Dynamic Power Simulator Utilizing Computational Fluid Dynamics and Machine Learning for Proposing Task Allocation in a Data Center

Kazumasa Kitada*, Yutaka Nakamura[†], Kazuhiro Matsuda[‡] and Morito Matsuoka^{‡ §}

*School of Information Science and Technology, Osaka University, Osaka, Japan

Email: k-kitada@ane.cmc.osaka-u.ac.jp

[†]Graduate School of Engineering Science, Osaka University, Osaka, Japan

Email: nakamura@is.sys.osaka-u.ac.jp

[‡]Cybermedia Center, Osaka University, Osaka, Japan

Email: kazuhiro.matsuda@ane.cmc.osaka-u.ac.jp

[§]Advanced Telecommunications Research Institute International, Kyoto, Japan

Email: matsuoka@cmc.osaka-u.ac.jp

Abstract—A dynamic power simulator for a data center was demonstrated by combining computational fluid dynamics (CFD) and machine learning. The total power consumption of the data center was simulated. The sensitivity of the temperature distribution along the virtual machine (VM) allocation was analyzed using a non-parametric process for the CFD. An allocation of server tasks was proposed for reducing the power consumption of the air conditioner installed in the data center. This simulation showed that the optimum operating temperature increases with the power usage effectiveness. These results indicate that the power simulator developed in this study is a powerful tool for dynamic power simulation and for estimation of better operation parameters, including VM allocation, from the aspect of power consumption.

Keywords–Data center; power simulator; computational fluid dynamics; virtual machine allocation; machine learning.

I. INTRODUCTION

In recent years, many data centers have been built to support the increased popularity of services such as video on demand and cloud storage. In line with this trend, the power consumption of data centers has increased sharply [1]. As a result, the power consumption of data centers has become a serious social problem, and studies on energy-efficient data centers have attracted considerable attention [2]-[6]. Various proposals have been put forward for reducing the power consumption of data center equipment such as air conditioners, power supply units, and servers [7]–[9]. Considering the total power consumption of a server, Khuller et al. proposed an energy management system in which the workload of all servers was concentrated in a portion of the servers and the other servers were turned off [10]. Iyengar et al. improved the energy efficiency of data centers by configuring the fan speed of air conditioners or shutting off air conditioners according to the temperature distribution in the data center [11]. However, reducing the power consumption of individual units without coordination among them is insufficient for reducing the total power consumption of the data center, in which various types of equipment operate in a strongly interrelated way.

ICT equipment and air conditioners account for most of the power consumption in a data center [12]. Therefore, cooperative control, such as simultaneous optimization of the server workload assignment and air conditioners, is essential for improving the energy efficiency of the data center. The power consumption of each piece of equipment depends not only on its own operational conditions but also on those of others. For example, the server power consumption depends on the inlet air temperature, which is in turn affected by the air-conditioner setting. The server workload assignment and operation set points of the air conditioners must be controlled in a coordinated manner to minimize the total energy consumption of the data center. To reduce the total power consumption, it is necessary to predict the state of the data center, such as its temperature distribution, given by the operational conditions such as virtual machine allocation. Because the heat emitted from the servers accounts for most of the heat emitted from the data center, it is essential to predict the temperature of the exhaust heat from the servers for predicting the total power consumption of the data center. However, predicting the heat emitted from the servers is difficult because many factors are interdependent. For example, the temperature of the exhaust heat from the server depends on the air volume passing through the server and the server's power consumption. Moreover, as the intake temperature and air volume vary according to the position of the server in the data center, the temperature of the exhaust from the server changes with the location even when the same workload is assigned to servers with the same power consumption.

Computational fluid dynamics (CFD) is a representative technique for simulating the temperature distribution when various parameters are interrelated [13]–[15]. However, CFD estimates only the temperature distribution and normally does not simulate the power consumption directly.

In this study, we developed a novel simulator that estimates the power consumption of the data center. This simulator was demonstrated by combining the temperature distribution simulation using CFD and machine learning techniques to predict the power consumption of the server and air conditioner. Power consumption of the servers and the air conditioner was estimated by regression models, such as a neural network. Their parameters were trained so that the estimates fit the exact power consumption of the data center in operation. The air temperature and fan rotational speed of the air conditioner were measured as learning data to build the power consumption model of the air conditioners. The procedure for estimating the total power consumption of the data center is as follows. First, the exhaust temperature for the servers is simulated using a CFD simulator with a task. Then, the total power consumption of the data center is calculated as a summation of the individual equipment power consumptions obtained from the CFD simulation results and the power consumption model

of each equipment.

By combining the proposed simulator and a sensitivity analysis of CFD, VM allocation optimization from the aspect of power consumption was demonstrated to equalize the exhaust temperature in the rack plane. Through VM allocation, we increased the blowing temperature of the air conditioner and reduced its power consumption. In addition, we evaluated the optimum temperature set point of the air conditioner to minimize the total power consumption of the data center with several power usage effectiveness (PUE) [12] values.

The remainder of this paper is organized as follows. Section II describes the power consumption simulator. Section III describes the procedure for VM allocation. Section IV discusses the evaluation of the proposed power consumption simulator. Finally, Section V presents our conclusions and possible areas of future work.

II. ENERGY SIMULATOR

Our research group manages an experimental data center as a test bed to demonstrate energy-saving technologies for a green data center. Figure 1(a) shows a general view of the testbed data center. In this data center, a system that reuses exhaust heat from the servers is implemented to reduce power consumption efficiency [16]. Figure 1(b) shows the equipment arrangement and airflow of a conventional data center with one hot aisle (HA) and one cold aisle (CA). On the other hand, the testbed data center has a conventional arrangement with cold and hot aisles and an additional super-hot aisle (SHA) to raise the exhaust heat temperature to around 50° C. The heat reuse efficiency reaches practical levels at this temperature.

A. Summary of power simulator

Figure 2(a) shows a conceptual diagram of the power simulator for the data center. In our previous study, the temperature prediction and power estimation were demonstrated using only the machine-learning-based power simulator [17] [18]. In that system, sensor data are required for temperature prediction as well as estimation of the power consumption of the data center.

In this study, the power consumption simulator consists of the CFD simulator and external programs to estimate the power consumption of individual pieces of equipment in the data center. Figure 2(b) shows a conceptual diagram of the power simulator, which comprises CFD and a power predictor, for the data center. It is not necessary to install sensors across the entire data center because the sensor data are not required for estimating the power consumption. As a result, this system is more practical than a power simulator based on machine learning alone. The server's power consumption is estimated by a neural network whose inputs are the CPU usage, intake temperature, and rotational speed of the fan of the air conditioner. The exhaust temperature of each server is estimated by the CFD simulator according to the predicted heating value. The power consumption of each server and that of the air conditioner are estimated using the calculation result of the CFD simulation. By using the CFD simulation result, the power consumption is estimated by the power model for the server and CFD simulation. The CFD simulation and power consumption estimation are performed in turn. The total power consumption of the data center is estimated by the power consumption of the servers and air conditioners. The power consumption of the air conditioner is also estimated by



(a) Tandem arrangement with cold aisle, hot aisle, and super-hot aisle to reuse exhaust heat from servers



(b) Conventional arrangement with cold aisle and hot aisle



using the learning data on the total power consumption of the servers, temperature from the air conditioner, air conditioner fan rotational speed, and fresh air temperature and humidity. After estimating the power consumption of all servers, the power consumption of the air conditioner is estimated using the air conditioner model. As a result, the power consumption of the entire data center and the temperature distribution in the data center were determined.

B. CFD simulation

CFD is an analytical technique for numerically solving a hydrodynamic governing equation [19]. For the CFD simulation, we follow the steps described below. First, the simulation



(b) CFD-based power simulator (this study)

Figure 2. Block diagrams of the power consumption simulator for a data center constructed using only machine learning and using the hybrid CFD and machine learning method

model is built, in which servers acting as heat sources and an air conditioner are emulated. Next, parameters including the heat value of the server, intake air temperature, and air volume are set. Then, the required estimation is done using the CFD simulator (Flow Designer Ver. 12; Advanced Knowledge Lab., Inc.). As a result of the simulation, the temperature at any point in the entire analysis domain can be evaluated. Two models are used for CFD analysis: steady-state analysis for the state in which the temperature is stable and unsteady analysis for the transition state before converging to a steady state. In this study, a simulation based on steady-state analysis is demonstrated.

C. Server power consumption model

This section describes the construction of the server power consumption model. The CPUs, Memory, hard disc drive (HDD) and internal fans account for most of the power consumption of a server. The power consumption of an HDD is almost constant, and additional power consumption occurs when the HDD is accessed. Memory also consumes a certain amount of energy only when it is accessed. As a result of our experiment to analyze the server power consumption, we found that the power consumption was around 4 W at most when 1 GB of memory is accessed. The change in the power consumption of the HDD is almost negligible even when the drive is accessed. Therefore, we assumed that the change in the server power consumption mainly depends on the CPU and internal fan in this case, because the amount of memory installed in each server was 1 GB at most.

The power consumption of the CPU changes with the usage ratio and its temperature. The CPU temperature depends on the CPU usage ratio, intake temperature of the server, and air speed passing through the server. The air speed passing through the server mainly depends on the rotational speed of the internal

TABLE I. SERVER INFORMAITON FOR POWER MODEL



Figure 3. A typical server power model constructed by the neural network method

fans. The rotational speed of the internal fan is controlled by using the CPU temperature and intake temperature of the server. In this way, the power consumption of the server is determined by several parameters that depend on each other. As a result, a simple linear regression model is not suitable for constructing the power model in the data center.

In this study, we use a neural network to construct the server power consumption model. The parameters of the neural network were determined by using the operation data of a server in the testbed data center. Table I shows the specifications of the server 's measured operation data. The CPU usage ratio and intake temperature were adopted as the explanatory variables to estimate the server power consumption. We also adopt the blowing temperature and fan rotational speed of the air conditioner as explanatory variables. The network shape and hyperparameters of learning are optimized. As a result, the server power consumption model is constructed. The root mean square error (RMSE) between the measured and the estimated values of the test data set is 8.41 kW. Figure 3 shows a server power model constructed by the neural network method. This figure shows the typical relationships among the measured power consumption of the server, its CPU usage ratio, and its intake temperature. The estimated error of the server power consumption reaches 6.7% for the test data.

D. Air-conditioner power model

Coefficient of performance (COP) is an index that indicates the power consumption efficiency of an air conditioner [20], and it is defined as follows:

$$COP = \frac{c}{P_{consume}} \tag{1}$$

where c and $P_{consume}$ are the coolability and total power consumption of the data center, respectively. Figure 4 shows the COP against the operational parameters of outlet air temperature and fan speed of the air conditioner as estimated by a support vector regression model. This support vector



Figure 4. Air-conditioner power model constructed by support vector regression method

regression model is trained by using the recorded data from our experimental data center [21] [22]. From this figure, it is clear that the COP dynamically changes against the operational parameters of the air conditioner. Therefore, this suggests that controlling the air conditioner using the prediction method enables the power consumption to be reduced.

III. TASK ALLOCATION

A. Procedure for Task Allocation

This section describes the procedure for determining the quantity of tasks to assign to each server when some workload is assigned to the entire data center. A VM is assigned and moved to each server as a task unit. We do not distinguish between different types of VMs, and we assume that there is no difference in the load between VMs. Turning off servers with no assigned task to reduce power consumption is not done in this allocation method.

VM allocation follows the steps described below.

- 1) Calculate the temperature distribution subject to a given initial allocation of workloads by order analysis of CFD simulation.
- 2) Conduct sensitivity analysis against the given objective with respect to the generated heat at each server. In this study, the objective is to achieve a uniform exhaust temperature distribution, and thus, the objective function is defined using the variance of the exhaust temperatures.
- 3) Find the "local optimal" heat pattern to minimize the variance using the sensitivity analysis result.
- 4) Find the VM allocation pattern that produces a generated heat pattern that is closest to the local optimum.
- 5) Reallocate VMs to realize the allocation pattern.

The VM allocation procedure for the data center model with a conventional arrangement is described in the following subsections.



Figure 5. Temperature distribution in the exhaust-side plane of the rack when VMs are assigned to servers randomly

| | | Sen | sitivity |
|------|------|------|----------|
| | | | 0.02 |
| | | | |
| | | | 0.00 |
| | | | 0.02 |
| | | | 0.02 |
| | | | |
| | | | -0.01 |
| | | | 0.01 |
| | | | |
| | _ | | 0.01 |
| | | | |
| | | | 0.01 |
| | | | |
| | | | |
| | | | 0.00 |
| | | | |
| | | | |
| | | | |
| | | | 0.00 |
| | | | |
| | | | |
| | | | -0.00 |
| | | | 0.00 |
| | | | |
| | | | |
| | | | -0.01 |
| | | | |
| | | | 0.01 |
| | | | -0.0 |
| | | | 0101 |
| | | | |
| | | | |
| | | | -0.0Z |
| | | | |
| | | | 0.00 |
| | | | -0.02 |
| | | | 0.02 |
| | | | |

Figure 6. Sensitivity map of each server for temperature distribution in the case of the randomly assigned VM shown in Fig. 5

B. Order analysis of the initial pattern of generated heat

As preparation for sensitivity analysis, order analysis for the initial workload is demonstrated. The initial pattern of generated heat is obtained for the VM allocation randomly assigned to the servers. Figure 5 shows the temperature distribution in the exhaust-side plane of the rack. CFD simulation is applied for the initial pattern of generated heat.

C. Sensitivity analysis for exhaust temperature in rack plane

A sensitivity analysis is a non-parametric process to estimate the intensity of the influence (sensitivity) of a small change in each parameter [23]. In this process, the variance with the temperature of each server in the exhaust-side plane is used as an objective function. The sensitivity analysis for the objective function is demonstrated, and then, the sensitivities for the variance in the rack exhaust temperature of each server are obtained. Figure 6 shows the value of the sensitivity of each server in a color bar. This sensitivity shows exhaust temperature variance in the rack plane when the heating value increases to 1 W.

D. Calculate target pattern of generated heat

By using the sensitivity, we find the servers' allocation pattern of generated heat that equalizes the exhaust temperature in the rack plane. The local optimal generated heat pattern that does not require VM reallocation is the target pattern. The target amount of generated heat $target_Power_i$ of server *i*

TABLE II. EXHAUST TEMPERATURE VARIANCE WHEN USING FIVE COEFFICIENTS FOR PATTERN OF GENERATED HEAT

| Coefficient | Variance |
|-------------|----------|
| 0 (initial) | 0.175 |
| 15000 | 0.479 |
| 30000 | 0.256 |
| 45000 | 0.107 |
| 60000 | 0.103 |

is calculated, and it changes in proportion to the sensitivity of the current amount of generated heat $Power_i$, as shown in (2). After finding the local optimum, we do not recalculate to find a better allocation pattern because the exhaust temperature distribution hardly improves from the viewpoint of the power consumption.

$$target_Power_i = Power_i + Sensitivity_i \times \alpha \qquad (2)$$

Here, coefficient α in (2) is a proportionality constant. The coefficient α is set to an appropriate value, and a pattern of generated heat that reduces the rack exhaust temperature variance is obtained.

The difference between the intake and the exhaust air temperature is $\Delta Temp_i$, and it depends on the heating value of a server $Power_i$, speed of air passing over a server $AirSpeed_i$, and the heat transfer coefficient between the air and the server $\eta_{heat_transfar}$, as shown in (3) [24].

$$\Delta Temp_i = \frac{Power_i}{Air_Speed_i} \times \eta_{heat_transfar}$$
(3)

Assuming that the speed of air passing over a server is constant, $\Delta Temp_i$ is proportional to the server heating value. The exhaust temperature variance in the rack plane is almost proportional to the square of the server heating value. As with the temperature, the objective function of sensitivity analysis is proportional to the square of the coefficient α . The order analysis for the patterns of generated heat calculated using two values of α is shown, and the exhaust temperature variance in the rack plane is calculated. Table II shows the exhaust temperature variance in the rack plane as estimated by order analyses using at least three patterns of generated heat, including the initial pattern. A quadratic function of the exhaust temperature variances in the rack plane and a coefficient using at least three pairs of variances are calculated. As a result, a coefficient is estimated so that the quadratic function approaches a local minimum. Figure 7 shows the quadratic curve for the five points shown in Table II and the local minimum of the quadratic curve. Here, an appropriate value for α is 57676 in the initial pattern of generated heat. Then, the target heating value of each server using this value of α and equation (2) is obtained.

E. VM reallocation

VM reallocation is demonstrated to make each server's amount of generated heat close to the target pattern. The difference between them is shown in equation (4).

$$diff_i = Power_i - target_Power_i \tag{4}$$



Figure 7. Quadratic function of rack exhaust temperature variances and coefficient

$$\sum |diff_i| \tag{5}$$

Here, $server_j$ means that the decrease in $diff_i$ becomes a maximum when one VM is removed from the server. $server_k$ means that the decrease in $diff_i$ becomes a maximum when one VM from the server is moved using the server power model. Then, a VM is moved from $server_j$ to $server_k$. This procedure is repeated until (5) does not decrease. As a result, this procedure greedily assigns server VMs.

IV. EVALUATION

A. Power estimation using proposed simulator

We estimated the power consumption of the data center by using the developed power simulation with the measured CPU usage ratio of each server and the air-conditioner setting in the testbed data center. We compared the result of the power simulation with the measured power consumption of the testbed data center. Our testbed data center includes many types of servers, including the servers for which a power model was constructed. Table III shows information about these servers. The power model of each server was constructed in the same manner as in the previous section. Table IV shows the power consumption of each piece of equipment, total power consumption estimated for the data center, and power consumption measured in the testbed data center. The power simulator estimated the power consumption of the data center with an average error rate of 5.34%.

B. VM allocation in data center

We describe the VM allocation procedure for the data center model with a conventional arrangement (Fig. 1(b)). This data center model comprises six racks with 40 servers per rack and air conditioners. RX300S4 servers in tableI are constructed in this data center model. The maximum number of VMs maxVM that each server runs is eight. The CPU usage cpu_usage_i of serveri was calculated by using the assigned number of VMs VM_i and (6).

$$cpuusage_i = \frac{inVM_i}{maxVM} \times 100 \tag{6}$$

TABLE III. SERVER INFORMATION FOR THE TESTBED DATA CENTER

| Server | CPU | Memory | HDD storage |
|----------|--------------------------------|--------|-------------|
| Server A | Intel(R) Xeon(R) CPU 2.40 GHz | 1 GB | 200 GB |
| Server B | Intel(R) Xeon(TM) CPU 3.20 GHz | 1 GB | 200 GB |
| Server C | Intel(R) Xeon(TM) CPU 3.80 GHz | 1 GB | 500 GB |
| Server D | Intel(R) Xeon(R) CPU 1.60 GHz | 1 GB | 250 GB |
| Server E | Intel(R) Xeon(R) CPU 2.00 GHz | 1 GB | 400 GB |
| Server F | Intel(R) Xeon(TM) CPU 3.80 GHz | 1 GB | 400 GB |
| Server G | Intel(R) Xeon(R) CPU 3.00 GHz | 1 GB | 1 TB |
| Server H | Intel(R) Xeon(TM) CPU 3.80 GHz | 1 GB | 500 GB |
| | | | |

TABLE IV. POWER CONSUMPTION OF EACH DEVICE ESTIMATED BY SIMULATION AND MEASUREMENT IN TESTBED DATA CENTER

| | Server power | Air-conditioner power | Data center total power |
|---------------------|--------------|-----------------------|-------------------------|
| Testbed data center | 57.15 kW | 5.85 kW | 65.70 kW |
| Simulation result | 56.63 kW | 5.74 kW | 62.37 kW |



Figure 8. Air temperature distribution of the data center with random pattern of generated heat (RMSE: 0.82)

The average utilization rate of the data center is 20%–30%; therefore, 650 VMs are assigned to 240 servers for the initial pattern of generated heat so that the mean utilization rate of the servers is 33%. Figure 8 shows the temperature distribution of the rack exhaust side as obtained from a CFD simulation using this pattern of generated heat. The rack exhaust temperature variance for this temperature distribution was 0.76. The proposed VM allocation procedure was implemented for this initial pattern of generated heat. 140 VMs were migrated, and a new pattern was obtained. Then, a CFD simulation was conducted using the newly obtained pattern of generated heat. The new temperature distribution of the rack exhaust side was obtained, and then the rack exhaust temperature variance decreased to 0.23. Figure 9 shows this temperature distribution. Table V shows the temperature distribution of the rack exhaust side from the analysis results before and after reallocation. The maximum rack exhaust temperature decreased from 38.4°C to 36.2°C by using the proposed VM allocation procedure. This indicates that the air-conditioner temperature can increase by 2.2°C relatively in comparison with that before the VM reallocation, and then the power consumption of the air conditioner might decrease by around 5%.

C. Minimizing total power consumption of data center

1) Suitable operation point of air conditioner: The power consumption of a server depends on its CPU usage, its intake



Figure 9. Air temperature distribution of the data center after VM reallocation (RMSE: 0.23)

TABLE V. RACK EXHAUST TEMPERATURE VARIANCE AND MAXIMUM RACK EXHAUST TEMPERATURE BEFORE AND AFTER VM REALLOCATION

| | Initial (random) | After reallocation |
|----------------------------------|------------------|--------------------|
| Exhaust temperature variance | 0.82 | 0.23 |
| Maximum exhaust heat temperature | 38.4°C | 36.2°C |

temperature, and other factors, and it increases when the intake temperature increases. The power consumption of the air conditioner depends on its blowing temperature and fan rotational speed, and it increases when the blowing temperature is reduced. Therefore, increasing the air conditioner's blowing temperature reduces its power consumption. On the other hand, as the intake temperature of the servers increases, the total power consumption of the servers increases because the leakage current at the processors or rotational speed of the internal fans increase. Therefore, it is expected that operational temperature has the most suitable operation point from the aspect of power consumption of the data center which is the sum of the power consumption of all servers and air conditioners. A reduction in the power consumption of the entire data center is expected by operating the air conditioner at the abovementioned blowing temperature. Figure 10 shows the total power consumption of the servers, power consumption of the air conditioner, and total power consumption of the data center as described in when the air temperature is varied from 15°C to 30°C and the air conditioner is operated. Figure 11



Figure 10. Power consumption of servers, power consumption of air conditioner, and total power consumption of data center versus air-conditioner blowing temperature



Figure 11. Total power consumption of data center versus air temperature, showing the temperature that minimizes the power consumption

shows an enlarged view of the total power consumption of the data center shown in Figure 10. The total power consumption of the data center is given by a concave function, and it reaches a minimum when the air blowing temperature of the air conditioner is set to 20.5° C and operated. In this manner, the air-conditioner blowing temperature that minimizes the power consumption of the data center was calculated by the power consumption simulator. A reduction in the power consumption of the data center is expected by using the abovementioned value of the air blowing temperature.

2) Change in operation point of air conditioner for PUE: We evaluate the change in the air conditioner's blowing temperature that minimizes the total power consumption of the data center for different PUE values. The PUE is an index that indicates the power consumption efficiency of the data center, and it is calculated by using the power consumption of all servers $Power_{server}$ and the total power consumption of the data center $Power_{DC}$, as given in (7).

$$PUE = \frac{Power_{DC}}{Power_{server}} \tag{7}$$

According to the power simulation result, when the air temperature was set to 20.5°C as calculated in IV-C1, the PUE



Figure 12. Power consumption of the air conditioner for PUE values of 1.2, 1.3, 1.5, 1.7, and 2.0



Figure 13. Total power of the data center and the air temperature that minimizes the power consumption of the data center for PUE values of 1.2, 1.3, 1.5, 1.7, and 2.0

was 1.18. We estimate and compare the power consumption of the data center for PUE values of 1.3, 1.5, 1.7, or 2.0 when the air conditioner blowing temperature was set to 20.5°C. Figure 12 shows the power consumption of the air conditioner for each PUE value when the air conditioner blowing temperature was varied.

The air-conditioning power that gives each PUE was calculated on the assumption that the power to give the PUE (1.18 at 20.5°) obtained in an actual data center varies in proportion to each temperature. It is assumed that the server power did not change in this case. In addition, Figure 13 shows the change in the total power consumption of the data center at each PUE versus the air-conditioner blowing temperature using the air-conditioner power shown in Fig. 12. The red points show the operating temperature point at which the data center's total power is minimized for each PUE. In this manner, an air conditioner's operation point that minimizes the total power consumption of the data center changes with the PUE. Moreover, it was found that the air temperature that minimizes the power decreases so that the PUE value is small. In other words, it is advantageous from the viewpoint of the power consumption to set a lower air temperature when the power efficiency is high. Thus, the developed power simulator is a powerful tool for VM allocation to reduce an air conditioner's power consumption. Furthermore, it can be used to determine the most suitable operation point from the viewpoint of the power consumption of the data center through the dynamic simulation of the power consumption.

V. CONCLUSION AND FUTURE WORK

A novel power simulator was developed for the dynamic and real-time estimation of the power consumption of a data center. This simulator estimates the data center's power consumption through a combination of CFD and machine learning. The prediction error was suppressed to around 5.3% for the power consumption. A VM assignment policy was proposed to reduce the power consumption of the air conditioners by narrowing the distribution of exhaust heat from the servers. In addition, the optimum operation temperature was proposed to reduce the power consumption by using the power simulator.

These results indicate that the power simulator developed in this study shows promise as a real-time and dynamic power simulation tool.

In this study, learning data were obtained from a real server located in a data center. However, several interactions occur between servers in a real environment. In future work, it is necessary for the power simulator to take the relationships among the servers.

ACKNOWLEDGMENT

This work was supported by a CO_2 emission reduction project of the Japanese Ministry of the Environment and NICT. A part of this work was collaborative research with NTT Network Innovation Laboratories. We thank Advanced Knowledge Laboratory, Inc., for support with the CFD techniques. We thank Takasago Thermal Engineering Inc. for support with the air conditioner and aisle arrangement. We also thank Fujitsu Ltd. for supporting server management and sharing information about the characteristics of the servers.

REFERENCES

- J. Koomey, "Growth in data center electricity use 2005 to 2010," A report by Analytical Press, completed at the request of The New York Times, Aug. 2011.
- [2] A. Banerjee, T. Mukherjee, G. Varsamopoulos, and S. Gupta, "Coolingaware and thermal-aware workload placement for green hpc data centers," in Proceedings of International Green Computing Conference, Aug. 2010, pp. 245–256.
- [3] T. D. Boucher, D. M. Auslander, C. E. Bash, C. C. Federspiel, and C. D. Patel, "Viability of dynamic cooling control in a data center environment," Journal of electronic packaging, vol. 128, no. 2, Jun. 2006, pp. 137–144.
- [4] A. Beloglazov, J. Abawajy, and R. Buyya, "Energy-aware resource allocation heuristics for efficient management of data centers for cloud computing," Future generation computer systems, vol. 28, no. 5, May 2012, pp. 755–768.
- [5] B. Battles, C. Belleville, S. Grabau, and J. Maurier, "Reducing data center power consumption through efficient storage," NetApp Technical Report, Feb. 2007.
- [6] L. Parolini, B. Sinopoli, and B. H. Krogh, "Reducing data center energy consumption via coordinated cooling and load management," in Proceedings of the Conference on Power Aware Computing and Systems, Dec. 2008.
- [7] M. Norota, H. Hayama, M. Enai, T. Mori, and M. Kishita, "Research on efficiency of air conditioning system for data-center," in Proceedings of International Telecommunications Energy Conference 2003, Oct. 2003, pp. 147–151.

- [8] C. Ge, Z. Sun, and N. Wang, "A survey of power-saving techniques on data centers and content delivery networks," IEEE Communications Surveys Tutorials, vol. 15, no. 3, Jul. 2013, pp. 1334–1354.
- [9] D. Cavdar and F. Alagoz, "A survey of research on greening data centers," in Proceedings of Global Communications Conference Exhibition and Industry Forum 2012, Dec. 2012, pp. 3237–3242.
- [10] S. Khuller, J. Li, and B. Saha, "Energy efficient scheduling via partial shutdown," in Proceedings of Society for Industrial and Applied Mathematics 2010, Jan. 2010, pp. 1360–1372.
- [11] M. Iyengar, R. Schmidt, and J. Caricari, "Reducing energy usage in data centers through control of room air conditioning units," in Proceedings of 12th IEEE Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems, Jun. 2010, pp. 1–11.
- [12] The Green Grid, "Guidelines for energy-efficient data centers," *available at http://www.thegreengrid.org/*.
- [13] C. D. Patel, C. E. Bash, C. Belady, L. Stahl, and D. Sullivan, "Computational fluid dynamics modeling of high compute density data centers to assure system inlet air specifications," in Proceedings of The Pacific Rim/ASME International Electronic Packaging Technical Conference and Exhibition, Jul. 2001, pp. 1–9.
- [14] C. Patel, R. Sharma, C. Bash, and A. Beitelmal, "Thermal considerations in cooling large scale high compute density data centers," in Proceedings of Proceedings of IEEE Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems, June 2002, pp. 767–776.
- [15] U. Sing, A. Singh, and A. Sivasubramaiam, "CFD-based operational thermal efficiency improvement of a production data center," in Proceedings of the First USENIX Conference on Sustainable Information Technology, ser. SustainIT'10, Feb. 2010, pp. 6–6.
- [16] Y. Taniguchi and et al., "Tandem equipment arranged architecture with exhaust heat reuse system for software-defined data center infrastructure," IEEE Transactions on Cloud Computing, vol. PP, no. 99, June 2015, pp. 1–13.
- [17] Y. Tarutani and et al., "Temperature distribution prediction in data centers for decreasing power consumption by machine learning," in Proceedings of IEEE 7th International Conference on Cloud Computing Technology and Science, Dec. 2015.
- [18] S. Tashiro and et al., "A network model for prediction of temperature distribution in data centers," in Proceedings of IEEE 4th International Conference on Cloud Networking (CloudNet), Oct. 2015, pp. 261–266.
- [19] J. D. Anderson and J. Wendt, Computational fluid dynamics. Springer, 1995, vol. 206.
- [20] J. D. Moore, J. S. Chase, P. Ranganathan, and R. K. Sharma, "Making scheduling" cool": Temperature-aware workload placement in data centers," in USENIX annual technical conference, General Track, 2005, pp. 61–75.
- [21] T. Deguchi and et al, "Impact of workload assignment on power consumption in software-defined data center infrastructure," in Proceedings of IEEE CloudNet 2014, Oct 2014, pp. 440–445.
- [22] T. Deguchi and et al., "A workload assignment policy for reducing power consumption in software-defined data center infrastructure," IEICE Transactions on Communications, vol. E99B, no. 2, Feb. 2016, pp. 347–355.
- [23] K. Momose and K. Ikejima, "Development of sensitivity-based decision support system for heat and fluid flow design : Variational approach to sensitivity analysis," in The society of Heating, Air-Conditioning and Sanitary Engineers of Japan, vol. 18, no. 1, 2006, pp. 403–406.
- [24] Y. Cengel, Ed., Heat transfer: a practical approach 2nd edition. Mcgraw-Hill, Oct. 2003, ISBN: 978-0072458930.