THERMAL PREDICTION MODELS FOR VIRTUALIZED DATA CENTER SERVERS BY USING THERMAL-PROFILES

Muhammad Tayyab Chaudhry[1], Chun Yong Chong[2], T.C. Ling[3],
Saim Rasheed[4], Jongwon Kim[5]

[1,2,3] University of Malaya, Kuala Lumpur, 50603, Malaysia,

[4] King Abdulaziz Univeristy, Jeddah, 21589, Kingdom of Saudi Arabia,

[5] Gwangju Institute of Science and Technology, Gwangju, 500-712, South Korea.

Email: Corresponding Author - mtayyabch@yahoo.com[1], cychong@um.edu.my[2] , tchaw@um.edu.my[3] ,
srahmed@kau.edu.sa[4] , jongwon@nm.gist.ac.kr[5]

ABSTRACT

The energy dissipated as heat for each utilization level of a data center server is empirically measured and stored as the thermal-profile. These thermal-profiles are used to predict the outlet temperatures of the related servers for current and future utilization. The predicted outlet temperature is an important parameter for energy efficient thermal-aware workload scheduling and workload migration in green data centers. This paper presents three models for outlet temperature prediction on virtualized data center servers based on thermal-profile. The best case scenario managed to predict the outlet temperature with a negligible error of 0.3 degree Celsius.

Keywords: data center; servers; monitoring; virtualization; thermal-profile; thermal-prediction;

1.0 INTRODUCTION

Monitoring systems in data centers look-over the environment and performance of tens and hundreds of servers periodically. There are various parameters (e.g., temperature and utilization) monitored for each server. This data is required by data center infrastructure management (DCIM) [1] tools and workload scheduling systems [2] to achieve energy efficient and proficient utilization on data center servers. Apart from the idle energy consumption, the total energy consumed and dissipated as heat by each server is increased for every utilization level increment and vice versa. The energy usage of a virtualized server involves the virtualized instances of operating systems called virtual machines (VMs). From now on, the word *server* is used interchangeably with *virtualized server.*

As long as the hardware configuration of a server is not altered, each heterogeneous server will dissipate a different but certain amount of heat for each discrete utilization level. This symbolic heat is empirically measured and stored as a thermal-profile for each server. The thermal-profile is used to predict the outlet temperature of a server by a given value of CPU usage through thermal-prediction modeling. A thermal prediction model eliminates at least one of the parameters e.g., outlet temperature, from the data center monitoring system without the loss of performance and accuracy. Similarly, the number of thermal sensors used for outlet temperature monitoring is reduced. The ability of a monitoring system to generate accurate thermal predictions offline, makes the prediction to become an essential parameter for thermal-aware workload scheduling and thermal-aware workload migration for load balancing.

Based upon thermal-profile, this paper presents multiple prediction models to predict outlet temperature of data center servers along with a matrix comparison among these models. The test results show that the outlet

1

temperature is accurately predicted by given values of CPU utilization with an average error of 0.3 degree Celsius. Thus, the outcome of workload scheduling and/or workload migration in terms of outlet temperature is predicted and used to minimize the overall heat generation in data centers. This result benefits in cooling energy saving [3] and thus contributes to the establishment of green data centers. The rest of the paper is arranged such that the Section 2 describes the up-to-date literature review and Section 3 reports the background and preliminaries of prediction modeling. Three different prediction models are presented in Section 4 with a thorough discussion, evaluation and comparison. Section 5 demonstrates the experimental setup and detailed results followed by a conclusion Section 6.

2.0 RELATED WORK

A server based power consumption profiling was performed by [4] and used as a reference to estimate the outlet temperature for workload scheduling. Different workload scheduling techniques were simulated based on a small data center with homogeneous uni-processor servers. However the profiling was performed in a limited range of processor frequency. It focused on heat recirculation [5] and energy consumption minimization for data centers. The complexity of the proposed optimization model will increase with the heterogeneity of the servers. The proposed thermal model lacked of linkage between the power consumption, server utilization level and the resultant heat. A basic thermal-profiling was performed in [6] to analyze the data center servers for thermal-stress with reference to inlet temperature change. The hotspots were predicted with the help of thermal-profiles of the servers. This was useful for cooling energy saving and establishment of green data centers. The work was further extended in [7] by empirically analyzing the effect of inlet temperature over outlet temperatures of the servers. By using the thermal-profiles of the servers, as proposed in our paper, the outlet temperature can be accurately predicted.

Thermal-profiles for different tasks were created and used for scheduling purposes in [8]. Tasks were categorized and scheduled either on cold or hot servers by using a thermal map. This map was an abstraction of the air temperature distribution across data center hall. By analyzing a snapshot of thermal map at a given time, the future changes in thermal map were predicted on the basis of thermal-profiles of the tasks at hand. However such approach was limited by the prerequisite of thermal-profile of the task. In addition, two homogenous tasks running over heterogeneous servers with the same starting temperature could reach different temperature after the completion of the tasks. Task-based thermal-profiles are often server specific and cannot be used for server monitoring across data center. By using the server based thermal-profiles, the outcome of running same task over different servers can be predicted accurately for server selection.

Rodero et al. [9] demonstrated that heat transferred from active servers caused an undesirable inlet temperature increment to neighboring servers. The authors confirmed that the rise in inlet temperature increases the outlet temperature of the server. Therefore, a HPC task profiling for servers was used and the server profiles were used for VM scheduling. However, they ignored the thermal aspect of power consumption by active server while making scheduling decisions. Thermal-profiles could be created for servers by using thermal-benchmark and used for monitoring as well as scheduling as described in latter sections. In addition to conventional onboard thermal sensors, the wireless thermal sensors were widely used for temperature monitoring in data centers. There were several algorithms for efficient placement of wireless sensors. These sensors were not applied in close proximity to each to reduce the packet loss and retransmission rate [10] and to avoid the sensing errors and battery depletion.

A server monitoring approach is proposed by [11, 12] by considering a window based distributed approach for cloud applications. This approach considered maximum value as threshold for a parameter being monitored. A time window was used to avoid unnecessary alarms due to short term threshold violations at different nodes. Each monitoring node kept a time window to verify the continuous violation of threshold before alarming the central manager node. This approach suffered from drawbacks e.g., skipping some violation events when multiple nodes were monitored in parallel and adding extra load for processing multiple parameters for each node. The processing load over the monitoring node and sending generated data over the network could be

2

Malaysian Journal of Computer Science. Vol. 29(1), 2016

reduced by lowering down the number of parameters being monitored without the loss of performance and accuracy.

Junmei et al. [13] proposed to place a few thermal sensors across server racks. It effectively monitored the temperature across data center. If sensors were efficiently placed, then they were effectively controlling the thermostat of data center cooling mechanism. Effective control of cooling leads to data center energy saving in cooling and the chances to generate hotposts (undercooling) and coldspots (overcooling) could be avoided. In other words, each sensor could accurately sense the temperature from multiple servers at optimum position. The chances of generating hotspots could be predicted and the cooling is temporally enhanced. The authors used the basic thermal-profile of data center servers by noting the lowest and highest outlet temperatures for each server. These thermal-profiles were used to evaluate the maximum cooling temperature across data center while avoiding hotspots. The lesser number of thermal sensors caused a certain amount of acceptable sensing error. At the same time, this error was minimized through repeated simulations for random outlet temperature distribution across all servers.

A sensor placement approach was reported by Xiaodong et al. [14] for detecting the hot servers through a monitoring system. They proposed computational fluid dynamic (CFD) analysis to find the optimum sensing region using wireless thermal sensors. A single sensor placed in optimum region was able to detect the presence of hot server inside the sensing region. They identified that one of the causes of hot server outlet was a rise in inlet temperature. However, the drawback of CFD based approaches was the complexity of implementation and the time to complete the simulations. Furthermore, the thermal sensor placement approaches can only detect the hot servers after a thermal threshold is violated. It was not possible to pin point a hot server due to number of scattered wireless sensors in each rack. On the other hand, a data center server with thermal-profile approach could be utilized for proactive scheduling as well as pin pointing the hot servers.

Thermal monitoring systems [15, 16] that relied upon thermal cameras could progress efficiently by detecting the hotspots from a distance, without the assistance of thermal sensors. Only three thermal cameras were required to accurately identify and locate the coordinates of the hotspot. The characteristics of hotspots were extracted and propagated over network, instead of sending multimedia images. Thus [15, 16] presented a network efficient thermal monitoring systems for data centers. However, thermal cameras had a drawback of huge cost of setup. Furthermore, multiple cameras were needed to detect the location of hotspots. If a hotspot was detected by a set of thermal cameras, an additional mechanism was required to find the exact server or set of servers which are inside the hotspot. A primitive mechanism of thermal image based thermal-profiles was used to identify different gradients of a hotspots in [15]. However it was not possible to predict the outlet temperatures of any server through thermal image based thermal-profiles without server usage data. Therefore the chances of using this approach for thermal-aware VM scheduling were minimum because of the shortage of temperature prediction capability.

The thermal state monitoring systems in data centers utilize a thermal-prediction to identify and verify the thermal conditions offline based on only two parameters: the inlet temperature and server utilization. This is achievable through outlet temperature calculation by using thermal-profiles. The traditional approaches of thermal monitoring by using scattered wireless thermal sensors are not able to correctly and precisely identify the outlet temperature of a particular server in situations such as hotspots. This is because every server may not have a thermal sensor over the outlet. The outlet temperature prediction is required for thermal-aware workload scheduling such as VM placement. A thermal-profile based outlet temperature prediction approach is able to reduce the load of monitoring nodes, the complexity of outlet temperature prediction and provide an accurate temperature input to thermal-aware workload schedulers. The fewer number of thermal sensors are required comparatively for thermal monitoring when the thermal prediction models are used to calculate and predict the outlet temperature.

3

Malaysian Journal of Computer Science. Vol. 29(1), 2016

3.0 BACKGROUND

Thermal-profiles are created by using customized scripts (i.e., thermal-benchmark) to manipulate multiple VM [6]. Thermal-benchmarks for data center servers are used to create thermal-profiles by manipulating the server-utilization as shown in Fig. (1). A thermal-profile is represented in tabular form with each tuple having three elements and represented as *{CPU(MHz), Inlet Temperature (Celsius), Outlet Temperature (Celsius)}.* This paper considers the virtualized servers with multiple VMs for thermal-benchmark. Each VM executes a CPU intensive benchmark instance.

The thermal benchmark for data center servers induces a workload pattern to imitate the real life computational load. This is performed by manipulating VM operations such as suspend and power-on with respect to time intervals. In this way, different workload patterns for the servers are generated as *spike, linear* and *step* as shown in Fig. (1). The *spike* consists of multiple VM powered-On at the same time and suddenly raising the server utilization to maximum. This is visible as a *spike* at left side (i.e., at 1-33 minutes) of Fig. (1). The *spike* ends when all the VM are suspended at the same time. In this paper, stable results were obtained when the servers are kept at full utilization for 30 minutes during each instance of thermal-benchmarks.
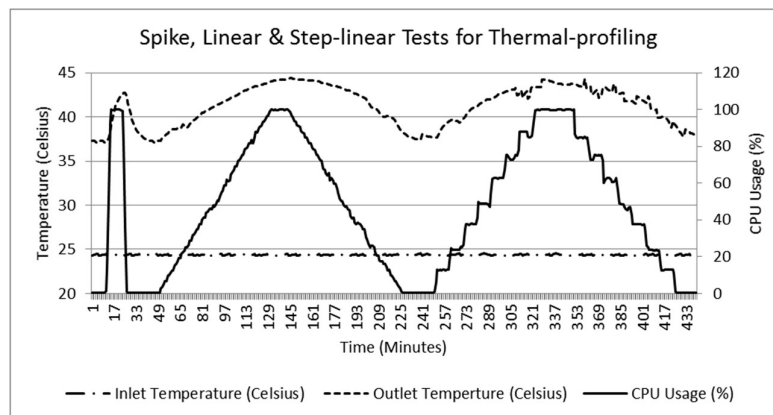

Fig. 1 Spike, step and linear thermal benchmarks used for thermal-profile.

The *linear* increase in server utilization is achieved by powering-on one VM at a time and using dynamic frequency scaling (DFS) to increase the frequency of virtual core from minimum to maximum. This is visible in the middle graph (i.e., at 145-255 minutes) of Fig. (1). By repeating this periodically for each VM until the utilization of the server reaches to maximum, the linear thermal-benchmark can raise the outlet temperature to maximum during a certain time. Then, after some predefined time interval at maximum utilization, the linear thermal-benchmark script lowers the server utilization by DFS of each VM from maximum to minimum and suspends each VM.
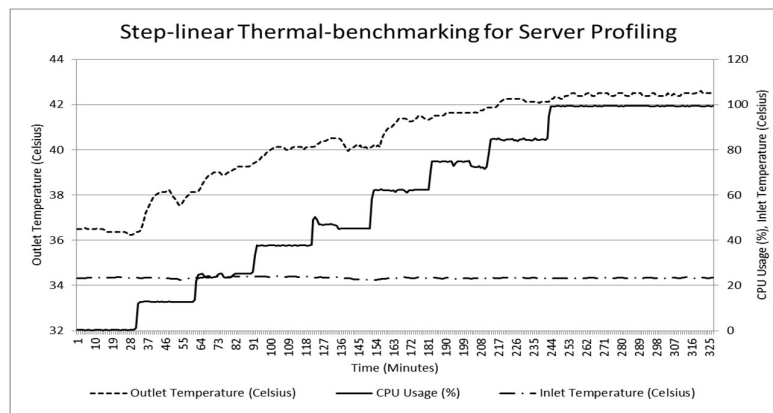

Fig. 2. Step-linear benchmarking for thermal-profiling of servers.

4

The corresponding outlet temperature follows the server utilization graph (i.e., at 145-255 minutes) as shown in Fig. (1). The *linear* outlet temperature is higher than *spike* at full utilization. This is because the total execution time for *spike*-based thermal-benchmark is many times lower than the *linear* thermal-benchmark. In Fig. (1) the *step linear* benchmark is shown at the right-side (i.e., at 257-337 minutes) of graph and each VM is powered-on after a predefined time interval to reach the maximum utilization of the server. After the predefined time interval when the server reaches maximum utilization, the VMs are suspended one after another in predefined time intervals to form a step-down type of pattern (i.e., at 337-417 minutes). The outlet temperature follows the execution pattern of the related server.

This paper utilizes the average maximum outlet temperature at each step (utilization level) to create the thermal-profile of the related server. Each tuple or slot in thermal-profile is related to a particular utilization level and the corresponding inlet temperature and outlet temperature. In order to record the outlet temperature at different utilization levels of the server, the *step-linear* thermal-benchmarking script powers-on each VM after the stable interval of 30 minutes to obtain a step like utilization pattern of the server. This is demonstrated in Fig. (2) for one of the test bed servers. The utilization level is represented by percentage in Fig. (1). For more accuracy, the server utilization is expressed in MHz as demonstrated in Fig. (2) and the result is used in the following sections. This paper shows that the outlet temperature of a server is calculated through inlet temperature and CPU usage by utilizing the thermal-profile. This reduces the need for monitoring the outlet temperature. Finally, the accurate temperature predictions can be generated by using different prediction models and justified by the concept of prediction modeling. In next section, three novel prediction models are presented for outlet temperature prediction of data center servers on the basis of thermal-profiles. In later section, the enhanced thermal benchmarks are used to verify the proposed prediction models. A performance matrix to evaluate the prediction model is also proposed in this paper with the following criteria:

- Accuracy of prediction with respect to actual outlet temperature. This is also represented by percentage error.
- Causes of prediction error that determines the degree of robustness of the model.
- Complexity of Calculation; determines the amount of computations and logic required to implement the model.
- Output of the prediction model is in the form of a range but an exact value with low prediction error value is highly recommended for thermal monitoring and workload scheduling.

4.0 THERMAL-PROFILE BASED PREDICTION MODELING ARCHITECTURE

This section explains the architecture of the proposed technique. For every server *i,* the CPU usage in MHz can be represented as CPU_{MHz}^i and inlet temperature as T_{inlet}^i then the outlet temperature T_{oulet}^i can be calculated offline. As demonstrated in Fig. (3), the proposed system gathers the CPU Usage, power consumption and inlet temperature of the data center server and stores it to database. The outlet temperatures of the servers are calculated through thermal-profiles.

In this paper, it is proposed to use the nett increase in outlet temperature $T_\Delta^i = (T_{oulet}^i - T_{inlet}^i)$ for each slot instead of keeping the values of inlet temperature and outlet temperature. This was empirically tested and proved that at different stable average inlet temperatures but same level of server utilization, the value of T_Δ^i remains unchanged for a particular server. Therefore, the outlet temperature of the server can be calculated at run time by using inlet temperature and utilization level. If TP^i is the thermal-profile of server *i* then the columns of TP^i can be represented by $TP^i \rightarrow CPU_{MHz}^i[n]$, and $TP^i \rightarrow T_\Delta^i[n]$ for a *n* slot thermal-profile. Outlet temperature prediction can be performed on the basis of multiple considerations and among them the intended use of the prediction temperature is the most important. The temperature prediction can be performed according to current CPU usage in MHz or the expected MHz in near future. There are three prediction models presented in this paper for virtualized servers according to their most likely applications. These are listed as follows:

- Worst-case Prediction Model (WCPM)

5

Malaysian Journal of Computer Science. Vol. 29(1), 2016

- Optimistic Prediction Model (OPM)
- Advanced Optimistic Prediction Model (EOPM)

The WCPM is the basic model which is enhanced to OPM. The EOPM is enhanced version of OPM.

One of the advantages of using thermal prediction models is the minimum number of thermal sensors required to implement it and still providing the up to date outlet temperature predictions. In order to understand this, consider if the CPU usage variation, caused by workload variation, for any server. This situation requires the presence of a stream of outlet temperatures from thermal sensors attached with each server. On the other hand, the thermal-profile based outlet temperature prediction models presented in this paper require only the inlet temperature and CPU utilization to successfully predict the outlet temperature with minimum error. Thus the thermal sensors deployed for outlet temperature sensing can no longer be needed. If the inlet temperature does not vary much as compared to outlet temperature of the servers, then the deployment of outlet temperature sensors is not required. Last but not least, the thermal prediction modeling, minimizes the inherent limitations of wireless sensor networks as discussed in [10]. In the following, each model is discussed one by one.

4.1 Worst-case Prediction Model (WCPM)

The worst-case prediction is the maximum possible outlet temperature of the server in the nearest slot of thermal-profile according to intended CPU MHz. The worst-case predicted outlet temperature $T_{worst\_case\ oulet}^i$ at server CPU utilization CPU_x^i can be given as in equation (1).

$$T_{WCPM\ oulet}^i = (TP^i \rightarrow T_\Delta^i[n]) + T_{inlet}^i + \beta \tag{1}$$

Where $CPU_x^i \leq TP^i \rightarrow CPU_{MHz}^i[n]$ and β is the prediction error

The value of β is maximum for WCPM because it estimates the maximum outlet temperature for every slot of thermal-profile according to given CPU utilization. Variation is inlet temperature is another reason of rise in the value of β.
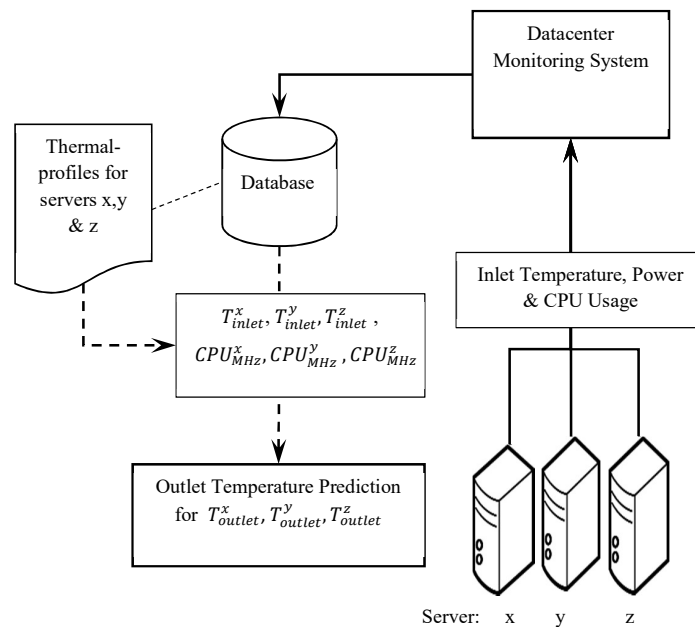

Fig. 3 Outlet temperature prediction using thermal-profile, inlet temperature and CPU usage for each server.

6

The most suitable application of WCPM is for server preference for batch scheduling of VM. When there are multiple VM with combined load to completely occupy a slot of thermal-profile, there is a less chance of underutilization of a server as compared to latter example given above. The value of β is also lower in this case. Another application of WCPM is for decision making of static scheduling of VM. If a VM or a batch of VM is to be scheduled permanently on a server, then the WCPM can give an overview of the highest possible outlet temperature in future. Last but not least, the WCPM gives a margin for backfilling of additional VM that can be scheduled within a slot of thermal-profile at a scheduling instance. The value of β can be used as backfilling margin. In future, a backfilling based thermal-aware scheduler will be presented for VM.

4.2 Optimistic Prediction Model (OPM)

The short coming of previous model regarding overestimation of outlet temperature is normalized by OPM. This model uses interpolation of $TP^i \rightarrow T_\Delta^i$ in order to bring the value of β as lower as possible. In order to perform interpolation, the fine grained information for per slot rise in outlet temperature compared to idle state is introduced in this section. The WCPM can be updated by considering that all the non-idle state slots of thermal-profile have a greater value of T_Δ^i than the idle state. Therefore for each slot, fine grained nett outlet temperature $T_{\Delta\Delta}^i$ can be defined as

$$TP^i \rightarrow T_{\Delta\Delta}^i[n] = (TP^i \rightarrow T_\Delta^i[n]) - (TP^i \rightarrow T_\Delta^i[0]) \qquad (2)$$
$$\text{Where } n > 0$$

Thus $T_{\Delta\Delta}^i[n]$ can be added in thermal-profile as a column and equation (1) can be re-written as

$$T_{WCPM\ oulet}^i = (TP^i \rightarrow T_{\Delta\Delta}^i[n]) + (T^{\ i} \rightarrow T_\Delta^i[0]) + T_{inlet}^i + \beta \qquad (3)$$
$$\text{Where } n \geq 0 \text{ and } CPU_x^i \leq TP^i \rightarrow CPU_{MHz}^i[n]$$

It should be noted that $T_{\Delta\Delta}^i[0] = 0$. The outlet temperature $T_{optimistic\ oulet}^i$ calculated through optimistic prediction model can be given as in equation (4)

$$T_{OPM\ oulet}^i = \left((TP^i \rightarrow T_{\Delta\Delta}^i[n]) * \frac{CPU_x^i}{TP^i \rightarrow CPU_{MHz}^i[n]} \right) + (TP^i \rightarrow T_\Delta^i[0]) + T_{inlet}^i + \beta \qquad (4)$$
$$\text{Where } n \geq 0 \ \text{ and } CPU_x^i \leq TP^i \rightarrow CPU_{MHz}^i[n] \text{ and } T_{\Delta\Delta}^i[0] = 0$$

The presence of β is due to the fluctuation of inlet temperature and the fact that the server may take some time to get heated and reach the predicted outlet temperature. Another reason for prediction error is the sharp rise and fall in CPU usage in amount exceeding the range of a single slot of thermal-profile. The experimental results prove that the value of β is significantly lowered in this model than the WCPM due to interpolation. In fact this paper recommends the use of OPM for real life monitoring of data center.

The OPM is hotspot safe because at maximum utilization in each slot of thermal-profile, the predicted outlet temperature from WCPM and OPM are almost the same. The most suitable application of OPM is the thermal anomaly detection across data center servers. Since this model uses the interpolation, the predicted temperature is calculated on the most recent CPU utilization of the server and is therefore faster in detecting thermal anomalies with error range of β. In future work, a fast and accurate thermal anomaly detection and management system for virtualized data centers will be presented. Also a more sophisticated and detailed work will be presented on thermal-benchmarks. The OPM can be used as a semi-offline thermal monitoring tool for servers as it can accurately calculate the outlet temperature from the given CPU usage and inlet temperature.

OPM model is not suitable for backfilling as it does not provide much margin to deploy a VM. Therefore the OPM has limited application for proactive thermal-aware scheduling of VM. Despite the short comings, the OPM provides the most up to date outlet temperature with low error. Thus it eliminates the need for a stream of actual outlet temperature readings because of low β. This paper shows that the OPM is successfully applicable

7

Malaysian Journal of Computer Science. Vol. 29(1), 2016

to virtualized servers and virtualized desktops. The OPM can be readily used with reactive thermal-aware scheduling algorithms for VM as well as for the prediction based, proactive thermal-aware scheduling.

4.3 Extended Optimistic Prediction Model (EOPM)

A common shortfall of WCPM and OPM is the presence of β which makes the predicted value as doubtful. A solution is to allow a range of β which can also be used for fine grained proactive scheduling of individual VM. The experimental results prove that the actual outlet temperature of a server is always within a range of β through EOPM. This range is always within the rounded up and rounded down values of the predicted temperature by OPM. The rounded up value of WCPM is used to limit the upper boundary of prediction. In other words, according to experimental results, the following is always true

$$RoundDown(T^i_{OPM\ oulet}, 0) \leq T^i_{OPM\ oulet} \leq RoundUp(T^i_{OPM\ oulet}, 0) \ \leq \ RoundUp(T^i_{WCPM\ oulet}, 0) \qquad (5)$$

It should be noted that the EOPM always consider the whole number values for rounded values of $T^i_{OPM\ oulet}$. Previously [9] considered the temperature values in whole numbers. The EOPM combines the benefits of WCPM and OPM and has less or no noticeable disadvantages except the complexity of calculations. The best application of EOPM is to remove the prediction error caused by inlet temperature variation. The EOPM is represented by two curves: *EOPM (L)* and *EOPM (U)*. The former is the rounded down OPM and the latter is the rounded up OPM.

Another application is to verify the potential of backfilling as given by WCPM. This is because the WCPM sometimes gives an overestimated value of predicted outlet temperature of the server. But EOPM verifies how much of backfilling capacity with-in the range of current slot of server thermal-profile is available. Another application of EOPM is to verify the prediction accuracy of OPM. Since the actual outlet temperature is always according to equation (5), therefore outlet temperature by OPM can be accurately verified. It is possible that the rounded up value of $T^i_{OPM\ oulet}$ becomes equal to the $T^i_{WCPM\ oulet}$ but it does not exceed. A shortcoming of EOPM is the excess of data it produces. As comparing to previous two models, the EOPM data consists of three streams of rounded values as shown in equation (5). But this model provides a fairly accurate range for β which is lower than that of WCPM. The WCPM always gives a value that is more than the actual outlet temperature whereas the EOPM gives a +- range or outlet temperature.

Table (1) presents a comparison of the prediction models presented in this paper. Each model has some drawbacks and best applications. The prediction error is related to overestimation and underestimation and Table (1) shows the absolute maximum error percentage for each thermal-prediction model proposed in this paper. As shown in Table (1), the EOPM has the least error. This is because this model provides a lower limit and upper limit range of outlet temperature and therefore is the most accurate of all models. While WCPM usually overestimates the outlet temperature. But WCPM and OPM models output a single valued outlet temperature prediction. Among these two models, OPM is more accurate because it uses interpolation.

Table 1 Comparison matrix of all prediction models

Prediction Model	Maximum Prediction Error (%)	Causes of Prediction Error	Complexity of Calculation	Output	Thermal-aware Applications
WCPM	12.0 %	Inlet temperature variation, lack of interpolation	Lowest	Exact	Static/Batch VM Scheduling, Backfilling
OPM	7.0 %	Inlet temperature variation, sharp rise/fall in CPU utilization	Medium	Exact	Individual VM scheduling, Dynamic scheduling, Offline monitoring
EOPM	3.5 %		Highest	+-Range	Same as WCPM, prediction error checking

8

5.0 EXPERIMENTAL SETUP AND RESULTS

The proposed prediction were models were applied to a variety of rack compatible servers which were heterogeneous computers ranging from rack compatible severs to desktops. The heterogeneity was on the basis of processor model, RAM and hard disk. As demonstrated in table (2), the prototype servers were categorized into four categories.

Table 2 Comparison matrix of all prediction models

Computer Type	Processor Model	Code Name
Server	Intel Xeon E5430 2.66GHz	SA
Server	Intel Xeon E5320 1.86 GHz	SB
Desktop	Core i7 3770 3.4 GHz (turbo boost disabled)	DA
Desktop		DB

The server DA had a legacy hard disk and server DB had the latest hard disk. The servers were running VMware ESXi 5.1 hypervisor. Each server had 12 VM each with single vCPU, 512 MB RAM and 10 GB hard disk. Each VM was running CPU intensive benchmark Prime95 [17]. The thermal benchmarking system performs the VM operations through vCenter API system calls.

For the creation of thermal-profile for any server, the volume of historic data required, depends upon the slots in thermal-profile. As demonstrated in Fig. (2), the 8 steps increment of server utilization yields a 9 slot thermal-profile with the first slot as the idle state. This will require the data of over 300 minutes. However, in this paper, 6 slot thermal-profiles were used as shown in Table (3). For that matter, the historic data required was of 180 minutes for each server.

5.1. Thermal-profiles

The thermal-profiles used for this paper consisted of 6 slots as demonstrated in table (3). The experiment results show that the increase in number of slots of thermal-profile increase the prediction accuracy for WCPM. However the OPM requires less number of slots but the improvement in accuracy is observed with more slots in thermal-profiles. For any server, the outlet temperature may not be in linear relationship with CPU usage and therefore interpolation may yield some error.

Table 2. Thermal-profiles used for implementation of prediction models.

Slot	CPU_{MHz}^{DA} (MHz)	T_{Δ}^{DA} (Celsius)	CPU_{MHz}^{DB} (MHz)	T_{Δ}^{DB} (Celsius)	CPU_{MHz}^{SA} (MHz)	T_{Δ}^{SA} (Celsius)	CPU_{MHz}^{SB} (MHz)	T_{Δ}^{SB} (Celsius)
0	200	4.68	200	3.86	200	10.54	200	13
1	3500	2.5	3500	2.3	5350	2.5	3750	2.3
2	6900	5	6900	4.13	7500	2.8	5600	3
3	10100	6.75	10100	5.34	10700	4.37	7600	4.25
4	13100	9	13100	6.5	16000	6.38	11200	5.5
5	15000	8	15000	7.9	21000	7.76	14800	6

5.2. Experimental results

This section presents the experimental results. As demonstrated in Fig. (3), different types of workload patterns were executed across all prototype servers. The five minute and ten minute aggregation detailed results are available at [18]. Thermal-profiles created earlier, were tested under enhanced type of thermal-benchmarking. As shown in Fig (4-15), DFS was periodically applied to each VM in order to create a diverse scenario in which each VM is undergoing ups and downs in utilization. The over all effect is the server (host) utilization is randomly being altered. With the periodic variations of each VM, the thermal-benchmarks demonstrated in Fig. (1) were executed to verify the accuracy of prediction models presented in this paper. The results presented here are related to ten minutes aggregated results and this paper recommends using the aggregated values for monitoring and scheduling of servers. First the results of desktop servers (DA and DB) are presented from Fig. (4-9). The EOPM is represented by EOPM(L) and EOPM (U) as demonstrated in Fig. (8, 9). The results consist

9

Malaysian Journal of Computer Science. Vol. 29(1), 2016

of aggregated values by ten minutes. The effect of conducting the workload of Fig. (4-5) over the servers DA and DB is the corresponding variation in outlet temperature.
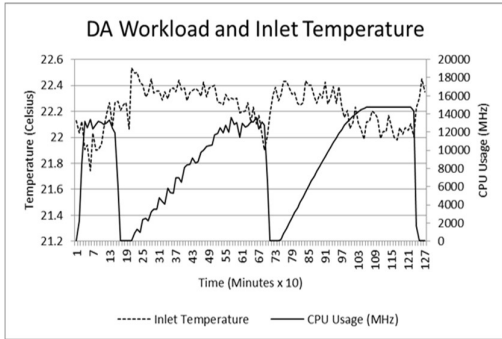

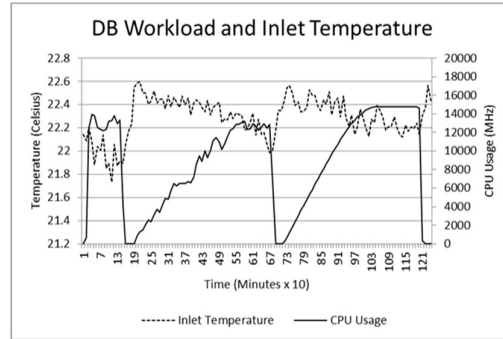Fig. 4 Server DA workload and inlet temperature

Fig. 5 Server DB workload and inlet temperature
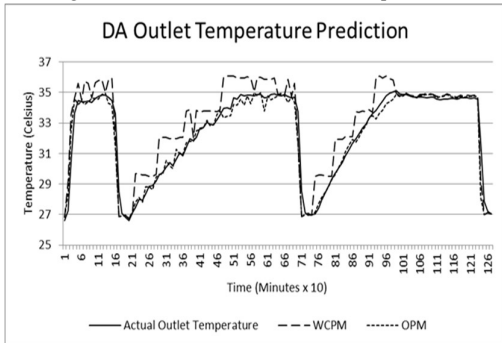

Fig. 6 Server DA outlet temperature prediction through WCPM and OPM.
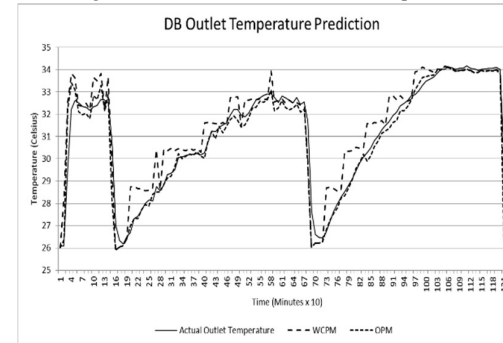

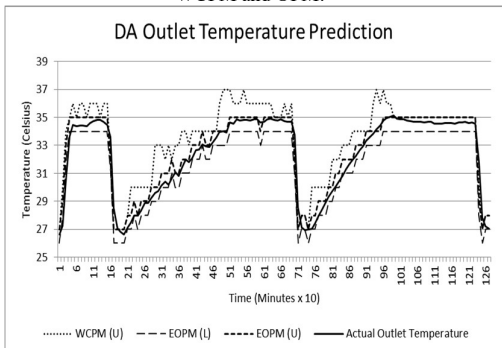Fig. 7 Server DB outlet temperature prediction through WCPM and OPM.

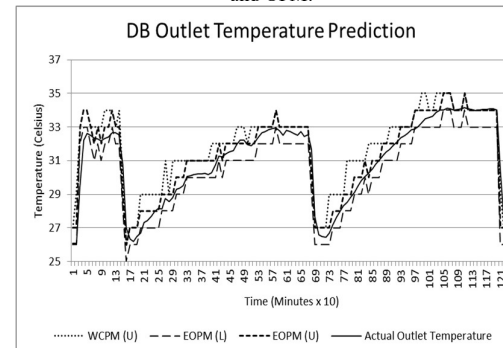Fig. 8 Server DA outlet temperature prediction through EOPM.

Fig. 9 Server DB outlet temperature prediction through EOPM.

It is observable that the outlet temperature calculated by OPM closely follows the actual outlet temperature. The OPM prediction closely follows the actual outlet temperature as demonstrated in Fig. (6, 7, 12, 13). The results show that the outlet temperature can be calculated offline with reasonable accuracy.
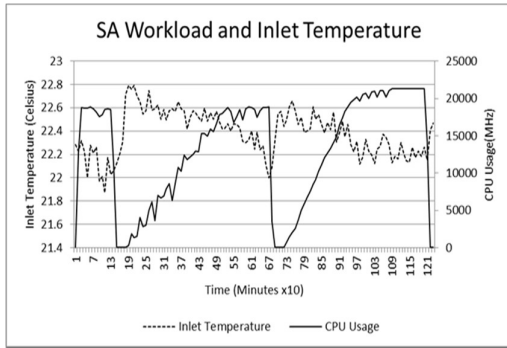
10
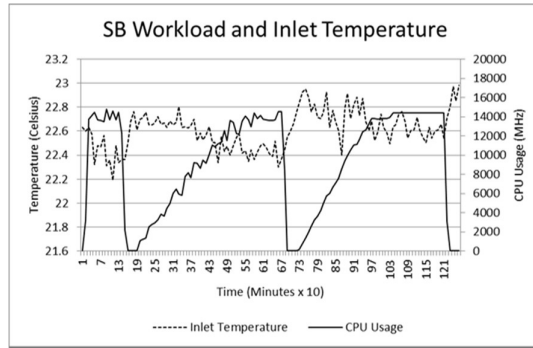
Fig. 10 Server SA workload and inlet temperature.

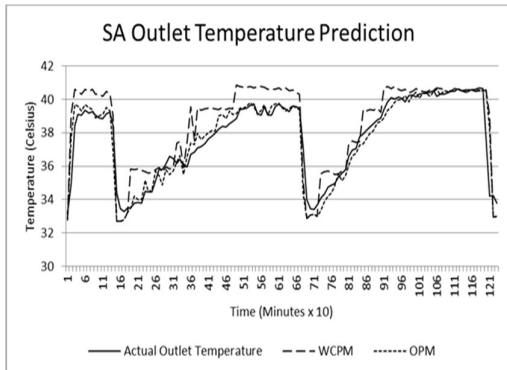
Fig. 21 Server SB workload and inlet temperature.

Fig. 32 Outlet temperature prediction for server SA through WCPM and OPM.
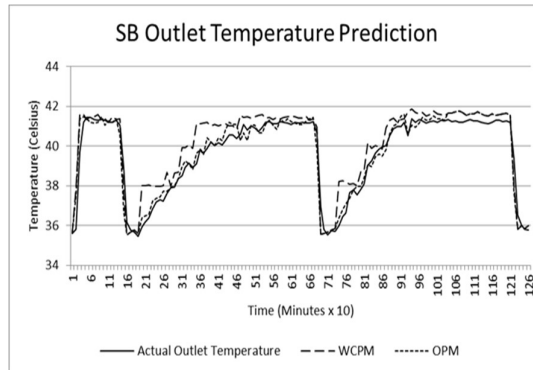

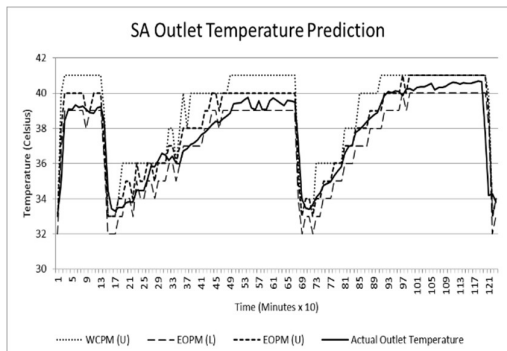Fig. 43 Outlet temperature prediction for server SB through WCPM and OPM.

Fig. 54 Server SA outlet temperature prediction through EOPM.
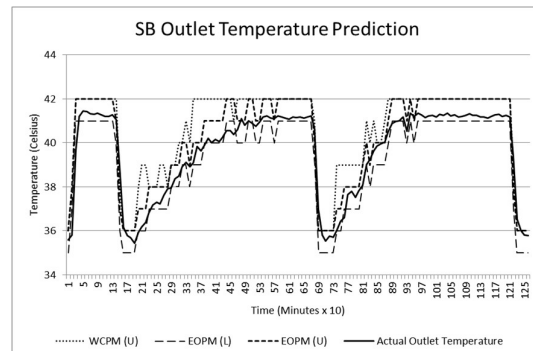

Fig. 15 Server SB outlet temperature prediction through EOPM.

The average prediction error in case of WCPM is always higher than the OPM. As demonstrated in Table (4) the prediction error for WCPM: β_{WCPM} and for OPM: β_{OPM}. The lesser the slots in thermal-profile and the number of VM to be powered-on is small, the higher will be the value of β_{WCPM}. Therefore the β_{WCPM} provides a macro-level of availed/not-availed potential of computing resources in terms of outlet temperature.

The experimental results for servers SA and SB are presented in Figs. (10-15). Overall the WCPM and OPM performed as explained in previous section. For all servers, the outlet temperature is always over-predicted by WCPM. These workload in Fig. (4, 5, 10 and 11) consists of 12 VM for each server. Each VM was powered-on after 30 minutes. The WCPM prediction clearly shows the unexplored margin for batch scheduling of VM. This is evident from the difference between actual outlet temperature and the predicted temperature in Fig. (6, 7, 12 and 13). As demonstrated in Fig. (8, 9, 14 and 15) the EOPM (U) and EOPM (L) provide a shortest range of β to contain the actual outlet temperature. For any server, EOPM certainly can demonstrate the micro-potential of outlet temperature that could be successfully exploited in any slot of thermal-profile. At every step, if the actual

11

outlet temperature of the server coincides with the EOPM (U), it means the respective slot of thermal-profile is fully utilized. Otherwise the closer the actual outlet temperature to EOPM (L), the more is the underutilization of the server or the underutilized micro-potential of the computing capacity of server in respective slot of thermal-profile.

As demonstrated in Table (4), the value of β_{WCPM} is always below 1 Celsius and for β_{OPM} is always below 0.5 Celsius. These are average values for each host. This shows that the OPM is accurate enough to be relied upon for monitoring and scheduling. The error results of detailed tests can further demonstrate this proposition. The value of β_{WCPM} is higher than β_{OPM} because the OPM model uses interpolation and WCPM predicts the worst case outlet temperature. Between the servers of same category (DA, DB and SA, SB), the difference of 0.3 Celsius in the values of β_{WCPM} is negligible and inherent to the thermal-profiles of the related servers. The aggregation and rounding of the values of outlet temperature and server utilization (during the creation of thermal-profiles and calculating the predictions) also contribute in prediction errors.

Table 3 Prediction error for WCPM and OPM

Server	β_{WCPM} (Celsius)	β_{OPM} (Celsius)
DA	0.9	0.31
DB	0.6	0.32
SA	0.9	0.44
SB	0.6	0.3

The output of prediction models will be used to implement a thermal-aware proactive VM scheduler in future. WCPM based predicted outlet temperature will be used as input to reactive thermal-aware scheduler and OPM based predicted outlet temperature will be used as input to a proactive thermal-aware workload scheduler. The results of the implementation of thermal-aware schedulers can later be compared. The EOPM based predicted range of outlet temperature will be used as input to a back-filling based thermal-aware workload scheduler. Thermal-aware migration of workload requires the outlet temperature prediction for the source and target servers for energy efficiency. In short, these models will be useful in establishing green data centers.

6.0 CONCLUSION AND FUTURE WORK
 This paper showed that the thermal-profiles of data center servers can be created and used for their outlet temperature prediction. This technique reduces the need for monitoring the outlet temperature, reduces the number of thermal sensors and subsequently minimizes the amount of outlet-temperature-data generated, transmitted and stored by the monitoring systems in data centers. Various prediction models: WCPM, OPM and EOPM were presented in this paper to predict the outlet temperature of the associated servers can also be utilized thermal-aware workload scheduling and simulation modeling. The OPM can predict the outlet temperature with the negligible error of 0.3 Celsius. This paper has shown that the thermal-profiles of servers can help in reducing the complexity of monitoring systems and the dependent systems such thermal-aware workload scheduling and thermal-aware load balancing in virtualized data centers. This will contribute towards the establishment of green data centers. In future, a proactive thermal-aware anomaly detection and workload scheduler will be proposed on the basis of thermal-prediction models presented in this paper. A more detailed thermal-benchmarking approach will be presented as an authentic and established methodology for thermal evaluation of the data center servers.

12

Malaysian Journal of Computer Science. Vol. 29(1), 2016

REFERENCES

[1]   M. Rouse. (August 2013). *data center infrastructure management (DCIM)*. Available:
        http://searchdatacenter.techtarget.com/definition/data-center-infrastructure-management-DCIM

[2]   N. Vasic, T. Scherer, and W. Schott, "Thermal-aware workload scheduling for energy efficient data
        centers," presented at the Proceedings of the 7th international conference on Autonomic computing,
        Washington, DC, USA, 2010.

[3]   M. T. Chaudhry, T. C. Ling, A. Manzoor, S. A. Hussain, and J. Kim, "Thermal-Aware Scheduling in
        Green Data Centers," *ACM Comput. Surv.,* vol. 47, pp. 1-48, 2015.

[4]   H. Dong and S. Tao, "Thermal-aware energy-efficient task scheduling for DVFS-enabled data centers,"
        in *Computing, Networking and Communications (ICNC), 2015 International Conference on*, 2015, pp.
        536-540.

[5]   Q. Tang, S. Gupta, and G. Varsamopoulos, "Thermal-Aware Task Scheduling for Data centers through
        Minimizing Heat Recirculation," in *Cluster Computing, 2007 IEEE International Conference on*,
        Austin, TX, U.S., 2007, pp. 129-138.

[6]   M. T. Chaudhry, T. C. Ling, Syed Asad HUSSAIN, ATIF MANZOOR, "Minimizing Thermal-stress
        for Data Center Servers Through Thermal-aware Relocation," *The Scientific World Journal,* 2014.

[7]   M. T. Chaudhry, T. Ling, S. A. Hussain, and X. Lu, "Thermal-aware relocation of servers in green data
        centers," *Frontiers of Information Technology & Electronic Engineering,* vol. 16, pp. 119-134, 2015.

[8]   L. Wang, S. Khan, and J. Dayal, "Thermal aware workload placement with task-temperature profiles in
        a data center," *The Journal of Supercomputing,* vol. 61, pp. 780-803, 2012/09/01 2012.

[9]   I. Rodero, H. Viswanathan, E. Lee, M. Gamell, D. Pompili, and M. Parashar, "Energy-Efficient
        Thermal-Aware Autonomic Management of Virtualized HPC Cloud Infrastructure," *Journal of Grid
        Computing,* vol. 10, pp. 447-473, 2012/09/01 2012.

[10]   A. Krause, C. Guestrin, A. Gupta, and J. Kleinberg, "Near-optimal sensor placements: maximizing
        information while minimizing communication cost," in *Information Processing in Sensor Networks,
        2006. IPSN 2006. The Fifth International Conference on*, 2006, pp. 2-10.

[11]   M. Shicong, A. K. Iyengar, I. M. Rouvellou, L. Ling, L. Kisung, B. Palanisamy*, et al.*, "Reliable State
        Monitoring in Cloud Datacenters," in *Cloud Computing (CLOUD), 2012 IEEE 5th International
        Conference on*, 2012, pp. 951-958.

[12]   M. Shicong, L. Ling, and W. Ting, "State Monitoring in Cloud Datacenters," *Knowledge and Data
        Engineering, IEEE Transactions on,* vol. 23, pp. 1328-1344, 2011.

[13]   Q. Junmei, L. Li, L. Liang, T. Yuelong, and C. Jiming, "Smart temperature monitoring for data center
        energy efficiency," in *Service Operations and Logistics, and Informatics (SOLI), 2013 IEEE
        International Conference on*, 2013, pp. 360-365.

[14]   W. Xiaodong, W. Xiaorui, X. Guoliang, C. Jinzhu, L. Cheng-Xian, and C. Yixin, "Intelligent Sensor
        Placement for Hot Server Detection in Data Centers," *Parallel and Distributed Systems, IEEE
        Transactions on,* vol. 24, pp. 1577-1588, 2013.

[15]   H. Liu, E. Lee, D. Pompili, and X. Kong, "Thermal camera networks for large datacenters using real-
        time thermal monitoring mechanism," *The Journal of Supercomputing,* vol. 64, pp. 383-408,
        2013/05/01 2013.

[16]   H. Viswanathan, E. K. Lee, and D. Pompili, "Self-Organizing Sensing Infrastructure for Autonomic
        Management of Green Datacenters," *Network, IEEE,* vol. 25, pp. 34-40, 2011.

[17]   C. Wei, G. Fengqian, and L. Yuan, "Policy Based Power Management in Cloud Environment with Intel Intelligent Power Node Manager," in *Enterprise Distributed Object Computing Conference Workshops (EDOCW), 2012 IEEE 16th International*, 2012, pp. 66-69.

[18]   M. T. Chaudhry. Experiment Results Network Efficient Monitoring: (pswd4allFiles: "thermal") https://www.mediafire.com/folder/a51zydugyxzq9/ [Online].