_{sup} that can be supplied given a certain workload pattern.

In order to isolate the cold aisles, the thermal-aware workload placement formulation can be used to minimize the heat recirculation in a certain flow pattern. In this formulation, the data center can be considered an open system with a cold air feed (temperature T_{sup}) that is affected by the chassis, which are heat sources. T_{in}^i and T_{out}^i denote inlet and outlet temperatures of i th chassis. The enthalpy change in each chassis is stated by $K_i(T_{out}^i - T_{in}^i)$ in which $K_i = \dot{m}_i c_p$ where $\dot{m}_i$ is mass flow rate of chassis i and c_p is the specific heat capacity of air. A workload distribution is described by vector $\vec{P}$, which quantifies the power dissipated by servers in each chassis. From $\vec{P}$, we determine the corresponding inlet and outlet temperatures $\vec{T}_{in}$ and $\vec{T}_{out}$. As explained by [43], the contribution of the flow pattern on heat recirculation can be characterized by a cross interference coefficient (CIC) matrix denoted by $\mathbf{A}$. In fact, $\mathbf{A}$ is a matrix representation of the heating contribution of all chassis outlet flows on each others' inlet temperatures in a certain flow pattern. Using the CIC matrix, the inlet temperatures of each chassis can be calculated as follows:

$$\vec{T}_{in} = \vec{T}_{sup} + \mathbf{D}\vec{P} \quad (1)$$

in which the heat distribution matrix, $\mathbf{D}$ is defined by

$$\mathbf{D} \equiv [(\mathbf{K} - \mathbf{A}^T \mathbf{K})^{-1} - \mathbf{K}^{-1}]. \quad (2)$$

where $\vec{T}_{in}$ and $\vec{T}_{sup}$ are the corresponding inlet temperature and the cold air supply vectors. The former contains T_{in}^i for each chassis. The latter is a vector with identical components equal to the cool air temperature supplied by the CRAC unit.

3.3. Preliminaries

3.3.1. Cooling Efficiency of the CRAC Unit

Traditionally, the energy efficiency of a cooling system is characterized by the coefficient of performance (CoP). The CoP is defined as the ratio of the amount of heat that is removed by the cooling system (Q) to the total amount of energy (E) that is consumed in the cooling system to remove that amount of heat from the system:

$$\text{CoP} = \frac{Q}{E}. \quad (3)$$

In the current study, the proposed coefficient is based on a typical water-chilled CRAC unit utilized in an HP Utility data center [39]. In this model, the CoP is a function of T_{sup} :

$$\text{CoP}(T_{sup}) = (0.0068T_{sup}^2 + 0.0008T_{sup} + 0.458) \quad (4)$$

3.3.2. VM Allocation and Power Model

We consider a large computing facility, consisting of N chassis. Each chassis $i \in \{1, \dots, N\}$ contains a set of servers which provide the total computing capacity of M_{cap}^i million instructions per second (MIPS). We denote the VM set as $V = \{VM_j(M_{req}^j, L_j) | j = 1, \dots, M\}$, where j th VM requests M_{req}^j million instructions per second and a task of L_j million instructions (MIs) is executed on j th VM. We consider that each VM executes a single task with a length of L_j . Thus, the task execution time is equal to the task instruction length (L_j) divided by the VM processing capacity (M_{req}^j). The total power consumption of the i th chassis with a CPU utilization of u_i can be calculated as [47]

$$P_{comp}^i = \alpha_i + \beta_i \cdot u_i, \quad (5)$$

where α_i and β_i represent the idle power consumption of the i th chassis and the extra power consumption from CPU utilization, respectively. The computing energy consumption of i th chassis can be defined as

$$E_{comp}^i = \int_{t_0}^{t_1} P_{comp}^i(u_i(t))dt = \alpha_i \cdot t_{max}^i + \beta_i \cdot \sum_{j \in V_i} (u_{ij} \cdot t_j), \quad (6)$$

where t_{max}^i , V_i , u_{ij} and t_j represent the maximum time that i th chassis is active, the set of VMs that are allocated to i th chassis, the utilization of j th VM on i th chassis (i.e., M_{req}^j / M_{cap}^i) and the fraction of time that j th VM is active (i.e., L_j / M_{req}^j), respectively.

3.3.3. Total Energy Consumption

The total energy consumption of a data center is defined as the sum of chassis energy, E_{comp} , and the cooling energy of the CRAC unit, E_{CRAC} [47]. The cooling energy consumption of the CRAC unit can be computed as $E_{CRAC} = E_{comp} / \text{CoP}(T_{sup})$. Thus, the total energy consumption of a data center can be written as

$$E_{total} = E_{comp} + E_{CRAC} = \left(1 + \frac{1}{\text{CoP}(T_{sup})}\right) \sum_{i=1}^N E_{comp}^i. \quad (7)$$

In order to prevent performance degradation or damage to the data center facilities, the CRAC unit must keep the inlet temperature of each chassis under the red-line temperature T_{red} . Thus, considering Equation (1) we have

$$\max_{1 \leq i \leq N} \{\vec{T}_{in}\} \leq T_{red}, \quad (8)$$

$$\Rightarrow T_{sup} + \max_{1 \leq i \leq N} \{\mathbf{D}\vec{P}\} \leq T_{red}, \quad (9)$$

$$\Rightarrow T_{sup} \leq T_{red} - \max_{1 \leq i \leq N} \{\mathbf{D}\vec{P}\}. \quad (10)$$

Thus, the maximum supply temperature from the CRAC unit is $T_{sup} = T_{red} - \max_{1 \leq i \leq N} \{\mathbf{D}\vec{P}\}$. Now, we can rewrite Equation (7) as

$$\begin{aligned} E_{total} &= E_{comp} + E_{CRAC} \\ &= \left(1 + \frac{1}{CoP(T_{red} - \max_{1 \leq i \leq N} \{\mathbf{D}\vec{P}\})}\right) \sum_{i=1}^N E_{comp}^i. \end{aligned} \quad (11)$$

4. Problem Formulation

In this subsection, we provide the non-linear integer formulation for the problem of thermal-aware VM allocation. We define an integer variable x_{ij} , which is set to 1 if the j th VM is allocated to i th chassis. Then, we define another integer variable y_i , which indicates whether the i th chassis is ON or OFF as follows:

$$y_i = \begin{cases} 0 & \text{if } x_{ij} = 0, \forall j \in \{1, \dots, M\} \\ 1 & \text{otherwise.} \end{cases} \quad (12)$$

The computing power consumption of the i th chassis can be modeled as

$$P_{comp}^i = \alpha_i \cdot y_i + \beta_i \cdot \sum_{j=1}^M x_{ij} \cdot u_{ij}. \quad (13)$$

We define $\vec{P}_{comp} = [P_{comp}^1, \dots, P_{comp}^N]^T$ as a column vector, representing the computing power consumption of each chassis and $t_{max}^i = \max_{1 \leq j \leq M} \{x_{ij} \cdot t_j\}$ as the maximum time that i th chassis is active. The mathematical optimization problem is defined as

$$\begin{aligned} \min. & \left(1 + \frac{1}{CoP(T_{red} - \max_{1 \leq i \leq N} \{\mathbf{D}\vec{P}_{comp}\})}\right) \\ & \times \sum_{i=1}^N (\alpha_i \cdot t_{max}^i + \beta_i \cdot \sum_{j=1}^M (x_{ij} \cdot u_{ij} \cdot t_j)), \end{aligned} \quad (14)$$

$$\text{subject to } \forall j: \quad \sum_{i=1}^N x_{ij} = 1, \quad (15)$$

$$\forall i: \quad \sum_{j=1}^M x_{ij} \cdot M_{req}^j \leq M_{cap}^i, \quad (16)$$

$$\forall i: \quad y_i \leq \sum_{j=1}^M x_{ij} \leq C \cdot y_i, \quad (17)$$

$$\forall i: \quad y_i \in \{0, 1\}, \quad (18)$$

$$\forall i, j: \quad x_{ij} \in \{0, 1\}. \quad (19)$$

The objective function (14) tries to minimize the total energy consumption of a data center including the computing and cooling energy consumption. In practice, there is a trade-off between consolidating VMs on fewer servers to save the idle power consumption and the load balancing of VMs

on a higher number of servers to avoid thermal hotspot creation. Our objective function accurately has taken these conflicting objectives into account.

The first constraint (15) ensures that each VM must be allocated to exactly one chassis. The second constraint (16) states that the requested MIPS of all VMs allocated to i th chassis must be less than or equal to the capacity of the chassis. In the third constraint (17), C is a significantly large number (that is, at least as large as the number of VMs). If for the i th chassis, $\sum_{j=1}^M x_{ij}$ equals 0 (i.e., no VM is allocated to i th chassis), then (16) implies that y_i must be 0. If $\sum_{j=1}^M x_{ij}$ is bigger than 0 for the i th chassis, then (16) states that y_i must be 1. The proposed non-linear integer formulation can obtain optimal VM placement solutions; however, it is an NP-hard problem and impractical for large-sized data centers. We therefore use a heuristic approach offering a comparably fast running time and still yielding near-optimal solutions.

5. The MITEC-GA Algorithm

In this section, we provide a genetic algorithm (MITEC-GA) to solve the VM allocation problem formulated in the previous section. Our proposed heuristic uses a vector of integer numbers, c , in order to implement the chromosome encoding. Each chromosome c represents a solution for the VM allocation problem. The allocation of j th VM to the i th chassis can be expressed by $c_j = i$. Thus, we have the following:

$$1 \leq c_j \leq N, \forall j = 1, \dots, M. \quad (20)$$

Let us define the available MIPS of i th chassis as

$$r_i = M_{cap}^i - \sum_{j:c_j=i} M_{req}^j. \quad (21)$$

A chromosome is a feasible solution for the problem if r_i is greater than or equal to 0 for all chassis; that is, no allocation exceeds chassis capacity.

Algorithm 1 details the proposed heuristic. The MITEC-GA begins by generating a pool of random population and calculating the fitness function for each candidate solution (lines 2–5). The algorithm uses E_{total} in Equation (7) as the fitness function. Our genetic algorithm runs in an iterative manner and repeats the following steps in each iteration:

- Selection: The roulette-wheel selection method is used to randomly select two parents from current population;
- Crossover: The new parents are generated by mixing selected chromosomes together and obtaining new allocations for some randomly selected VMs (lines 15–24);
- Mutation: The new solutions are formed by changing the new parents in a random way (lines 25–39).

We define a weighting factor w_j for j th VM in a candidate solution c which is directly proportional to M_{req}^j, β_{c_j} while inversely proportional to r_{c_j} . The higher the value of w_j , the more likely the allocation of j th VM changes in the solution c . We define w_j as

$$w_j = \frac{M_{req}^j \times \beta_{c_j}}{r_{c_j} + \epsilon}. \quad (22)$$

This weighting factor represents the importance of dynamic energy consumption. When the remaining capacity of a chassis decreases, the utilization of the chassis increases; as a consequence, the dynamic energy consumption of the chassis increases. Similarly, when the requested MIPS of a VM is large, it means that this VM affects the dynamic energy consumption of the chassis more than other VMs. Thus, a large VM has a higher chance to be placed on another chassis resulting in higher load balancing and lower consolidation, which can avoid hot-spot thermal issues and consequently reduce

cooling energy costs. In order to mutate chromosomes, the MITEC-GA first sorts VMs according to their weighting factor in descending order. The algorithm then substitutes the allocation of randomly selected VMs with lower weighting factors for the allocation of VMs with higher weighting factors.

Algorithm 1 MITEC-GA

```

1: procedure GENETIC
2:   for  $i = 1$  to  $popSize$  do
3:      $population[i] \leftarrow$  select a random feasible solution
4:      $fitness[i] \leftarrow$  calculate the fitness function
5:    $CurPop \leftarrow population$ 
6:   for  $i = 1$  to  $num\_of\_iter$  do
7:      $selParents \leftarrow$  select two solutions from  $CurPop$ 
8:     using roulette-wheel selection
9:      $crossParents \leftarrow$  call CROSSOVER( $selParents$ )
10:     $mutateSol \leftarrow$  call MUTATE( $crossParents$ )
11:    Apply the fitness function on  $mutateSol$ .
12:     $CurPop \leftarrow$  select  $n$  solutions from  $CurPop$ 
13:    and  $mutateSol$  with low energy consumption
14:   $FinalSolution \leftarrow$  the solution within  $CurPop$  with best fitness
15: function CROSSOVER( $selParents S_1$  and  $S_2$ )
16:  Select two random numbers, ( $rand1 < rand2$ )
17:  for  $i = rand1$  to  $rand2$  do
18:     $c_1, c_2 \leftarrow S_1[i], S_2[i]$ 
19:     $m_{c_1}^a, m_{c_2}^a \leftarrow$  available MIPS in chassis's  $c_1$  and  $c_2$ 
20:    if  $m_{c_1}^a \geq M_{req}^i$  then
21:      Substitute  $c_1$  for  $c_2$  in  $S_2$ 
22:    if  $m_{c_2}^a \geq M_{req}^i$  then
23:      Substitute  $c_2$  for  $c_1$  in  $S_1$ 
24:  return  $S_1$  and  $S_2$ 
25: function MUTATE( $crossParents S_1$  and  $S_2$ )
26:  for all  $VM_j$  do ▷ Calculate weighting factors
27:     $w_1[j] \leftarrow (M_{req}^j \times \beta_{S_1[j]}) / (r_{S_1[j]} + \epsilon)$ 
28:  Sort  $w_1$  in descending order
29:  for all  $VM_j$  in the first half of  $w_1$  do
30:     $selPart \leftarrow$  set of all VMs in the last tenth of  $w_1$ 
31:    repeat
32:       $sel\_VM \leftarrow$  select a VM randomly from  $selPart$ 
33:      if  $M_{req}^{VM_j} \geq M_{cap}^{S_1[sel\_VM]}$  then
34:        Substitute  $S_1[sel\_VM]$  for  $S_1[VM_j]$  in  $S_1$ 
35:        break
36:      else
37:        Increment the  $selPart$ 's space by one
38:    until  $selPart$ 's space reach to  $VM_j$ 
39:  Do the same things for  $w_2$  and  $S_2$ 

```

6. Simulation Results

In this section, the details of our simulation study and the performance results are presented.

6.1. Simulation Setup

In order to evaluate the the proposed scheme, CloudSim is modified and used to simulate thermal-aware VM allocation policies. We consider that a data center consists of two rows of five racks, each containing 50 chassis with 10 servers. The servers are HP D1080 and D2050. There are 35 chassis of D1080 and 15 chassis of D2050 servers. The D1080 server has a power consumption of 2020 W when idle and 2520 W at 100% CPU utilization. Similarly, the D2050 server has a power consumption of 1590 W at the idle state and 2490 W at 100% CPU utilization. The CRAC supplies chilled air at 20 °C into the data center through the raised floor vents (Figure 1). Thermal intervals of 15 min (that is, sufficient time for data center thermal stability) are added to the simulator. In each interval, we calculate the maximum supply temperature from the CRAC unit, T_{sup} , based on Equation (10) and measure the energy consumption based on Equation (7).

In order to characterize the workload model, we use Amazon Elastic Compute Cloud (EC2) Standard Instances: Small Instance with one EC2 compute unit (1000 MIPS); Medium Instance with two EC2 compute units (2000 MIPS); Large Instance with four EC2 compute units (4000 MIPS); and Extra Large Instance with eight EC2 compute units (8000 MIPS). We consider different numbers of VMs in our simulations to achieve different data center utilizations (i.e., from 300 VMs to 620 VMs, which means from 48% data center utilization to 98% data center utilization). In the simulations, we assign one job with a random MIs to each VM. The MIs of each job follows a normal distribution from 750,000 MIs to 4,500,000 MIs [48,49]. We vary the mean of the normal distribution for each type of instances. For example, for Small Instance VMs, the mean is 1,500,000 MIs, and for Extra Large Instance VMs, it is 3,750,000 MIs. All the results are averaged over 50 runs for randomly-generated VMs.

We compare our heuristic with four thermal-aware greedy algorithms and one power-aware heuristic. Minimum Inlet Temperature (MinTinlet) is a greedy algorithm which allocates each VM to a chassis with the minimum inlet temperature. Minimum of Maximum Inlet Temperature (MinMaxTinlet) is a greedy algorithm which allocates each VM to a chassis which causes the minimum increase in the maximum inlet temperature. The third greedy algorithm is Minimum Summation Differential Inlet Temperature (MinSumDTinlet), which allocates each VM to a chassis which causes a minimum increase in the sum of inlet temperatures of all chassis (before and after allocation) within the data center. The main objective is to cause a minimum change in the inlet temperature of all chassis. LRH [50] is another greedy algorithm that assigns VMs to idle servers based on the chassis' contribution to heat recirculation. Finally, we use one of the implemented VM allocation policies in the CloudSim, "PowerVmAllocationPolicySimple" [51], which allocates each VM to the first chassis with sufficient MIPS capacity.

6.2. Simulation Results

The performance of our proposed heuristic is depicted in Figure 2. On average, more than 30% energy savings are achieved by our heuristics in higher utilizations in comparison with other greedy VM allocation heuristics. This is due to the fact that using the consolidation technique may have an effect on heat recirculation phenomena and increase the total energy consumption. The numbers of powered-off chassis for greedy algorithms are considerably greater than the numbers of powered-off chassis for MITEC-GA; thus, powered-on chassis are over-utilized in these greedy algorithms. MITEC-GA turns on more chassis; however, powered-on chassis are not fully utilized, resulting in less of an effect on the heat recirculation phenomena and consequently less dynamic energy consumption.

As shown in Figures 3 and 4, MITEC-GA can efficiently reduce both computing and cooling energy. It can reduce more cooling energy especially when the utilization of servers is high (i.e., VM number = 620). This is because MITEC-GA turns on more chassis to obtain load balancing within the system, avoiding the creation of hotspots. This load balancing allows higher temperatures for supplied cooled air without exceeding the red-line temperature in each server for a given workload.

The supply temperatures of different algorithms are also shown in Figures 5 and 6. Thermal-aware greedy algorithms provide higher supply temperatures; however, they are not as efficient as MITEC-GA, which considers both thermal and computing power simultaneously. Simulation results also suggest that using a consolidation technique may have an effect on heat recirculation phenomena and increase the total energy consumption. As depicted in Table 1, the numbers of off-chassis for greedy algorithms are greater than the numbers of off-chassis of MITEC-GA; thus, on-chassis are over-utilized in these greedy algorithms. The effect of over-utilizing the chassis can be seen in Figure 2, where their energy consumption is significantly higher than MITEC-GA. MITEC-GA turns on more chassis; however, on-chassis are not fully utilized, they consume less dynamic energy consumption, and they have less effect on heat recirculation phenomena.

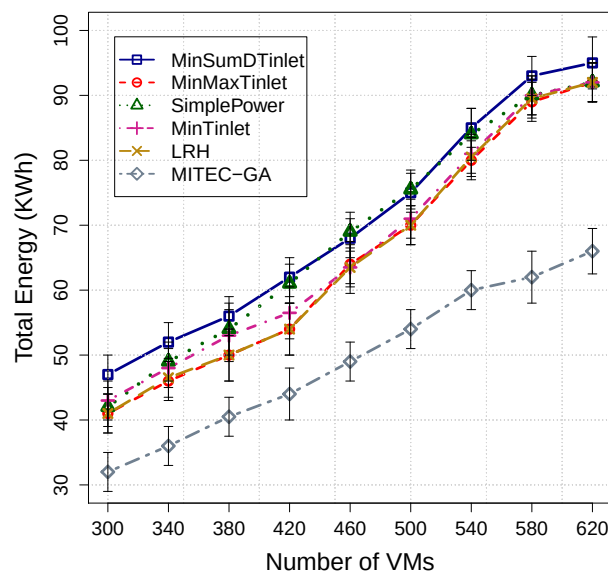


Figure 2. Total energy consumption vs. Virtual Machine (VM) numbers. MITEC-GA: minimizing the total energy consumption of a heterogeneous data center—genetic algorithm.

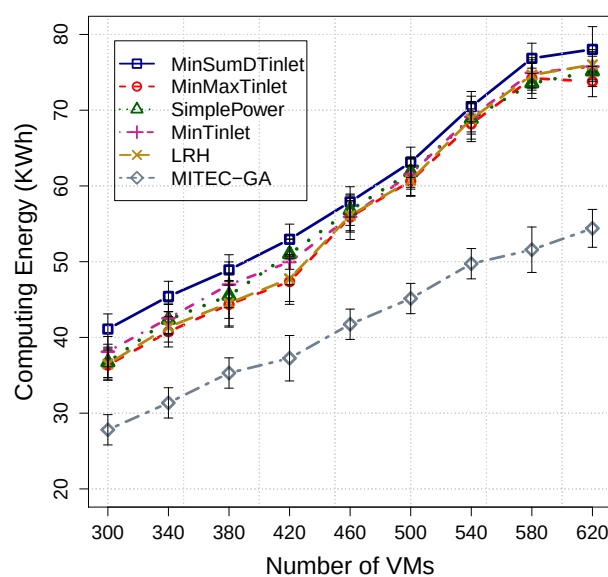


Figure 3. Computing energy consumption vs. VM numbers.

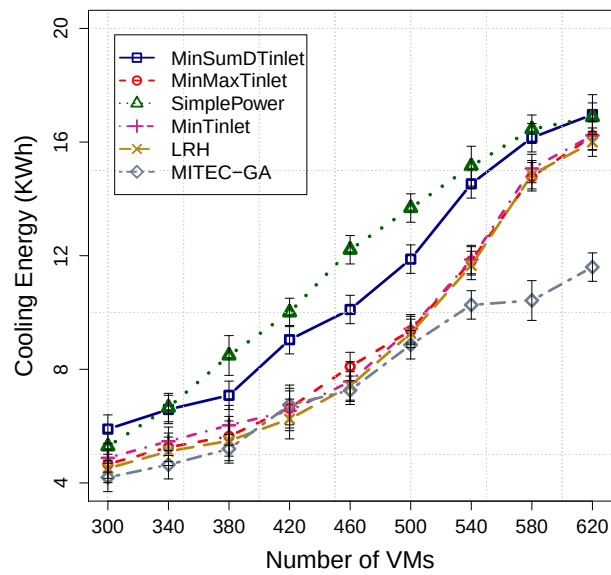


Figure 4. Cooling energy consumption vs. VM numbers.

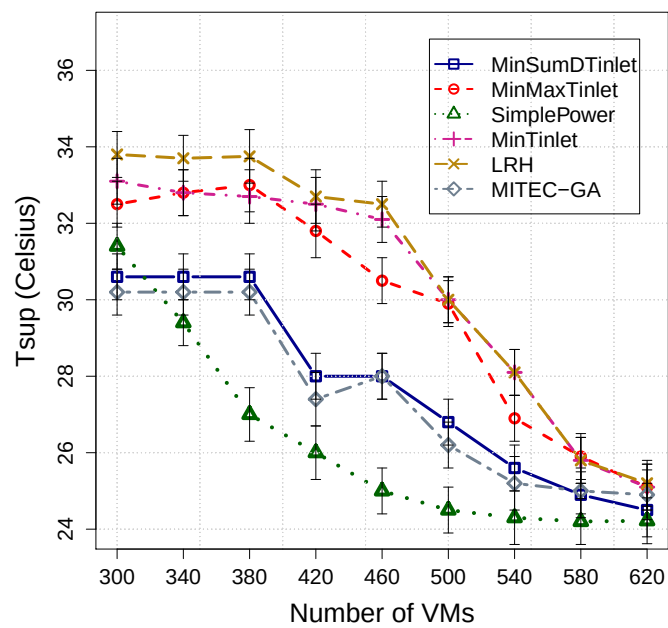


Figure 5. Supply temperature vs. VM numbers.

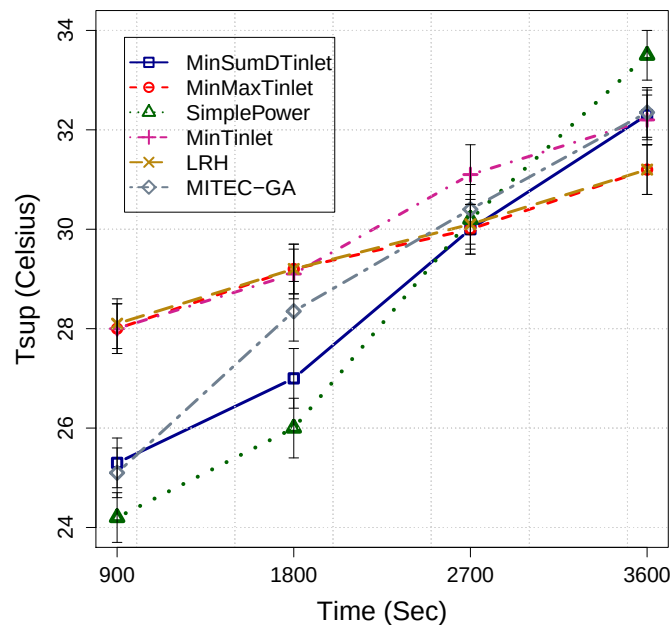


Figure 6. Supply temperature of different algorithms over time.

In terms of the cost–benefit analysis, reducing energy consumption can reduce the power capacity-related costs as well as the energy costs. Assuming \$0.12 energy cost per kWh [52], by the energy saving of up to 30% which is achieved compared to thermal-aware greedy algorithms and power-aware VM allocation heuristics, MITEC-GA can save around 131,400–254,040 kWh of energy yearly, meaning \$15,768 to \$30,484 savings on energy bills per year.

Table 1. Numbers of off-chassis in different VM allocation algorithms for different numbers of VMs.

Number of VMs	MinSumDTinlet	MinMaxTinlet	SimplePower	MinTinlet	LRH	MITEC-GA
300	25	25	25	25	25	23
340	22	22	22	22	22	20
380	19	19	19	18	19	17
420	15	15	15	15	15	13
460	12	12	12	12	12	10
500	9	9	9	9	9	7
540	5	5	5	5	5	4
580	2	2	2	2	2	0
620	1	1	1	1	1	0

7. Conclusions

In this paper, we formulated the problem of minimizing the total energy consumption of a heterogeneous data center (MITEC) as a non-linear integer optimization problem. We provided a thermal-aware VM allocation heuristic, MITEC-GA, based on the genetic algorithm. The MITEC-GA is a meta-heuristic optimization technique, which randomly searches the feasible solution space and finds a near-optimal solution for the allocation of thermal-aware VMs. The algorithm operates in an iterative manner. In each iteration, the MITEC-GA tries to generate the solutions that are a better fit than the current solution based on the total energy consumption of a data center. By the repeated modification of the existing solution, the MITEC-GA evolves the current solution toward the optimal

solution. Experimental results reveal the effectiveness of our proposed heuristic and demonstrate that, on average, more than 30% energy savings are achieved by our heuristic in comparison with other greedy VM allocation heuristics. In future works, we will consider an on-line thermal-aware VM allocation problem and use VM migration techniques to achieve more energy savings. We will also consider the dynamic effect of fan speeds on heat recirculation phenomena and try to provide an on-line fan speed control mechanism.

Author Contributions: Conceptualization, A.A., A.K. and S.M.G.; Data curation, A.A.; Formal analysis, A.A. and S.M.G.; Funding acquisition, S.M.G.; Investigation, A.A.; Methodology, A.A.; Project administration, A.A., A.K. and S.M.G.; Resources, A.A. and S.M.G.; Software, A.A.; Supervision, A.K. and S.M.G.; Validation, A.A.; Visualization, S.M.G.; Writing—original draft, A.A.; Writing—review & editing, S.M.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: We would like to express our special thanks of gratitude to Seyed Majid Zahedi at the University of Waterloo for his support and valuable advice on this project, which inspired us to improve the quality of this article.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Senyo, P.K.; Addae, E.; Boateng, R. Cloud computing research: A review of research themes, frameworks, methods and future research directions. *Int. J. Inf. Manag.* **2018**, *38*, 128–139. [\[CrossRef\]](#)
2. Amoretti, M.; Zanichelli, F.; Conte, G. Efficient autonomic cloud computing using online discrete event simulation. *J. Parallel Distrib. Comput.* **2013**, *73*, 767–776. [\[CrossRef\]](#)
3. Basmadjian, R. Flexibility-Based Energy and Demand Management in Data Centers: A Case Study for Cloud Computing. *Energies* **2019**, *12*, 3301. [\[CrossRef\]](#)
4. Georgilakis, P.S. Review of Computational Intelligence Methods for Local Energy Markets at the Power Distribution Level to Facilitate the Integration of Distributed Energy Resources: State-of-the-art and Future Research. *Energies* **2020**, *13*, 186. [\[CrossRef\]](#)
5. Shuja, J.; Gani, A.; Shamshirband, S.; Ahmad, R.W.; Bilal, K. Sustainable cloud data centers: a survey of enabling techniques and technologies. *Renew. Sustain. Energy Rev.* **2016**, *62*, 195–214. [\[CrossRef\]](#)
6. Sebastio, S.; Trivedi, K.S.; Alonso, J. Characterizing machines lifecycle in google data centers. *Perform. Eval.* **2018**, *126*, 39–63. [\[CrossRef\]](#)
7. Hameed, A.; Khoshkbarforoushha, A.; Ranjan, R.; Jayaraman, P.P.; Kolodziej, J.; Balaji, P.; Zeadally, S.; Malluhi, Q.M.; Tziritas, N.; Vishnu, A.; et al. A survey and taxonomy on energy efficient resource allocation techniques for cloud computing systems. *Computing* **2016**, *98*, 751–774. [\[CrossRef\]](#)
8. Malla, S.; Christensen, K. A survey on power management techniques for oversubscription of multi-tenant data centers. *ACM Comput. Surv. (CSUR)* **2019**, *52*, 1–31. [\[CrossRef\]](#)
9. Avgerinou, M.; Bertoldi, P.; Castellazzi, L. Trends in data centre energy consumption under the european code of conduct for data centre energy efficiency. *Energies* **2017**, *10*, 1470. [\[CrossRef\]](#)
10. Diouani, S.; Medromi, H. Survey: An Optimized Energy Consumption of Resources in Cloud Data Centers. *Int. J. Comput. Sci. Inf. Secur. (IJCSIS)* **2018**, *16*.
11. Yeo, S.; Lee, H.H. Using mathematical modeling in provisioning a heterogeneous cloud computing environment. *Computer* **2011**, *44*, 55–62.
12. Wu, C.J. Architectural thermal energy harvesting opportunities for sustainable computing. *IEEE Comput. Archit. Lett.* **2014**, *13*, 65–68. [\[CrossRef\]](#)
13. Naserian, E.; Ghoreyshi, S.M.; Shafiei, H.; Mousavi, P.; Khonsari, A. Cooling aware job migration for reducing cost in cloud environment. *J. Supercomput.* **2015**, *71*, 1018–1037. [\[CrossRef\]](#)
14. Lee, E.K.; Viswanathan, H.; Pompili, D. Proactive thermal-aware resource management in virtualized HPC cloud datacenters. *IEEE Trans. Cloud Comput.* **2017**, *5*, 234–248. [\[CrossRef\]](#)
15. Liu, L.; Li, C.; Sun, H.; Hu, Y.; Xin, J.; Zheng, N.; Li, T. Leveraging heterogeneous power for improving datacenter efficiency and resiliency. *IEEE Comput. Archit. Lett.* **2015**, *14*, 41–45. [\[CrossRef\]](#)
16. Li, Y.; Wang, X.; Luo, P.; Pan, Q. Thermal-aware hybrid workload management in a green datacenter towards renewable energy utilization. *Energies* **2019**, *12*, 1494. [\[CrossRef\]](#)

17. Nada, S.; Said, M. Effect of CRAC units layout on thermal management of data center. *Appl. Therm. Eng.* **2017**, *118*, 339–344. [[CrossRef](#)]
18. Bai, Y.; Gu, L. Chip temperature-based workload allocation for holistic power minimization in air-cooled data center. *Energies* **2017**, *10*, 2123. [[CrossRef](#)]
19. Moazamigoodarzi, H.; Gupta, R.; Pal, S.; Tsai, P.J.; Ghosh, S.; Puri, I.K. Modeling temperature distribution and power consumption in IT server enclosures with row-based cooling architectures. *Appl. Energy* **2020**, *261*, 114355. [[CrossRef](#)]
20. He, Z.; He, Z.; Zhang, X.; Li, Z. Study of hot air recirculation and thermal management in data centers by using temperature rise distribution. In *Building Simulation*; Springer: Berlin/Heidelberg, Germany, 2016; Volume 9, pp. 541–550.
21. Patel, C.D.; Bash, C.E.; Belady, C.; Stahl, L.; Sullivan, D. Computational fluid dynamics modeling of high compute density data centers to assure system inlet air specifications. In Proceedings of the Pacific Rim ASME International Electronic Packaging Technical Conference and Exhibition (IPACK), Kauai, HI, USA, 8–13 July 2001; pp. 8–13.
22. Moore, J.; Chase, J.S.; Ranganathan, P. Weatherman: Automated, online and predictive thermal mapping and management for data centers. In Proceedings of the IEEE International Conference on Autonomic Computing (ICAC), Dublin, Ireland, 13–16 June 2006; pp. 155–164.
23. Sharma, R.K.; Bash, C.E.; Patel, R.D. Dimensionless Parameters For Evaluation Of Thermal Design And Performance Of Large-Scale Data Centers. In Proceedings of the 8th ASME/AIAA Joint Thermophysics and Heat Transfer Conference, St Louis, MO, USA, 24–26 June 2002; pp. 1–11.
24. Ferreto, T.C.; Netto, M.A.S.; Calheiros, R.N.; De Rose, C.A.F. Server consolidation with migration control for virtualized data centers. *J. Future Gener. Comput. Syst.* **2011**, *27*, 1027–1034. [[CrossRef](#)]
25. Cioara, T.; Anghel, I.; Salomie, I. Methodology for energy aware adaptive management of virtualized data centers. *Energy Effic.* **2017**, *10*, 475–498. [[CrossRef](#)]
26. Raj, V.M.; Shriram, R. Power management in virtualized datacenter—A survey. *J. Netw. Comput. Appl.* **2016**, *69*, 117–133. [[CrossRef](#)]
27. Rosikiewicz, J.; McKelvey, R.T.; Mittell, A.D. Virtual Machine Data Replication. U.S. Patent 8,135,748, 2012.
28. Li, H.; Zhu, G.; Cui, C.; Tang, H.; Dou, Y.; He, C. Energy-efficient migration and consolidation algorithm of virtual machines in data centers for cloud computing. *Computing* **2016**, *98*, 303–317. [[CrossRef](#)]
29. Nguyen, T.H.; Di Francesco, M.; Yla-Jaaski, A. Virtual machine consolidation with multiple usage prediction for energy-efficient cloud data centers. *IEEE Trans. Serv. Comput.* **2017**.
30. Farahnakian, F.; Pahikkala, T.; Liljeberg, P.; Plosila, J.; Hieu, N.T.; Tenhunen, H. Energy-aware VM consolidation in cloud data centers using utilization prediction model. *IEEE Trans. Cloud Comput.* **2016**. [[CrossRef](#)]
31. Shirvani, M.H.; Rahmani, A.M.; Sahafi, A. A survey study on Virtual Machine migration and server consolidation techniques in DVFS-enabled cloud datacenter: Taxonomy and challenges. *J. King Saud-Univ. Comput. Inf. Sci.* **2020**, *32*, 267–286.
32. Wang, S.; Qian, Z.; Yuan, J.; You, I. A DVFS based energy-efficient tasks scheduling in a data center. *IEEE Access* **2017**, *5*, 13090–13102. [[CrossRef](#)]
33. Ghoreyshi, S.M. Energy-efficient resource management of cloud datacenters under fault tolerance constraints. In Proceedings of the 2013 International Green Computing Conference Proceedings, Arlington, VA, USA, 27–29 June 2013; pp. 1–6.
34. Pakbaznia, E.; Pedram, M. Minimizing data center cooling and server power costs. In Proceedings of the 14th ACM/IEEE International Symposium on Low Power Electronics and Design, San Francisco, CA, USA, 19–21 August 2009; pp. 145–150.
35. Lin, M.; Wierman, A.; Andrew, L.L.; Thereska, E. Dynamic right-sizing for power-proportional data centers. In Proceedings of the 2011 IEEE INFOCOM, Shanghai, China, 10–15 April 2011; pp. 1098–1106.
36. Chen, Y.; Das, A.; Qin, W.; Sivasubramaniam, A.; Wang, Q.; Gautam, N. Managing server energy and operational costs in hosting centers. In Proceedings of the 2005 ACM SIGMETRICS International Conference on Measurement and Modeling of Computer Systems, Banff, AB, Canada, 6–10 June 2005; pp. 303–314.
37. Lucchese, R. *Cooling Control Strategies in Data Centers for Energy Efficiency and Heat Recovery*; Luleå University of Technology: Luleå, Sweden, 2019.

38. Beitelmal, A.H.; Patel, C.D. Thermo-fluids provisioning of a high performance high density data center. *J. Distrib. Parallel Databases* **2007**, *21*, 227–238. [\[CrossRef\]](#)
39. Moore, J.D.; Chase, J.S.; Ranganathan, P.; Sharma, R.K. Making scheduling "Cool": Temperature-aware workload placement in data centers. In Proceedings of the 2005 USENIX Annual Technical Conference, Anaheim, CA, USA, 10–15 April 2005; pp. 61–75.
40. Patel, C.D.; Bash, C.E.; Sharma, R.; Beitelmal, M.; Friedrich, R. Smart Cooling of Data Centers. In Proceedings of the Pacific RIM/ASME International Electronics Packaging Technical Conference and Exhibition (IPACK), Maui, HI, USA, 6–11 July 2003; pp. 129–137.
41. Bash, C.E.; Patel, C.D.; Sharma, R.K. Efficient thermal management of data centers—Immediate and long-term research needs. *J. HVAC&R Res.* **2003**, *9*, 137–152.
42. Sharma, R.K.; Bash, C.E.; Patel, C.D.; Friedrich, R.J.; Chase, J.S. Balance of power: dynamic thermal management for Internet data centers. *IEEE Internet Comput.* **2005**, *9*, 42–49. [\[CrossRef\]](#)
43. Tang, Q.; Mukherjee, T.; Gupta, S.K.S.; Cayton, P. Sensor-based fast thermal evaluation model for energy efficient high-performance datacenters. In Proceedings of the Fourth International Conference on Intelligent Sensing and Information Processing (ICISIP), Bangalore, India, 15–18 December 2006; pp. 203–208.
44. Tang, Q.; Gupta, S.K.; Varsamopoulos, G. Energy-efficient thermal-aware task scheduling for homogeneous high-performance computing data centers: A cyber-physical approach. *IEEE Trans. Parallel Distrib. Syst.* **2008**, *19*, 1458–1472. [\[CrossRef\]](#)
45. Pakbaznia, E.; Ghasemazar, M.; Pedram, M. Temperature-aware dynamic resource provisioning in a power-optimized datacenter. In Proceedings of the Conference on Design, Automation and Test in Europe (DATE), Dresden, Germany, 8–12 March 2010; pp. 124–129.
46. Abbasi, Z.; Varsamopoulos, G.; Gupta, S.K.S. Thermal aware server provisioning and workload distribution for internet data centers. In Proceedings of the 19th ACM International Symposium on High Performance Distributed Computing (HPDC), Chicago, IL, USA, 21–25 June 2010; pp. 130–141.
47. Tang, X.; Liao, X.; Zheng, J.; Yang, X. Energy efficient job scheduling with workload prediction on cloud data center. *Clust. Comput.* **2018**, *21*, 1581–1593. [\[CrossRef\]](#)
48. Sun, X.; Su, S.; Xu, P.; Jiang, L. Optimizing multi-dimensional resource utilization in virtual data center. In Proceedings of the 2011 4th IEEE International Conference on Broadband Network and Multimedia Technology, Shenzhen, China, 28–30 October 2011; pp. 395–400.
49. Sun, X.; Su, S.; Xu, P.; Chi, S.; Luo, Y. Multi-dimensional resource integrated scheduling in a shared data center. In Proceedings of the 2011 31st International Conference on Distributed Computing Systems Workshops, Minneapolis, MN, USA, 20–24 June 2011; pp. 7–13.
50. Mukherjee, T.; Banerjee, A.; Varsamopoulos, G.; Gupta, S.K.S.; Rungta, S. Spatio-temporal thermal-aware job scheduling to minimize energy consumption in virtualized heterogeneous data centers. *J. Comput. Netw.* **2009**, *53*, 2888–2904. [\[CrossRef\]](#)
51. Beloglazov, A.; Buyya, R. Optimal online deterministic algorithms and adaptive heuristics for energy and performance efficient dynamic consolidation of virtual machines in Cloud data centers. *J. Concurr. Comput. Pract. Exp.* **2012**, *24*, 1397–1420. [\[CrossRef\]](#)
52. Rasmussen, N. *Implementing Energy Efficient Data Centers*; American Power Conversion: South Kingstown, RI, USA, 2006.

