Frontiers of Information Technology & Electronic Engineering
www.zju.edu.cn/jzus; engineering.cae.cn; www.springerlink.com
ISSN 2095-9184 (print); ISSN 2095-9230 (online)
E-mail: jzus@zju.edu.cn

Thermal-aware relocation of servers in green data centers

Muhammad Tayyab CHAUDHRY[‡1], T. C. LING[1], S. A. HUSSAIN[2], Xin-zhu LU[1]

([1]*Department of Computer System & Technology, Faculty of Computer Science & Information Technology, University of Malaya, Kuala Lumpur 50603, Malaysia*)

([2]*Department of Computer Science, COMSATS Institute of Information Technology, Lahore 54000, Pakistan*)

E-mail: mtayyabch@yahoo.com; tchaw@um.edu.my; asadhussain@ciitlahore.edu.pk; luxinzhu2013@siswa.um.edu.my

Received May 13, 2014; Revision accepted Aug. 8, 2014; Crosschecked Dec. 30, 2014

Abstract: Rise in inlet air temperature increases the corresponding outlet air temperature from the server. As an added effect of rise in inlet air temperature, some active servers may start exhaling intensely hot air to form a hotspot. Increase in hot air temperature and occasional hotspots are an added burden on the cooling mechanism and result in energy wastage in data centers. The increase in inlet air temperature may also result in failure of server hardware. Identifying and comparing the thermal sensitivity to inlet air temperature for various servers helps in the thermal-aware arrangement and location switching of servers to minimize the cooling energy wastage. The peak outlet temperature among the relocated servers can be lowered and even be homogenized to reduce the cooling load and chances of hotspots. Based upon mutual comparison of inlet temperature sensitivity of heterogeneous servers, this paper presents a proactive approach for thermal-aware relocation of data center servers. The experimental results show that each relocation operation has a cooling energy saving of as much as 2.1 kW·h and lowers the chances of hotspots by over 77%. Thus, the thermal-aware relocation of servers helps in the establishment of green data centers.

Key words: Servers, Green data center, Thermal-aware, Relocation
doi:10.1631/FITEE.1400174 **Document code:** A **CLC number:** TP393; TK12

1 Introduction

Data centers around the world consume an enormous amount of electric power each year. An average data center consumes the equivalent amount of electricity as 25 000 homes in USA (ABB, 2013). The cost of electricity expenditure exceeds the total capital expenditure over the working life of servers. Apart from computing, a major amount of electricity is also consumed in cooling the servers. This is because a data center has a closed environment and the electrical power consumed by information technology (IT) equipment is converted into heat (LD Didactic Gmbh) and an equal amount of power is needed to remove that heat and to maintain a proper working environment via cooling. Data centers must apply energy saving techniques to go green as a large part of the electricity is generated by burning fossil fuels. Considering the hot/cold aisle rack arrangement over raised floor design of a data center (U.S. Environmental Protection Agency, 2007), the cooling cost can be as much as the computing cost in terms of electricity used.

Power usage efficiency (PUE) is the ratio of total electricity usage by the data center to the electricity used for computing. If the cooling infrastructure consists of mechanical chillers only (Liu *et al.*, 2012), the PUE value may be equal to 2.0 unless power saving practices are adopted. Recent studies have shown a slight decrease in data center PUE worldwide to 1.93 (Koomey, 2011). The traditional approach of server consolidation to save computing power may result in utilization of few servers to their limits. However, the electricity consumption and the resulting heat dissipation from these servers reach

‡ Corresponding author
ORCID: Muhammad Tayyab CHAUDHRY, http://orcid.org/0000-0001-9485-0054

maximum in a small area of the data center. Rise in inlet air temperature can further increase the outlet temperature to such a limit that a hotspot is formed. A hotspot may trigger an otherwise idle cooling mechanism to start cooling or it may prolong the cooling process for an already active cooling mechanism. In both cases, the cooling is boosted for a larger area than that of the hotspot and more power is spent for cooling than consumed by the computing tasks inside that hotspot. By avoiding the chances of hotspots, the extra burden on cooling infrastructure can be avoided and power can be saved.

There are multiple factors which may combine to provide suitable condition for a hotspot. Among these factors is the physical phenomenon of cold air getting warmer as it reaches the inlet of the servers mounted near the top of racks. Furthermore, some systems such as the legacy servers are less power efficient and thus dissipate more heat. The processor is the most power consuming and thus the most heat dissipating hardware equipment on the motherboard (U.S. Environmental Protection Agency, 2007). Legacy processor architecture lacks the adaptive power usage capability and consumes more power compared to modern processors (Huck, 2011) and therefore dissipates more heat (Masiero, 2012). Servers which consume more power when idle are more prone to give rise to hotspots than others. A server is considered power efficient if it consumes less power when idle and provides more computing power per watt in terms of MHz per watt when active.

The chances of hotspot can be foreseen by analyzing the heat dissipation of different servers at various locations inside a data center with respect to inlet air temperature. This is due to the fact that some servers may dissipate more heat at higher inlet temperature while some servers may not be that sensitive, owing to the hardware architecture. Based upon this fact, if the physical location of each server is determined according to the inlet temperature sensitivity, the peak outlet temperature of the servers and the chances of hotspots can be reduced. For the servers that are already mounted in racks, inlet temperature sensitivity analysis can help rearrange the locations of a set of servers. Hotspot avoidance based relocation of data center servers can be a part of capacity planning or it can be done in parallel with traditional cooling mechanism optimization based techniques (Lee *et al.*, 2012). The server relocation technique proposed in this paper is a novel approach for minimizing the chances of hotspots and greening the data centers.

2 Related work

A power profiling based thermal map prediction and equipment relocation technique was reported by Jonas *et al.* (2010). Thermal map prediction was based on power profiles of the server chassis. On the basis of the fact that every chassis makes a contribution to the heat recirculation of all the chassises in the data center, an equipment relocation algorithm was proposed. However, it is practically complex to access the heat recirculation contribution coefficient for each of the hundreds of chassises in a data center. The complexity of server relocation can be reduced by focusing on the servers inside hotspot regions. If the servers with high electricity consumption or those having high utilization rate are placed at top of the racks where the inlet temperature is high, then this will increase, rather than decrease, the maximum outlet temperature.

If a power saving technique such as diskless booting (Tu *et al.*, 2010) is used, the servers will dissipate even less heat if they are located in a way that the inlet temperature has minimum effect on increasing the outlet temperatures of the servers. If the power consumption profiles of servers are created so that the least power is used to execute a given computing load and to ensure performance and profit (Kusic *et al.*, 2009), the scheduling algorithm can save more power if hotspots are avoided.

Instead of using a neural network to predict the outlet temperature, the thermal profiles can be used to predict the thermal map and chance of hotspots (Jonas *et al.*, 2007; 2010). Inlet temperature variation may lead to hotspot and cause an additional burden to the cooling mechanism. Data center energy efficiency and power consumption based scheduling techniques can perform better if the computational workload is distributed among servers on the basis of comparatively low inlet temperature preference (Tang *et al.*, 2007; Mukherjee *et al.*, 2009; Ahuja, 2012). Similarly, the reduction in recirculation of heat is more effective if the servers are arranged according to their sensitivities to inlet temperature hike (Tang *et al*., 2007).

The research aiming to increase the thermostat setting (Banerjee *et al.*, 2010; 2011) for cold air in a data center or to model the thermal map (Tang *et al.*, 2006; Ahuja *et al.*, 2011) should consider the optimization of server locations as a prerequisite to implementation. This is also applicable to recent ASHRAE standards (ASHRAE, 2011) for enhanced inlet temperatures. The coefficients of heat recirculation and extraction for the data center servers (Tang *et al*., 2006) are sensitive to the inlet temperature increment, and the values of coefficients should not be affected by this phenomenon. If servers are maximally used through backfilling (Wang *et al.*, 2009a; 2009b), the chances of hotspots are increased with the increase in inlet temperature. Therefore, the inlet temperature should be considered before backfilling or the servers should be relocated accordingly and then backfilled to avoid hotspots. Similarly, the server consolidation techniques for minimizing the number of active servers should choose only those servers that will not cause hotspots. This is because the consolidated servers will be at their peak utilization all the time (Corradi *et al.*, 2011).

Task-temperature profiles used for thermal-aware workload scheduling should consider the effect of inlet temperature sensitivity of the physical servers upon the scheduling outcome in terms of the thermal map to be unexpected (Wang *et al.*, 2012). The thermal profiling based techniques introduced by Rodero *et al.* (2010; 2012) cannot estimate or create the generic profiles of all the homogeneous servers which are located at different inlet air temperatures across a data center. Hence, there is a gap in research related to the thermal-aware arrangement of data center servers. This paper presents a thermal-aware server location evaluation and relocation of virtualized data center servers to optimum locations. The proposed approach results in prevention of possible hotspots and cooling energy saving, as well as increased effectiveness of thermal-aware scheduling techniques for those centers.

Cooling-aware workload placement with performance constraints was proposed by Sansottera and Cremonesi (2011), in which the rise in inlet temperature is due to heat recirculation. The heat recirculation was considered to be due to the air flow. The temperature of hot air from each server was declared to be due to power consumption according to computational workload on that server. Various test case scenarios were analyzed by computational fluid dynamics (CFD) simulations with different power consumption levels for the servers in order to profile each server for heat recirculation. This approach is based upon the prior work of Tang *et al*. (2006). These profiles were used to evaluate the highest possible thermostat setting and the lowest possible heat recirculation, and maximize the performance. This approach leads to the utilization of each server according to the thermal profile. The simulated scenario of the data center has two computer room air conditioning (CRAC) units at two opposite boundaries of the data center hall. One of these CRAC units was turned off so that hot air would be removed less efficiently from that region and heat recirculation may occur. In such a case, the servers that have a high heat recirculation impact are always underutilized. The servers consume up to 60% of peak power in the idle state as shown by our experiments. Therefore, it is not energy efficient to keep some servers idle or underutilized. Instead, we propose to identify the servers that are affected by heat recirculation and identify the outlet temperature at various utilization levels. Then such servers can be relocated at other locations inside the data center, so that these servers can be maximally used with comparatively low outlet temperature at the new location.

Among pioneer work, Tang *et al*. (2006) created the heat recirculation profiles and heat-exit profiles of the data center servers by using various power consumption levels in CFD simulations. These profiles can be used to predict the thermal map of the data center, given a power distribution vector and a heat recirculation coefficient matrix. This is a fast method for thermal prediction. However, the CFD simulations are time consuming (usually costing dozens of hours). Also, heterogeneous servers do not necessarily consume the same amount of power at the same level of CPU utilization and therefore do not have the same outlet temperatures despite the same inlet temperature. This makes the CFD-based profiling approach prone to hardware related limitations that can be verified only through experiments upon real hardware as we show in this paper.

3 Data center energy modeling preliminaries

By the law of energy conservation, the number of watts of electrical power consumed is converted into an equivalent number of joules of thermal energy (LD Didactic Gmbh). From now on the words 'power', 'electricity', and 'energy' are used interchangeably (energy is consumed per unit of time). If $E^i_{\text{computing}}$ is the electricity consumed by data center server i, this energy is converted to E^i_{joules} as

$$E^i_{\text{computing}}=E^i_{\text{joules}}. \tag{1}$$

As explained in Moore *et al.* (2005), power consumed by water-chilled CRAC units at HP labs is calculated with reference to the cold air set temperature and is called the coefficient of performance (COP). The COP is the ratio of the heat (Q) to the amount of work done (w):

$$\text{COP}=Q/w. \tag{2}$$

The COP has a numerical value which increases with the increase in supplied cold air temperature. The electrical energy E^i_{cooling} consumed to remove the heat dissipated by server i by supplying the cold air at a set temperature T^i_{received} can be written as

$$E^i_{\text{cooling}}=E^i_{\text{joules}}/\text{COP}(T^i_{\text{received}}). \tag{3}$$

Eq. (3), however, differs from Moore *et al.* (2005) as it considers the inlet air temperature that each server is receiving.

Due to heat recirculation, the cold air gets hot when it travels towards servers from the perforated tiles of a hollow floor. The hike in inlet air temperature has a direct impact on outlet air temperature for each server. The servers near the top of each rack are the victims of this phenomenon. These servers will dissipate more heat because the inlet air temperature increases despite the fact that they might not be fully used. So, the servers at the top of the racks put more burden on the cooling system than the servers near the floor, because of the rise in inlet air temperature. The COP curve (Moore *et al.*, 2005) fails to give a solution to this.

Therefore, the total electricity consumption E^i_{Total} for running server i can be written as

$$E^i_{\text{Total}}=E^i_{\text{computing}}+E^i_{\text{cooling}}. \tag{4}$$

Using Eq. (3),

$$E^i_{\text{Total}}=E^i_{\text{computing}}+E^i_{\text{joules}}/\text{COP}(T^i_{\text{received}}), \tag{5}$$

or using Eq. (1),

$$E^i_{\text{Total}}=E^i_{\text{computing}}+E^i_{\text{computing}}/\text{COP}(T^i_{\text{received}}). \tag{6}$$

The total energy consumption of a data center can then be written as

$$\begin{aligned} E_{\text{Total}} &= \sum_{i=1}^{n} E^i_{\text{Total}} = \sum_{i=1}^{n}\left(E^i_{\text{computing}} + \frac{E^i_{\text{computing}}}{\text{COP}(T^i_{\text{received}})}\right) \\ &= \sum_{i=1}^{n} E^i_{\text{computing}}\left(1+\frac{1}{\text{COP}(T^i_{\text{received}})}\right). \end{aligned} \tag{7}$$

The total electrical energy consumed by a data center (Eq. (7)) contains the electrical energy consumed by physical servers and the cooling system. Thus, with the knowledge of electricity consumption of the servers, the data center's cooling energy and therefore the total energy consumption can be calculated. This paper proposes to calculate the cooling energy consumption for each server based upon the inlet air temperature at that server. This means that the COP value should not be taken at the set inlet air temperature (T_{set}) of CRAC. A useful fact is that the COP value increases with the increase in inlet air temperature. So, a server that is receiving the cold air at a higher temperature will be responsible for a smaller share in total cooling energy consumption according to Eq. (7). On the other hand, the servers having high temperature of inlet air (T^i_{received}) will have a corresponding increase in the outlet air temperature as

$$\Delta T^i=T^i_{\text{received}}-T_{\text{set}}, \tag{8}$$

where ΔT^i is the increase in inlet temperature of server i and causes the equivalent increase in outlet temperature of the server ($T^i_{\text{outlet(increased)}}$). The original outlet temperature can be given by

$$T^i_{\text{outlet}}=T^i_{\text{outlet(increased)}}-\Delta T^i. \tag{9}$$

The increased outlet temperature of a server due to the increase in inlet temperature not only puts extra burden on the cooling mechanism but also may form hotspots. The former effect is independent of workload on the server while the latter occurs when the server is executing the workload. The servers with an inlet air temperature higher than the set value ($T^i_{\text{received}}>T_{\text{set}}$) means that the cooling energy being wasted for any server i is

$$\Delta E^i_{T\_\text{inlet}}=E^i_{T\_\text{set}}-E^i_{T\_\text{received}}, \tag{10}$$

where $E^i_{T\_\text{set}}$ is the cooling energy used to supply the cold air to server i at temperature T_{set}, and $E^i_{T\_\text{received}}$ shows the energy used to supply air at temperature T^i_{received}. This is because $T^i_{\text{received}}>T_{\text{set}}$ and therefore $E_{T\_\text{set}}>E^i_{T\_\text{received}}$. So, $\Delta E^i_{T\_\text{inlet}}$ is the energy wasted for server i. As a result, the outlet temperature rises by ΔT^i (Eq. (8)). This causes an equivalent energy to be wasted to cool the outlet air of server i that is extra hot by ΔT^i. Therefore, the total cooling energy wasted is the sum of cooling energy wasted and the extra cooling energy spent for all servers, and can be given by

$$\begin{aligned} E_{\text{cooling\_wasted}} &= \sum_{i=1}^{n}(\Delta E^i_{T\_\text{inlet}} + \Delta E^i_{T\_\text{inlet}}) \\ &= \sum_{i=1}^{n} 2\Delta E^i_T. \end{aligned} \tag{11}$$

The value of $\Delta E^i_{T\_\text{inlet}}$ will be equal to or more than zero depending upon the position of the server with respect to the floor. In this paper, we propose to minimize the energy wastage on cooling as given by Eq. (11). The maximum allowed inlet temperature can be represented by $T_{\max}$ beyond which either the hotspot can occur or the server hardware may fail.

We define a problem statement for equipment relocation as

Problem If there exists server i such that $T^i_{\text{received}}>T_{\text{set}}$, find server j satisfying

(1) $T^j_{\text{received}}<T^i_{\text{received}}$

(2) $T^j_{\text{outlet}}<T^i_{\text{outlet}}$

(3) $T^i_{\text{inlet}}<T_{\max}$

(4) $T^j_{\text{inlet}}<T_{\max}$

subject to post relocation conditions:

(C1) $T^i_{\text{outlet(relocated)}} < T^i_{\text{outlet}}$

(C2) $T^j_{\text{outlet(increased)}}<T^i_{\text{outlet}}$

(C3) $T^j_{\text{received}}-T_{\text{set}}<T^i_{\text{received}}-T_{\text{set}}$

(C4) min $E_{\text{cooling\_wasted}}$.

The energy wasted in cooling (Eq. (10)) can be minimized by equipment relocation. Since this energy wastage is curable and therefore should not be included in calculating the total energy consumption, Eq. (7) can be generalized to

$$E_{\text{Total}} = \sum_{i=1}^{n} E^i_{\text{computing}}\left(1+\frac{1}{\text{COP}(T_{\text{set}})}\right) - E_{\text{cooling\_wasted}}. \tag{12}$$

Using Eqs. (6) and (11), the value of $E_{\text{cooling\_wasted}}$ can be obtained. The increase in cooling energy wastage results in a rise in data center PUE. $E_{\text{cooling\_wasted}}$ can be calculated through the thermodynamics model given by Tang *et al.* (2006), which requires the blow rate of the server fans. The servers used for demonstration (Table 1) in this study have dynamic fan rates and thus the calculation of $E_{\text{cooling\_wasted}}$ becomes complicated. Instead, the cooling cost was calculated indirectly through the approach proposed by Moore *et al.* (2005). Thus, the cooling cost is calculated with reference to the COP of the set temperature of the CRAC unit. We have extended this model for the calculation of $E_{\text{cooling\_wasted}}$ before applying the relocation algorithm. The energy wasted is calculated as the difference between the cooling energy for the supplied air temperature and the temperature of the cold air received by the servers, demonstrated as follows:

$$\begin{aligned} E_{\text{cooling\_wasted}} &= \sum_{i=1}^{m} 2\Delta E^i_T \\ &= \sum_{i=1}^{m} 2\left(\frac{E^i_{\text{computing}}}{\text{COP}(T_{\text{set}})} - \frac{E^i_{\text{computing}}}{\text{COP}(T^i_{\text{received}})}\right). \end{aligned} \tag{13}$$

This paper proposes to lower the cooling energy wastage by adjusting the location of the servers. While it may not be possible to totally eliminate $E_{\text{cooling\_wasted}}$, there is a possibility of normalizing the cooling energy wastage due to the increase in inlet temperature by lowering the average outlet temperature of the affected server(s) through relocation.

Consider that the same volume of air at temperature T_1 is heated to T_2. Then the heat at temperature T_2 is greater than that at T_1 (BBC, 2014). Applying the same concept on server outlets, if the temperature of the server outlet is lowered, this would signify the lowering of heat. As seen from Eq. (9), if the server is relocated to a location with comparatively low inlet temperature, the amount of heat dissipated is lower because the temperature of the outlet is lower at the new location. Thus, the cooling load is lowered. For the server which is relocated to the region of high inlet temperature as a location exchange, if the outlet temperature is lower than that of the previous server at the same location, the overall cooling load of both servers is decreased. The more homogeneous and the lower the outlet temperatures of the servers after relocation, the lower the cooling load.

From Eq. (9), it can be inferred that the outlet temperature depends upon the inlet temperature and server utilization level. If the server utilization remains the same, a change in inlet temperature has a direct impact upon outlet temperature. This allows prediction of the outlet temperature of server i at current inlet temperature (T^{i}_{received}) with respect to the inlet temperature of server j. Thus,

$$T^{i}_{\text{predicted\_outlet}}=(T^{i}_{\text{outlet}}-T^{i}_{\text{received}})+T^{j}_{\text{received}}. \quad (14)$$

The predicted outlet temperature of servers i and j can be used to evaluate the current location of each server for the possibility of hotspots. In this study, Eq. (14) is used to predict an array of outlet temperature values for any server.

4 Server relocation algorithm

This section presents the server relocation algorithm based upon the analysis and comparison of various variables related to performance, power, and temperature. Identical experiments were run on all servers. The experimental results of a server can be applied over homogeneous servers as well. The relocation algorithm (Algorithm 1) can be applied on two servers at a time. Therefore, for simplicity, two heterogeneous servers (type A and type B) are considered in the algorithm (Table 1). One member from each type of servers is chosen for the implementation.

Table 1 Specifications for server types

Server type	Processor
A	Intel Xeon CPU E5430 2.66 GHz
B	Intel Xeon CPU E5320 1.86 GHz

Algorithm 1 Server relocation

Input: $\boldsymbol{T}^{\text{A}}_{\text{idle\_inlet}}$, $\boldsymbol{T}^{\text{B}}_{\text{idle\_inlet}}$, $\boldsymbol{T}^{\text{A}}_{\text{inlet}}$, $\boldsymbol{T}^{\text{B}}_{\text{inlet}}$, $\boldsymbol{T}^{\text{A}}_{\text{outlet}}$, $\boldsymbol{T}^{\text{B}}_{\text{outlet}}$, $T_{\max}$

1 **if** $T^{\text{A}}_{\text{idle\_inlet}}<T^{\text{B}}_{\text{idle\_inlet}}$ && $T^{\text{A}}_{\text{idle\_inlet}}<T_{\max}$ && $T^{\text{B}}_{\text{idle\_inlet}}<T_{\max}$ **then**
2   $\Delta T_{\text{inlet}}=T^{\text{B}}_{\text{inlet}}-T^{\text{A}}_{\text{inlet}}$
3   $\Delta T_{\text{outlet}}=T^{\text{B}}_{\text{outlet}}-T^{\text{A}}_{\text{outlet}}$
4   **if** $\Delta T_{\text{inlet}}>0$ && $\Delta T_{\text{outlet}}>0$ **then**
5   **if** $\Delta T_{\text{outlet}}\geq\Delta T_{\text{inlet}}$ **then**
6   Calculate $\boldsymbol{T}^{\text{B}}_{\text{predicted\_outlet}}$ using Eq. (14)
7   Calculate $\boldsymbol{T}^{\text{A}}_{\text{predicted\_outlet}}$ using Eq. (14)
8   **if** $T^{\text{A}}_{\text{predicted\_outlet}}\leq T^{\text{B}}_{\text{outlet}}$ && $T^{\text{B}}_{\text{predicted\_outlet}}\leq T^{\text{B}}_{\text{outlet}}$ **then**
9   Switch locations of type A and type B servers
10   **end if**
11   **end if**
12   **end if**
13 **end if**

Note: $T^{\text{A}}_{\text{idle\_inlet}}$, $T^{\text{B}}_{\text{idle\_inlet}}$, $T^{\text{A}}_{\text{inlet}}$, $T^{\text{A}}_{\text{outlet}}$, $T^{\text{B}}_{\text{inlet}}$, $T^{\text{B}}_{\text{outlet}}$, ΔT_{inlet}, ΔT_{outlet}, $T^{\text{A}}_{\text{predicted\_outlet}}$, and $T^{\text{B}}_{\text{predicted\_outlet}}$ represent any element in $\boldsymbol{T}^{\text{A}}_{\text{idle\_inlet}}$, $\boldsymbol{T}^{\text{B}}_{\text{idle\_inlet}}$, $\boldsymbol{T}^{\text{A}}_{\text{inlet}}$, $\boldsymbol{T}^{\text{A}}_{\text{outlet}}$, $\boldsymbol{T}^{\text{B}}_{\text{inlet}}$, $\boldsymbol{T}^{\text{B}}_{\text{outlet}}$, $\Delta\boldsymbol{T}_{\text{inlet}}$, $\Delta\boldsymbol{T}_{\text{outlet}}$, $\boldsymbol{T}^{\text{A}}_{\text{predicted\_outlet}}$, and $\boldsymbol{T}^{\text{B}}_{\text{predicted\_outlet}}$, respectively

The proposed algorithm requires the idle state inlet temperatures and other parameters of two servers at each run. The idle state inlet temperatures are required to identify the difference in inlet temperatures and to check that the inlet temperature is less than the vendor specified maximum temperature (line 1). It is supposed that $T_{\max}$ is the same for all servers. Otherwise, in line 1 there may be different values for $T_{\max}$. The experiments performed on the pair of servers generate the data such as $T^{\text{A}}_{\text{inlet}}$, $T^{\text{B}}_{\text{inlet}}$, $T^{\text{A}}_{\text{outlet}}$, and $T^{\text{B}}_{\text{outlet}}$. The algorithm continues if type A server is located at the region with comparatively low inlet temperature than type B server. Lines 4 and 5 are performed to confirm that the differences in inlet and outlet temperatures of the two servers are available and that the outlet temperature difference is larger than inlet temperature difference. This provides an opportunity for predicting the outlet temperatures in lines 6 and 7. If the predicted temperature of type A server after relocation is lower than the outlet temperature of type B server before relocation and the predicted temperature of type B server after relocation is lower than the outlet temperature before relocation, the algorithm suggests switching locations of the servers.

5 Experimental setup

The proposed approach was tested over a set of heterogeneous servers, running a VMware ESXi 5.0 (VMware Inc., 2009) hypervisor. We used virtualized servers (hosts) as the hypervisor can give the detailed performance data and the frequency of the server processor can be manipulated at runtime. This helps in simulation for various processor frequencies to simulate the routine at which servers are actually used in the data center. Data are from three sources, thermal sensors, smart power meters, and the virtualized hosts, and are aggregated per host and per minute to match the time and hosts.

Servers were grouped according to their processor models (Table 1). The members of each server group are homogeneous. For implementation, two servers, one of type A and one of type B, were used.

To monitor the inlet and outlet air temperatures, external USB thermal sensors were used. The power consumption of each host was measured by USB smart power meters. Each server has up to eight virtual machines (VMs). Microsoft C# script was used as the workload booster to manipulate the VM operations. Each VM was running a CPU intensive benchmark Prime95 (Mersenne Research, Inc., 2012) and was kept in a suspended state. Each VM has a single virtual CPU. Two servers (one of type A and one of type B) were kept idle for about 10 h to prove the correlation between inlet and outlet temperatures.

Fig. 1 shows that the close correlation occurs between the inlet and outlet temperatures of the prototype servers in idle state. Experimental results presented in the next section will show that this relationship holds when the servers are active and this is a basic matrix of evaluating the post relocation outlet temperatures from the set of servers involved. We performed three sets of experiments involving one server from group A and another from group B at various CPU frequencies for server groups A and B (Table 2). Each experiment set contained at least two servers. Since the servers were heterogeneous and the difference between the processor frequencies was 0.8 GHz, dynamic frequency scaling was used to vary the maximum flips of the servers according to Table 2. Experiment sets 2 and 3 had the servers running at the same processor frequency and could approximately represent the scenarios when the servers were running underutilized, while in experiment set 1 both servers underwent maximum utilization.

All the experiment sets took 3 min idle time to start, next used the virtualized hosts (according to the frequency limit of experiment sets) for 30 min, and then brought the host to idle state to cool down to idle state temperature for 20 min. Altogether, each

Table 2 Sets of experiments performed

Experiment set	Processor frequency (GHz)	
	Type A server	Type B server
1	2.66	1.86
2	1.86	1.86
3	1.00	1.00

Each experiment set involves both types of servers. Type A server has a quicker processor than type B server

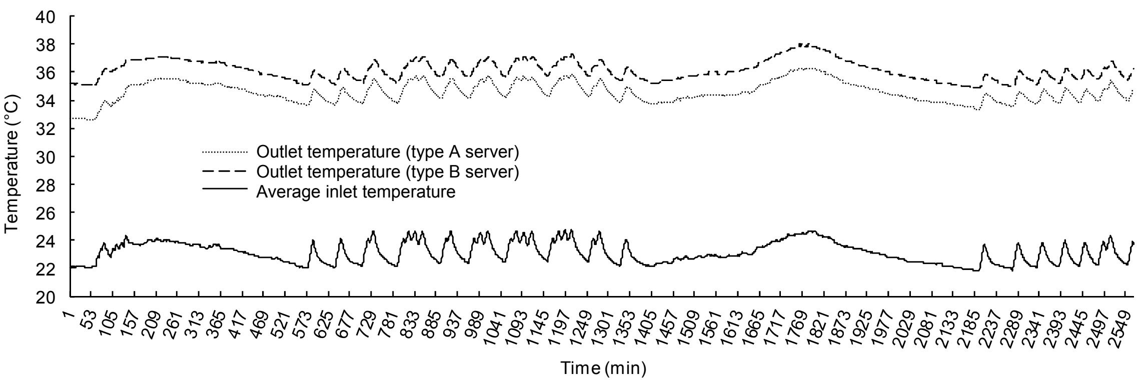

Fig. 1 Correlation between inlet temperature and outlet temperatures of type A and type B servers

experiment took approximately 55 min. The set temperature (T_{set}) was 21 °C.

The initial locations of the servers satisfy that type A server was placed in a colder area and type B server located in a hotter area. Table 3 shows the idle state of the servers. On average, type A server used less power when in idle state than type B server and this causes the average outlet temperature to be higher than that of the type A server in idle state (Table 3). However, type A server was receiving the inlet temperature at T_{set}, which was lower than the inlet temperature of type B server. This can be another reason of comparatively low outlet temperature of type A server. The COP for type A server was also lower than that of type B server according to the inlet temperature receieved.

Table 3 Idle state statistics of both servers

Server type	Average inlet temperature (°C)	COP ($T^i_{received}$)	Average idle state power consumption (W)	Average idle state outlet temperature (°C)
A	21.2	3.5	205	31.8
B	22.7	4.0	232	35.4

Type B server uses more electricity in the idle state

Thus, on the basis of the above observations, a hypothesis can be formed that the maximum temperature from hot air outlet of type A server will be lower than the maximum temperature from hot air outlet of type B server if the maximum power consumed by type A server is equal to the maximum power consumed by type B server, provided that type A server has a lower inlet temperature than type B server.

Consider two heterogeneous servers consuming equal amounts of electricity while executing similar workload. If the outlet temperatures of these two servers are not similar, the reason can be dissimilar inlet air temperatures. If type B server has a less powerful processor than type A server, then it will be giving less processing per unit of power consumption than type A server. If the processor of type A server has a higher maximum frequency and consumes less power in idle state and equal power at any level of processor utilization as compared to type B server (indicated by the outlet air temperature), then it will be worth predicting the outlet temperature of both servers after relocation on the basis of inlet temperature difference.

In idle state of servers, suppose that vectors $\boldsymbol{T}^{A}_{idle\_inlet}$ and $\boldsymbol{T}^{B}_{idle\_inlet}$ represent the inlet air time lapsed temperatures' series of type A and type B servers, respectively. Assume that $\boldsymbol{T}^{A}_{idle\_inlet}$ and $\boldsymbol{T}^{B}_{idle\_inlet}$ cover a reasonable time to make inference. If $\boldsymbol{P}^{A}_{idle}$ and $\boldsymbol{P}^{B}_{idle}$ are the idle power consumption vectors of type A and type B servers, respectively, during the experiments mentioned in Table 2, the outlet temperatures of the servers were saved in vectors $\boldsymbol{T}^{A}_{outlet}$ and $\boldsymbol{T}^{B}_{outlet}$ for type A and type B servers, respectively. The inlet temperatures of type A and type B servers were recorded in vectors $\boldsymbol{T}^{A}_{inlet}$ and $\boldsymbol{T}^{B}_{inlet}$, respectively.

Given that $\boldsymbol{T}^{A}_{idle\_inlet}<\boldsymbol{T}^{B}_{idle\_inlet}$, then $\boldsymbol{T}^{A}_{max\_outlet}<\boldsymbol{T}^{B}_{max\_outlet}$ provided that $\boldsymbol{P}^{A}_{max}\leq\boldsymbol{P}^{B}_{max}$. $\boldsymbol{T}^{A}_{max\_outlet}$ is the maximum outlet temperature vector from type A server when it consumes maximum power $\boldsymbol{P}^{A}_{max}$, and $\boldsymbol{T}^{B}_{max\_outlet}$ is the maximum outlet temperature vector of type B server when it consumes maximum power $\boldsymbol{P}^{B}_{max}$. If this hypothesis is proved through experiments given in Table 2, the next step will lead to an algorithm for equipment relocation.

6 Experimental results and discussion

In this section, the experimental results are presented and discussed with respect to the server relocation algorithm presented in the previous section. The input variables in Algorithm 1 are referred to Figs. 2 and 3, in which the servers undergo the experiment of maximum workload and CPU utilization. For Figs. 2–4, line 1 of Algorithm 1 is true for both hosts as the idle state inlet temperature and the maximum outlet temperature of type A server are always less compared with type B server. This indicates that if the difference in outlet temperature is dependent upon the difference in inlet temperatures, after switching places, when the same experiment is performed, type A server will have lower outlet temperature than type B server at the same location. Thus, the locations of type A and type B servers can be exchanged.

To verify that the inlet temperature and the outlet temperatures of both servers throughout the experiment remained such that type B server always had the higher inlet and outlet temperatures than type A server, the operations of lines 2 and 3 were performed. If the difference between the inlet temperatures ($\Delta \boldsymbol{T}_{\text{inlet}}$) remains above zero (line 3), it means that the outlet temperatures of both servers will satisfy that the type A server will have lower outlet temperature than the type B server. The outlet temperature difference ($\Delta \boldsymbol{T}_{\text{outlet}}$) remains around 2 °C (Fig. 5). The difference between $\Delta \boldsymbol{T}_{\text{outlet}}$ and $\Delta \boldsymbol{T}_{\text{inlet}}$ shows how much more heat is being dissipated from type B server than the heat difference between the inlet temperatures. This difference is quite significant in Fig. 14, when both servers ran experiment sets 2 and 3. Combined results of experiment sets 2 and 3 are presented in Figs. 11–16. According to Eq. (1), all the electricity is to be converted into heat. Therefore, unless type A server consumes more electricity than type B server, the outlet temperature of type A server will remain lower than that of type B server and the logical comparison (line 4) will be true. The power consumptions of both servers (Figs. 2 and 3) show that the maximum power consumption of type A server is always less than or equal to that of type B server throughout the experiment. Hence, the hypothesis presented earlier is proved (Fig. 5). The big hump of $\Delta \boldsymbol{T}_{\text{outlet}}$ (Fig. 5) is the sudden rise in the outlet temperature of type B server, while type A server took a while to get heated. This may be due to the fact that type B server consumes more energy in idle state than type A server and when the experiment set is conducted over type B server, the rate of rise in electriciy consumption of type B server increases more sharply than type A server during a short interval.

CPU utilization and effective time frequency of both servers are shown in Figs. 6 and 7, respectively. Both servers underwent maximum utilization of CPU, although they differed in the maximum time frequency. Before making decisions to switch locations of these two servers, the temperature predicion should be made. This is to foresee the effect of relocation (lines 6 and 7, Figs. 8–10). The inlet temperature difference ($\Delta \boldsymbol{T}_{\text{inlet}}$) was added to the outlet temperature of type A server to raise it and subtracted from type B server outlet temperature to lower it.

As shown in Fig. 10, the predicted temperature of type A server ($\boldsymbol{T}^{\text{A}}_{\text{predicted\_outlet}}$) is lower than the outlet temperature of type B server ($\boldsymbol{T}^{\text{B}}_{\text{outlet}}$), and the predicted outlet temperature of type B server ($\boldsymbol{T}^{\text{B}}_{\text{predicted\_outlet}}$) is also lower than $\boldsymbol{T}^{\text{B}}_{\text{outlet}}$. This means that if the experiment is repeated after switching the locations of both servers, not only type B server will

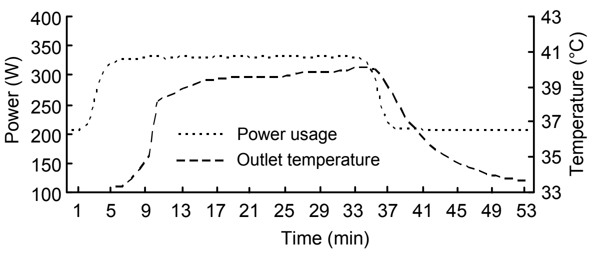

Fig. 2 Power usage and outlet temperature of type A server under experiment set 1

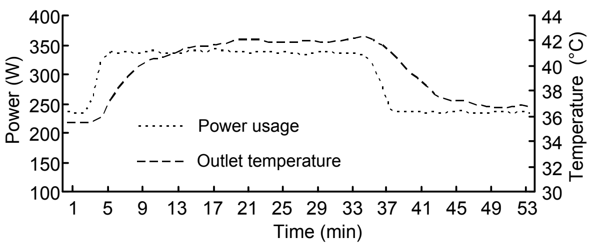

Fig. 3 Power usage and outlet temperature of type B server under experiment set 1

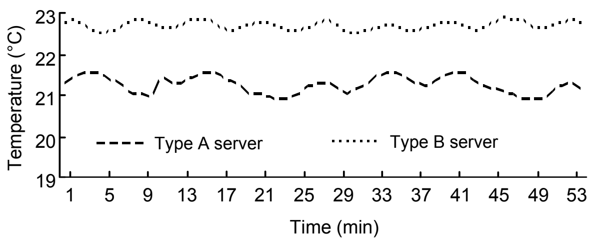

Fig. 4 Inlet temperatures of type A and type B servers under experiment set 1 (type B server is receiving hotter air at inlet)

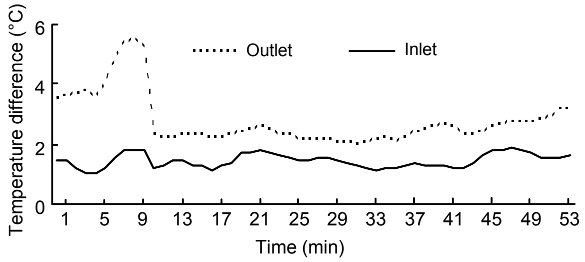

Fig. 5 The difference between outlet temperatures is always higher than the difference between inlet temperatures for type A server and type B server at initial locations under experiment set 1

have a lower outlet temperature (compared to its previous location), but also type A server will dissipate less heat at the new location than type B server at the old location.

The logical comparisons (line 8) are the post relocation conditions that were presented earlier in the problem statement and they guarantee that the relocation operation will result in hotspot avoidance and overall reduction in cooling load with homogeneous outlet temperatures of both servers. A significant difference between $\Delta \boldsymbol{T}_{\text{inlet}}$ and $\Delta \boldsymbol{T}_{\text{outlet}}$ (Fig. 14) shows a major imbalance in heat dissipation when the servers are underutilized. Type A server is underutilized in experiment sets 2 and 3, while type B server is fully utilized in experiment set 2 and underutilized in experiment set 3. However, type B server always consumes more energy than type A server and provides less processing potential per unit of electricity consumed than type A server. Type B server dissipates more heat even when both servers run at almost identical CPU frequencies in experiment sets 2 and 3.

Now we move onto the verification of post relocation considerations mentioned earlier in problem statement. The same sets of experiments were performed over type A and type B servers after switching their locations. The results are shown in Figs. 17–28. Figs. 17 and 18 indicate that type B

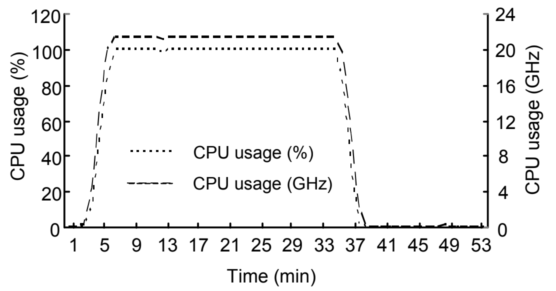

Fig. 6 CPU usage of type A server under experiment set 1

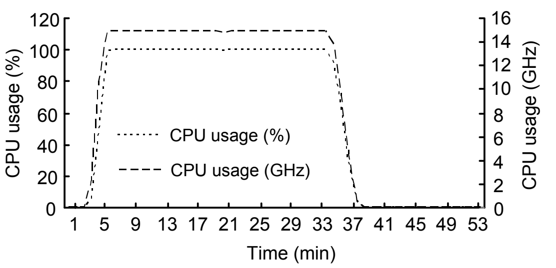

Fig. 7 CPU usage of type B server under experiment set 1

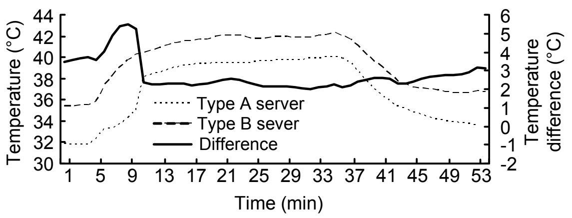

Fig. 8 Outlet temperatures of type A and type B servers, and their difference at current location under experiment set 1

The hump in outlet temperature difference is due to a sudden rise in the outlet temperature of type B server

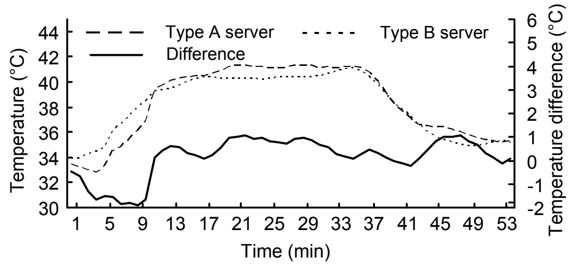

Fig. 9 Predicted outlet temperatures of type A and type B servers, and their difference after relocation on the basis of inlet temperatures under experiment set 1

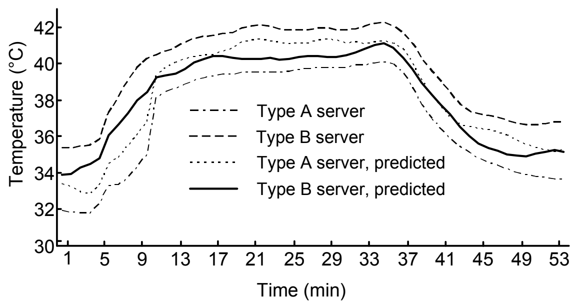

Fig. 10 Current and predicted outlet temperatures of type A and type B servers under experiment set 1

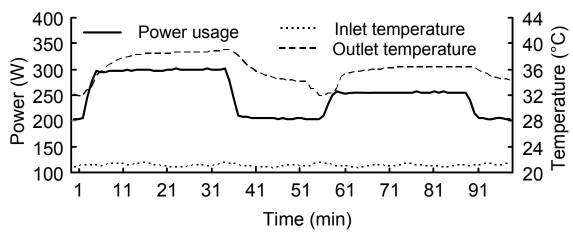

Fig. 11 Power consumption and inlet and oulet temperatures of type A server running at lower frequencies to match the pocessor of type B server under experiment sets 2 and 3

server shows a reduction in outlet temperature, whereas type A server outlet temperature is increased as compared to Figs. 2 and 3 with the increase in inlet temperature. However, the power consumptions of both servers follow the same trend as that before relocation (Figs. 2 and 3) when experiment sets 1–3 were repeated, so did the inlet air temperatures (Figs. 4, 20, and 25). The hump of Fig. 5 for $\Delta \boldsymbol{T}_{\text{outlet}}$ is flattened in Fig. 19, showing a positive change in difference between outlet temperatures. This shows that although the power consumption of type B server shoots up in the start of the experiment set just as in Fig. 3, the rise in outlet temperature of type B server in Fig. 18 is balanced by the higher rate of rise in outlet temperature of type A server in Fig. 17. This hump was responsible for an error in the prediction of Fig. 9 for a short interval of 4 min.

Figs. 19 and 26 show that the predicted $\Delta \boldsymbol{T}_{\text{outlet}}$ follows closely that of actual run. As shown in Figs. 21 and 27, the predicted temperatures ($\boldsymbol{T}^{\text{A}}_{\text{predicted\_outlet}}$ and $\boldsymbol{T}^{\text{B}}_{\text{predicted\_outlet}}$) follow closely the actual outlet temperatures ($\boldsymbol{T}^{\text{A}}_{\text{outlet}}$ and $\boldsymbol{T}^{\text{B}}_{\text{outlet}}$). Since the predicted outlet air temperature of type A server was less than that of type B server, the locations of these two servers were switched. Fig. 21 shows that $\boldsymbol{T}^{\text{B}}_{\text{predicted\_outlet}}$ is more accurate than $\boldsymbol{T}^{\text{A}}_{\text{predicted\_outlet}}$.

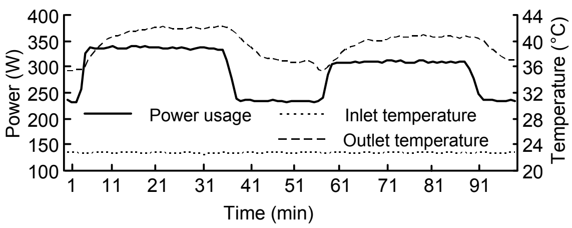

Fig. 12 Power consumption and inlet and oulet temperatures of type B server running at frequencies equal to those of type A server under experiment sets 2 and 3

The temperature is higher than that of type A server

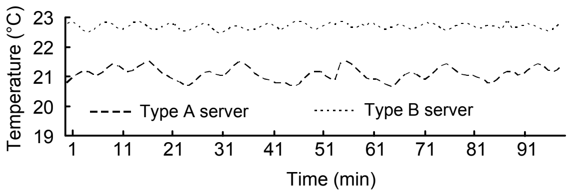

Fig. 13 Inlet temperatures of type A and type B servers when type B server is receiving hotter air at inlet under experiment sets 2 and 3

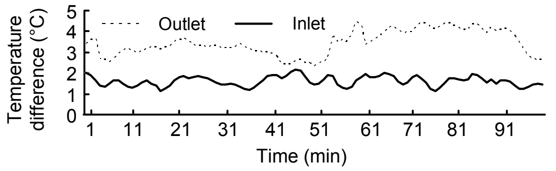

Fig. 14 Inlet and outlet temperature differences between type A and type B servers before moving under experiment sets 2 and 3

The outlet temperature difference is much higher than the inlet temperature difference, but the relocation algorithm does not use this

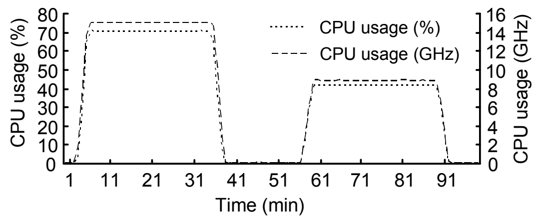

Fig. 15 CPU usage of type A server under experiment sets 2 and 3

Type A server running underutilized to match the processor frequency of type B server

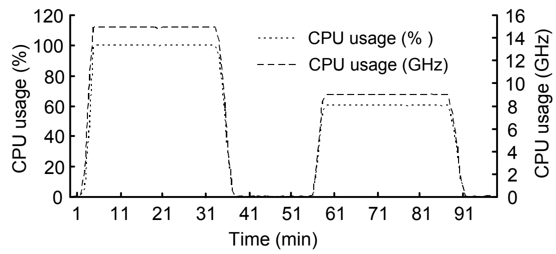

Fig. 16 CPU usage of type B server under experiment sets 2 and 3

Type B server running at its full processor in experiment set 2 and at 60% of its maximum frequency in experiment set 3

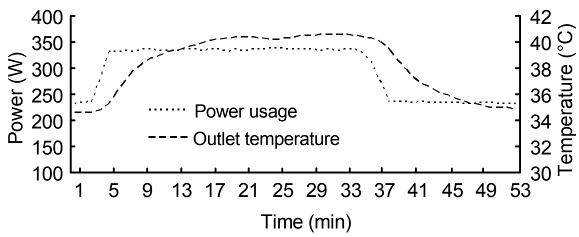

Fig. 17 Power usage and outlet temperature of type A server under experiment set 1

Type A server shows an increase in outlet temperature, but the power consumption is the same as that before moving

The reason is that $\boldsymbol{T}^{A}_{\text{predicted\_outlet}}$ was calculated based upon $\Delta\boldsymbol{T}_{\text{outlet}}$, which depended upon $\boldsymbol{T}^{B}_{\text{outlet}}$ at the previous location. Therefore, according to line 8 of Algorithm 1, the servers could be relocated even if $\boldsymbol{T}^{A}_{\text{predicted\_outlet}}$ is equal to $\boldsymbol{T}^{B}_{\text{outlet}}$ before relocation. This is proved by Figs. 27 and 28.

The results of experiment sets 2 and 3 in Figs. 27 and 28 prove the post relocation scenario. The significant difference between $\Delta\boldsymbol{T}_{\text{outlet}}$ (Fig. 14) and that of experiment set 3 (Fig. 26) is due to the fact that type B server dissipates more heat at low CPU utilization than type A server even when $\boldsymbol{T}^{A}_{\text{inlet}} \geq \boldsymbol{T}^{B}_{\text{inlet}}$. The averaged results of the experiments before and after relocation are summarized in Tables 4–6. The power consumptions of these two servers are the same before and after relocation. However, the change in outlet temperatures is notable. The calculation for $E_{\text{cooling\_wasted}}$ according to Eq. (13) shows that type B server is wasting energy at a rate of 22–24 W/min while type A server is not wasting any energy due to having a proper inlet

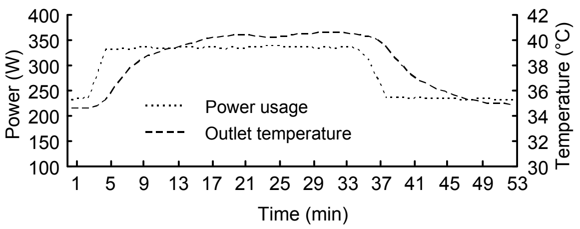

Fig. 18 Power usage and outlet temperature of type B server under experiment set 1

Type B server shows a decrease in outlet temperature, but the power consumption is the same as that before moving

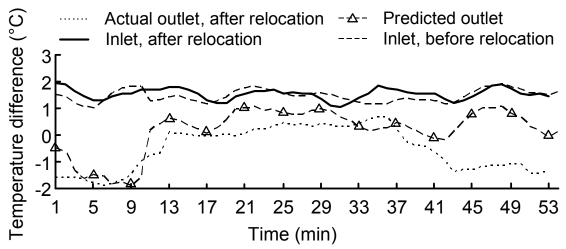

Fig. 19 Differences between type A and type B servers in terms of actual oulet temperature after relocation, predicted outlet temperature, and inlet temperatures before and after relocation under experiment set 1

The predicted difference in outlet temperatures is more accurate when the servers are running at maximum utilization

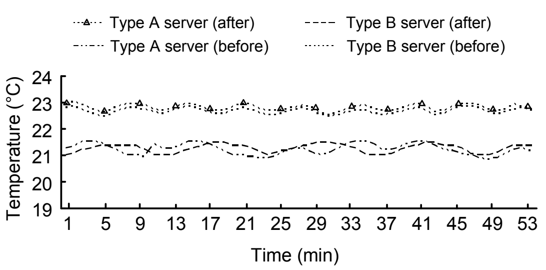

Fig. 20 Inlet temperatures of type A and type B severs before and after relocation under experiment set 1

The inlet temperatures at both locations were almost the same

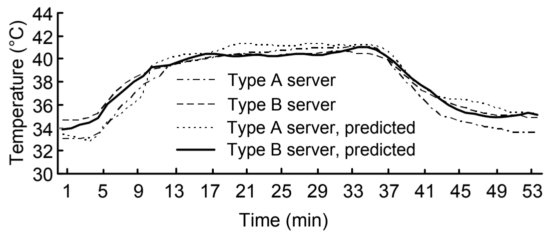

Fig. 21 Actual (after relocation) and predicted (before relocation) outlet temperatures of type A and type B servers under experiment set 1

The actual temperatures of the servers are within the predicted temperature ranges

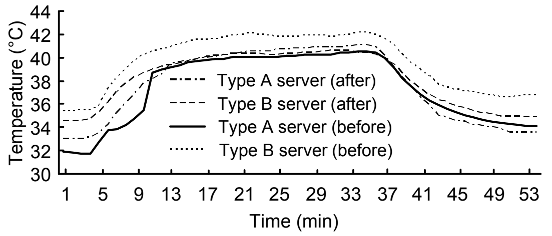

Fig. 22 Outlet temperatures of type A and type B servers before and after relocation under experiment set 1

The outlet temperatures of both types of servers are more homogeneous after relocation

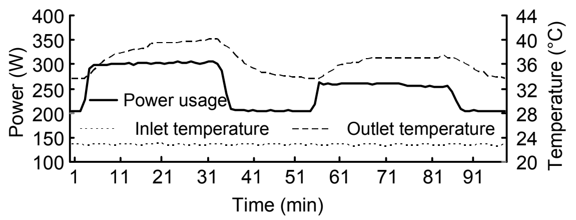

Fig. 23 Power consumption and inlet and outlet temperatures of type A server under experiment sets 2 and 3

The power consumption is the same as that before relocation, but the outlet temperature is higher after relocation

Table 4 Experimental results before relocation

Server type	Experiment set before relocation	Average inlet temperature (°C)	Average outlet temperature (°C)	Average computing power consumed (W/min)	$E^i_{computing}$/COP(T_{set}) (W/min)	$E^i_{computing}$/COP($T^i_{received}$) (W/min)	$E_{cooling\_wasted}$ (W/min)	Total energy consumed (W/min)
A	1	21.2	38.7	329.0	94.0	94.0	0.0	423.0
B	1	22.7	41.3	336.0	96.0	84.0	24.0	444.0
A	2 and 3	21.2	37.0	276.0	79.0	79.0	0.0	355.0
B	2 and 3	22.7	40.7	320.0	91.0	80.0	22.0	422.0

Table 5 Experimental results after relocation with calculations of decrease in cooling burden

Server type	Experiment set after relocation	Average inlet temperature (°C)	Average outlet temperature (°C)	Average computing power consumed (W/min)	Difference in inlet temperatures before and after relocation (°C)	Difference in outlet temperatures before and after relocation (°C)	Cooling energy saving (W/min)
A	1	22.8	39.4	328.0	1.6	0.7	11
B	1	21.2	39.6	333.0	−1.5	−1.7	24
A	2 and 3	22.7	37.2	280.0	1.5	0.2	10
B	2 and 3	21.2	39.0	321.0	−1.5	−1.7	22

Table 6 Decrease in occurrence of hotspots after relocation

Experiment set	Difference between average outlet temperatures before relocation (°C)	Difference between average outlet temperatures after relocation (°C)	Reduction in chance of hotspots after relocation (%)
1	2.6	0.2	77.0%
2 and 3	3.7	1.8	48.6%

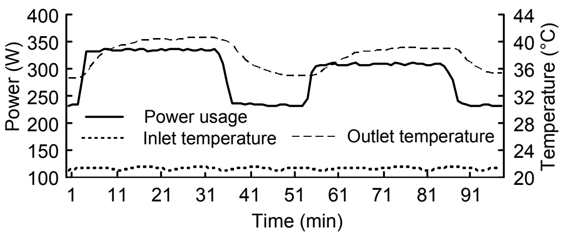

Fig. 24 Power consumption and inlet and outlet temperatures of type B server under experiment sets 2 and 3

The power consumption is the same as that before relocation, but the outlet temperature is lower after relocation

temperature. As demonstrated in Table 5, after relocation, type A server has a smaller increase in outlet temperature as compared to inlet temperature.

Therefore, type A server compensated for increase in inlet temperature and therefore $E_{cooling\_wasted}$ was reduced to half for type A server by as much as 11 W. Type B server had no $E_{cooling\_wasted}$ due to lower inlet temperature.

Therefore the saving in $E_{cooling\_wasted}$ for type B server is from 22 to 24 W (Table 4). Overall the relocation process saves over 5% of the cooling energy for the relocated servers, which is over 2.1 kW·h each for the working life of the servers. The relocation of servers brings homogeneity in outlet temperatures, which reduces the chance of hotspots by up to 77.0% (Table 6).

This can also be regarded as an improvement in cooling energy wastage. The proactive approach proposed in this paper saves the cooling energy that otherwise would be wasted on cooling a hotspot region. If the relocation algorithm is performed on the server sets after relocation, it will not predict favorable outlet temperatures from the servers. Hence, the algorithm performs well in reducing the outlet temperature and the chance of hotspots once implemented.

The total energy consumed after relocation is calculated and compared (Table 7). Since there is no major change in computing energy consumed before and after relocation, the cooling energy calculation on the basis of COP (Moore *et al.*, 2005) will not show the lowering of cooling burden due to relocation and/or decrease in outlet temperature of the servers. This is not only a limitation of COP based calculation but also cannot be easily proved through thermodynamics laws (Moore *et al.*, 2005; Tang *et al.*, 2006).

This is because the thermodynamics laws are applied on the difference between the outlet and inlet temperatures for heat calculation, instead of the intensity of the temperature. Deriving new thermal laws or thermal engineering equations for heat calculation is out of the scope of this paper. Therefore, the cooling energy savings given in Table 5 are subtracted from the total energy consumed after relocation to mark the benefits of comparatively low outlet temperature of the relocated server and the homogeneity of the outlet temperatures of the relocated servers. This is demonstrated in Table 7.

Recommendations for server relocation: Based upon the experiments and results, we present the best practices for server relocation inside data centers. The following recommendations will help the data center manager to identify, analyze, predict, and perform a relocation to save energy and to minimize cooling energy wastage:

1. The chances of legitimate relocation requirement are higher between a set of servers where a subset of the servers has higher inlet and outlet temperatures than another subset when all servers are idle.

2. If the server subsets have heterogeneous processors, then check if the servers with hotter outlets in idle state are also consuming larger energy than the servers in other subset(s). This will add the chance of relocation as the higher outlet temperature in idle state may give rise to hotspots when servers are used.

3. As a next step, the server subsets should be marked and put to experimental test load. The test load boosts the server's utilization to various levels. This step can be skipped if the daily usage of server CPU is available covering a reasonable time. However, this step is necessary if the servers are to be mounted for the first time. Data centers seldom keep the per minute performance records of thousands of

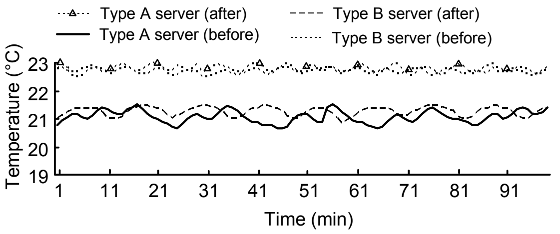

Fig. 25 Inlet temperatures of type A and type B servers before and after relocation under experiment sets 2 and 3
The inlet temperature does not change after relocation at both locations

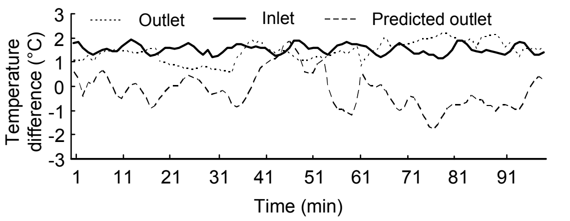

Fig. 26 Differences of inlet, outlet, and predicted outlet temperatures of type A and type B servers under experiment sets 2 and 3
The outlet temperatures of both servers are more homogeneous after relocation. The fall in experiment set 3 is due to type B server dissipating more heat than type A server even after relocation

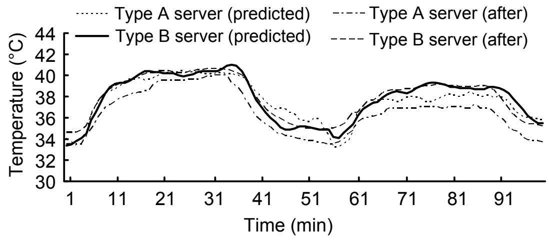

Fig. 27 Outlet and predicted outlet temperatures of type A and type B servers under experiment sets 2 and 3
The actual maximum temperatures of the servers are within the predicted temperature ranges

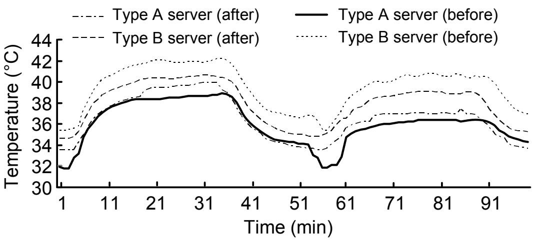

Fig. 28 Outlet temperatures of type A and type B servers before and after relocation under experiment sets 2 and 3
Both are more homogeneous after relocation

Table 7 Comparison of total energy consumption before and after relocation

Server type	Experiment set	Average computing power consumed (W)	$E^i_{\text{computing}}$/ COP(T_{set}) (W/min)	$E_{\text{cooling\_wasted}}$ (W/min)	Cooling energy saving (W/min)	Total energy consumed after relocation (W/min)	Total energy consumed before relocation (W/min)
A	1	328	94	11	11	422	423
B	1	333	95	0	24	404	444
A	2 and 3	280	80	10	10	360	355
B	2 and 3	321	91	0	22	390	422

servers and keep the aggregated records instead. Therefore, it is better to perform the experiment sets when there is an indication of inlet/outlet temperature variance.

4. If there is more than one server in each subset, the relocation algorithm should be applied between all the combinations of paired servers by taking one server from each subset. The relocation algorithm gives the predicted temperatures of the pair of servers. This can reduce the complexity of comparing servers.

5. Each server should be identified with the highest predicted change in inlet and outlet temperatures. To make server pairs, a good indicator is that these two servers use the same amount of maximum electricity.

6. Relocation is favorable if a small change in inlet temperature can bring more change in outlet temperature. The ratio of CPU time frequency to power consumed is a supporting value for the predicted temperatures. A server having a low value of CPU (MHz/W) will dissipate more heat than a server with higher CPU (MHz/W) if the maximum power usages of both servers are equal to each other.

7. Note that at higher inlet temperature, the outlet temperatures rise at a higher rate than at lower inlet temperature for the same server. The outlet temperatures of the relocated pair of servers should be more homogeneous and the post relocation conditions defined earlier in the problem statement should be satisfied.

7 Conclusions

In this paper we present an energy model to represent the cooling energy wastage by inlet temperature variations. The rise in inlet temperature can lead to hotspot causing increased outlet temperature of the data center servers. This increases the PUE of the data center due to energy wastage in cooling. This is highlighted through the data center energy modeling presented in this paper. The server relocation algorithm can successfully optimize the location of each server to lower the extra burden on the cooling mechanism. The proposed approach can lower the chance of hotspots and improve the cooling energy wastage by over 77%, lower the cooling load through thermal-aware server relocation, leading to a energy saving of 2.1 kW·h throughout the service time span of relocated servers. In short, the particular contributions of this paper are:

1. An energy model is presented to explain the effect of rise in inlet temperature of each server and the effect of this upon the total power consumption of the data center.

2. A proactive algorithm for server relocation is proposed to (1) avoid hotspots, (2) lower the peak temperature of hot air from the outlets of the relocated server set, and (3) homogenize the outlet temperatures of the set of relocated servers. These will result in lower cooling load, avoidance of hotspots, ensuring equipment safety, and help maintain green data centers.

3. Recommendations or best practices for server relocation are presented which will help the data center manager to identify, analyze, and perform a relocation to save power and to minimize cooling power wastage.

References

ABB, 2013. Efficient DC Power Supply for Data Centers. Available from http://www.electricalreview.co.uk/features/9475-efficient-dc-power-supply-for-data-centres

Ahuja, N., 2012. Datacenter power savings through high ambient datacenter operation: CFD modeling study. Proc. 28th Annual IEEE Semiconductor Thermal Measurement and Management Symp., p.104-107. [doi:10.1109/STHERM.2012.6188833]

Ahuja, N., Rego, C., Ahuja, S., *et al.*, 2011. Data center efficiency with higher ambient temperatures and optimized cooling control. Proc. 27th Annual IEEE Semiconductor Thermal Measurement and Management Symp., p.105-109. [doi:10.1109/STHERM.2011.5767186]

ASHRAE, 2011. 2011 Thermal Guidelines for Data Processing Environments—Expanded Data Center Classes and Usage Guidance. Available from http://ecoinfo.cnrs.fr/IMG/pdf/ashrae_2011_thermal_guidelines_data_center.pdf

Banerjee, A., Mukherjee, T., Varsamopoulos, G., *et al.*, 2010. Cooling-aware and thermal-aware workload placement for green HPC data centers. Int. Green Computing Conf., p.245-256. [doi:10.1109/GREENCOMP.2010.5598306]

Banerjee, A., Mukherjee, T., Varsamopoulos, G., *et al.*, 2011. Integrating cooling awareness with thermal aware workload placement for HPC data centers. *Sustain. Comput. Inform. Syst.*, **1**(2):134-150. [doi:10.1016/j.suscom.2011.02.003]

BBC, 2014. Energy Transfer and Storage. Available from http://www.bbc.co.uk/bitesize/ks3/science/energy_electricity_forces/energy_transfer_storage/revision/1/

Corradi, A., Fanelli, M., Foschini, L., 2011. Increasing cloud power efficiency through consolidation techniques. Proc. IEEE Symp. on Computers and Communications, p.129-134. [doi:10.1109/ISCC.2011.5984005]

Huck, S., 2011. Measuring Processor Power TDP vs. ACP. Available from http://www.intel.com/content/dam/doc/white-paper/resources-xeon-measuring-processor-power-paper.pdf

Jonas, M., Varsamopoulos, G., Gupta, S.K.S., 2007. On developing a fast, cost-effective and non-invasive method to derive data center thermal maps. Proc. IEEE Int. Conf. on Cluster Computing, p.474-475. [doi:10.1109/CLUSTR.2007.4629269]

Jonas, M., Varsamopoulos, G., Gupta, S.K.S., 2010. Non-invasive thermal modeling techniques using ambient sensors for greening data centers. Proc. 39th Int. Conf. on Parallel Processing Workshops, p.453-460. [doi:10.1109/ICPPW.2010.67]

Koomey, J., 2011. Growth in Data Center Electricity Use 2005 to 2010. Analytics Press, Oakland, CA.

Kusic, D., Kephart, J.O., Hanson, J.E., *et al.*, 2009. Power and performance management of virtualized computing environments via lookahead control. *Cluster Comput.*, **12**(1):1-15. [doi:10.1007/s10586-008-0070-y]

LD Didactic Gmbh. Converting Electrical Energy into Heat Energy—Measuring with the Joule and Wattmeter. LD DIDACTIC GmbH, Germany.

Lee, E.K., Kulkarni, I., Pompili, D., *et al.*, 2012. Proactive thermal management in green datacenters. *J. Supercomput.*, **60**(2):165-195. [doi:10.1007/s11227-010-0453-8]

Liu, Z., Chen, Y., Bash, C., *et al.*, 2012. Renewable and cooling aware workload management for sustainable data centers. *ACM SIGMETRICS Perform. Eval. Rev.*, **40**(1):175-186. [doi:10.1145/2318857.2254779]

Masiero, M., 2012. CPU Charts 2012: 86 Processors from AMD and Intel, Tested. Tom's Hardware.

Mersenne Research, Inc., 2012. Great Internet Mersenne Prime Search (GIMPS). Available from http://www.mersenne.org/freesoft/

Moore, J., Chase, J., Ranganathan, P., *et al.*, 2005. Making scheduling "cool": temperature-aware workload placement in data centers. Proc. USENIX Annual Technical Conf., p.61-75.

Mukherjee, T., Banerjee, A., Varsamopoulos, G., *et al.*, 2009. Spatio-temporal thermal-aware job scheduling to minimize energy consumption in virtualized heterogeneous data centers. *Comput. Netw.*, **53**(17):2888-2904. [doi:10.1016/j.comnet.2009.06.008]

Rodero, I., Lee, E.K., Pompili, D., *et al.*, 2010. Towards energy-efficient reactive thermal management in instrumented datacenters. Proc. 11th IEEE/ACM Int. Conf. on Grid Computing, p.321-328. [doi:10.1109/GRID.2010.5698002]

Rodero, I., Viswanathan, H., Lee, E.K., *et al.*, 2012. Energy-efficient thermal-aware autonomic management of virtualized HPC cloud infrastructure. *J. Grid Comput.*, **10**(3):447-473. [doi:10.1007/s10723-012-9219-2]

Sansottera, A., Cremonesi, P., 2011. Cooling-aware workload placement with performance constraints. *Perform. Eval.*, **68**(11):1232-1246. [doi:10.1016/j.peva.2011.07.018]

Tang, Q., Mukherjee, T., Gupta, S.K.S., *et al.*, 2006. Sensor-based fast thermal evaluation model for energy efficient high-performance datacenters. Proc. 4th Int. Conf. on Intelligent Sensing and Information Processing, p.203-208. [doi:10.1109/ICISIP.2006.4286097]

Tang, Q., Gupta, S.K.S., Varsamopoulos, G., 2007. Thermal-aware task scheduling for data centers through minimizing heat recirculation. Proc. IEEE Int. Conf. on Cluster Computing, p.129-138. [doi:10.1109/CLUSTR.2007.4629225]

Tu, C.Y., Kuo, W.C., Teng, W.H., *et al.*, 2010. A power-aware cloud architecture with smart metering. Proc. 39th Int. Conf. on Parallel Processing Workshops, p.497-503. [doi:10.1109/ICPPW.2010.73]

U.S. Environmental Protection Agency, 2007. Report to Congress on Server and Data Center Energy Efficiency Public Law 109-431. Available from http://www.energystar.gov/ia/partners/prod_development/downloads/EPA_Datacenter_Report_Congress_Final1.pdf?6133-414f

VMware Inc., 2009. VMware vSphere Basics ESXi 5.0 vCenter Server 5.0 (in Chinese). Available from http://pubs.vmware.com/vsphere-50/topic/com.vmware.ICbase/PDF/vsphere-esxi-vcenter-server-50-basics-guide.pdf

Wang, L., von Laszewski, G., Dayal, J., *et al.*, 2009a. Thermal aware workload scheduling with backfilling for green data centers. Proc. IEEE 28th Int. Performance Computing and Communications Conf., p.289-296. [doi:10.1109/PCCC.2009.5403821]

Wang, L., von Laszewski, G., Dayal, J., *et al.*, 2009b. Towards thermal aware workload scheduling in a data center. Proc. 10th Int. Symp. on Pervasive Systems, Algorithms, and Networks, p.116-122. [doi:10.1109/I-SPAN.2009.22]

Wang, L., Khan, S.U., Dayal, J., 2012. Thermal aware workload placement with task-temperature profiles in a data center. *J. Supercomput.*, **61**(3):780-803. [doi:10.1007/s11227-011-0635-z]