A Heat-Recirculation-Aware VM Placement Strategy for Data Centers

Hao Feng[1],Yuhui Deng[1,2]∗,Yi Zhou[3]
1.Department of Computer Science, Jinan University, Guangzhou, China, fh76020634@163.com
2.The State Key Laboratory of Computer Architecture, Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China, tyhdeng@jnu.edu.cn
3.Department of Computer Science, Columbus State University, Columbus, American, zhou_yi@columbussatate.edu

***Abstract*—Data centers consisted of a great number of IT devices (e.g., servers, switches and etc.) which generates a massive amount of heat emission. Due to the special arrangement of racks in the data center, heat-recirculation often occurs between nodes. It can cause a sharp rise in temperature of the equipment coupled with local hot spots in data centers. Existing VM placement strategies can minimize energy consumption of data centers by optimizing resource allocation in terms of multiple physical resources (e.g., memory, bandwidth, cpu and etc.). However, existing strategies ignore the role of heat-recirculation in the data center. To address this problem, in this study, we propose a heat-recirculation-aware VM placement strategy and design a Simulated Annealing Based Algorithm (SABA) to lower the energy consumption of data centers. Different from the existing SA algorithm, SABA optimize the distribution of the initial solution and the way of iteration. We quantitatively evaluate SABA's performance in terms of algorithm efficiency, the activated servers and the energy saving against with XINT-GA algorithm (Thermal-aware task scheduling Strategy), FCFS (First-Come First-Served), and SA. Experimental results indicate that our heat-recirculation-aware VM placement strategy provides a powerful solution for improving energy efficiency of data centers.**

***Index Terms*—data centers, heat recirculation, Virtual Machine, Cloud Computing.**

I. Introduction

Data centers have been widely utilized to provide various services such as Internet-based services and Cloud Computing [1]. Nowadays, data centers consist of thousands of servers are frequently virtualized into a large number of virtual machines (VMs), each of which is committed to serve one particular application or customer. Virtualization of tremendous servers leads to significant energy consumption as well as higher carbon emissions [2]–[4]. For example, Google has deployed more than 16 data centers comprising of up to 900,000 servers around the world. These data centers consumes almost 260 million watts of electric power which can be used to consistently power more than 200,000 homes [5]. To reduce energy consumption of data centers, a common approach is to exploit different VM placement strategies. Hence, a large amount of work has been devoted to investigate VM scheduling algorithms. However, most scheduling algorithms on VMs are dedicated to optimizing the utilization of physical resources (e.g., bandwidth resources, CPU usage, network resources, memory usage), whereas, few techniques were concerned with energy consumption of the cooling system in the data center. Studies of 44 data centers conducted by Salim *et al* showed that the air conditioning system (CRAC) consumes 40% of the total data center energy consumption, this statistic even increased as high as 60% in some least efficient data centers [6]. To address this problem, in this study, we propose an efficient VMs scheduling strategy by considering the heat-recirculation to reduce the energy consumption of the air conditioning system as well as the possibility of local hot spots. Our main contributions are summarized as follows:

- We investigate a novel VM placement strategy, which takes into account heat-recirculation to reduce the energy consumption of a data center. The VMs placement strategy, based on heat-recirculation effect, reduces cooling cost of CRAC in data centers.
- We introduce a Simulated Annealing Based Algorithm (SABA), a heuristic intelligent algorithm, to solve the NP-hard VM placement problem. SABA outperforms the Simulated Annealing algorithm (SA) by adjusting the distribution of the initial solution, which helping in obtaining an approximation of the optimum with fewer iterations.
- We conduct a series of simulation-based experiments to verify the effectiveness of our SABA algorithm. Experimental indicates that compared with the existing solution (e.g., FCFS and XINT-GA), our SABA algorithm can significantly improve CRAC efficiency and reduce energy consumption in the data centers.

The rest of the paper is organized as follows. Section 2 summarizes related work. An introduction and definition of heat-recirculation for data centers is presented in Section 3. Section 4 illustrates our innovative VM placement strategy and demonstrates our SABA algorithm. Section 5 validates the effectiveness of our VMs placement strategy as well as the effectiveness of SABA by a set of experiments. Finally, Section 6 concludes the paper.

II. Related Work

In the past few years, modern data centers consume a dramatic amount of energy. Prior studies have been focused on scheduling algorithms of VMs around the world and most of these attention are paid to balance the computing resources of

*Corresponding author: Yuhui Deng, tyhdeng@jnu.edu.cn

978-3-9819263-4-7/DATE20/©2020 EDAA

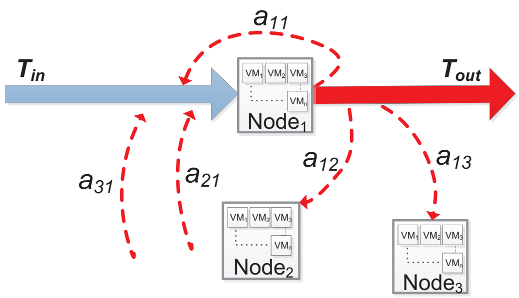


Fig. 1. Heat-recirculation of three nodes.

physical machines and the requirements of VMs [7]–[9]. These studies paid much attention to reduce the use rate of PMs' resources by optimizing the VMs placement in data centers. Unfortunately, those energy-saving solutions are inadequate to boost energy-efficiency of data centers. Our study proposed a thermal-aware VMs placement strategy which distinctly distinguish our energy-saving approach from aforementioned technique.

Very recently, studies revealed that tasks can be scheduled according to the heat-recirculation of servers to save energy consumption in data center [10]–[13]. These techniques adopt the common thermal-aware strategy - optimizing resource allocation by considering heat-recirculation - to cut back energy consumption in data centers. Unfortunately, those scheduling strategies are mostly focusing on tasks (or jobs), ignoring the impact of VMs on energy efficiency of data centers. Different from existing works, we integrated scheduling strategy with VMs factor to reduce energy consumption by leveraging the heat-recirculation.

III. Heat-recirculation

Tang *et al*. model and formulate heat-recirculation by a collection of mathematical equations and denotations [14]. The value of heat-recirculation from $Node_i$ to $Node_j$ is denoted as $a_{ij}Q^i_{out}$(as show in Fig. 1.). Since the heat-recirculation is assumed to be stable, the cross-interference coefficient a_{ij} is accordingly assumed to be a constant. Matrix A=$[a_{ij}]_{n\times m}$ denotes the cross-interference matrix of all nodes, thus A is a constant matrix, which can be used to formalize the heat-recirculation among nodes.

The cross-interference matrix A is calculated by Equation 1 [14]:

$$K\cdot(\vec{T}_{out}-\vec{T}_{sup}) = A'\cdot K\cdot(\vec{T}_{out}-\vec{T}_{sup})+\vec{P}. \quad (1)$$

where A' denotes the transposed matrix of A. $\vec{T}_{out}$ represents the outlet air temperature matrix of each node in the data center and $\vec{T}_{out}$ is expressed as an equation - $\vec{T}_{out}$ = $[T^1_{out}, T^2_{out}, \cdots T^n_{out}]$. $\vec{T}_{sup}$ represents the air temperature matrix supplied by CRAC of data centers and $\vec{T}_{sup}$ conforms to equation ($\vec{T}_{sup}$ = $[T_{sup}, T_{sup}, \cdots T_{sup}]$).

$\vec{P}$ is the energy needed by each node in data centers; $\vec{P}$ is formally expressed as $\vec{P}$ = $[P_1, P_2, \cdots P_n]^T$. K represents the diagonal matrix of the thermodynamic matrix. Formally, we have $K_{n\times n}$ = $diag(K_1, K_2, \cdots K_n)$, where K_i=$\rho f_i C_p$, C_p is the specific heat of air (typical value: $1005JKg^{-1}K^{-1}$), ρ is the air density (typical value: $1.19Kg/m^3$), and f_i is the flow rate (the speed of the air) of $Node_i$ (typical value: 520, *CFM*=0.2454m^3/s).

We assume that the VMs are homogeneous. Given a task of a total number of VMs (denoted as C_{tot}) which arrived at a data center in one period. Distributing these VMs on different servers will generate different CRAC energy consumption owing to the heat-recirculation effect within numerous nodes. For a multicore system, we can easily split the total VMs by the number of cores of each server. A scheduler assigns the VMs into n nodes, and each node will execute a "subtask" of VMs with size c_i.

Each node can bear at most m VMs, the scheduling results should be under the constraints:

$$\sum_{i=1}^{n} c_i - C_{tot} = 0, and\ c_i \leq m. \quad (2)$$

The relationship of the maximal temperature of airflow supplied by the CRAC T_{sup} and VMs placement strategy can be expressed as:

$$maximize\ T_{sup}, and\ \vec{t}_{sup} = D\vec{b} + D\vec{c}a - \vec{t}_{in}, \quad (3)$$

where a denotes the energy consumption by server $Node_i$ with m VMs (the utilization of this CPU is 100%) running on it, $\vec{b}$ represents the basic energy consumption of node $Node_i$, and D is the heat distribution matrix; D can be derived from A as:

$$D = [(K - A^T K)^{-1} - K^{-1}]. \quad (4)$$

Because $\vec{t}_{in}$, $D\vec{b}$ and a can be obtained in a data center, we can calculate a suitable VMs placement vector $\vec{c} = \{c_1, c_2, ...c_n\}$ to maximize the CRAC outlet airflow temperature. Moreover, Wang *et al.* evaluate CRAC energy consumption as a function of supply temperatures T_{sup} [15]. Wang *et al.* ' findings indicate that the percentage of energy savings achieves up to 4.3-9.8% per 1°C rise in temperature.

IV. Thermal-aware VMs placement strategy

In recent years, there have been various types of scheduling schemes and algorithms for VMs placement. However, most scheduling algorithms on VMs are devoted to cut down energy consumption by optimizing physical resources usage (such as bandwidth resources, network resources, CPU usage, and memory usage, etc.). Unfortunately, such strategies ignore the influence of the cooling system on energy consumption in data centers.

Equation 3 outlined above suggests that various VM scheduling strategies causes different CRAC temperature T_{sup}, which results in a different cooling energy consumption for a data center. This observation motivates us to investigate an efficient VMs placement strategy by which CRAC can provides a higher T_{sup} to meet the need for cooling down the servers.

A. The Simulated Annealing algorithm

It is widely known that finding an efficient VMs placement strategy is an NP-hard problem, we firstly investigate the Simulated Annealing algorithm (SA) to solve this NP-hard problem.

SA is an iterative heuristic algorithm which is employed to obtain an approximately optimal solution of a function. Meanwhile, *SA* is a variant of greedy algorithms (which make the optimal choice at each step as it attempts to find the overall optimal way to solve the entire problem). In our study, we tailor the SA algorithm to solve VMs scheduling problem. Our goal is to reduce CRAC energy consumption by obtaining the highest T_{sup} (See the details in Equation 3 in Section 3). In a heat-recirculation environment, various parameters and functions of the SA algorithm are reassigned as two following parts:

1) *Setting of Control Parameters*:
   The main parameters needed to be set are, namely, the initial temperature T_0, the rate of cooling q, the number of iterations and the end temperature T_{red}.
2) *Initial Solution*:
   We use random arrangement to generate one initial solution $\vec{c}_1$. Given the need of total VMs C_{tot} by workloads, the distribution of VMs ($\forall \vec{c}_1$in C_{tot}) needs to be guaranteed ($C_{tot} = \sum_{j=1}^{n} c_j$).

B. The SABA algorithm

Derived from the SA algorithm, we develop a novel algorithm - the Simulated annealing based algorithm (SABA) - which takes the heat-recirculation of VMs scheduling into consideration. Compared with the SA algorithm, our SABA effectively reduces the number of iterations and runtime of the SA algorithm by following improvements:

1) In the step of generating the initial solution, we assign a relatively large value to the position near the top of the rack. By doing so, compared to SA, the SABA algorithm generates the initial value which accelerates its process toward obtaining the final optimal solution at very first steps.
2) In the step of generating a new solution, if the new solution assigns fewer values to the location near the top rack after the transformation strategy, the transformation will not be performed. That is, the SABA algorithm prevents the current solution from a change in direction which is not conducive for SABA to obtain the optimal global solution.
3) Comparing with the SA algorithm, if the CPU utilization is not high, SABA will start fewer nodes to schedule VMs. Therefore, SABA increases the utilization of nodes while ensuring the completion of tasks.

V. Experiment Simulation

In our extensive experiments, we build a simulation resembling a real-world data center to evaluate the effectiveness of our heat-recirculation-aware VM placement strategy.

A. Performance Evaluation

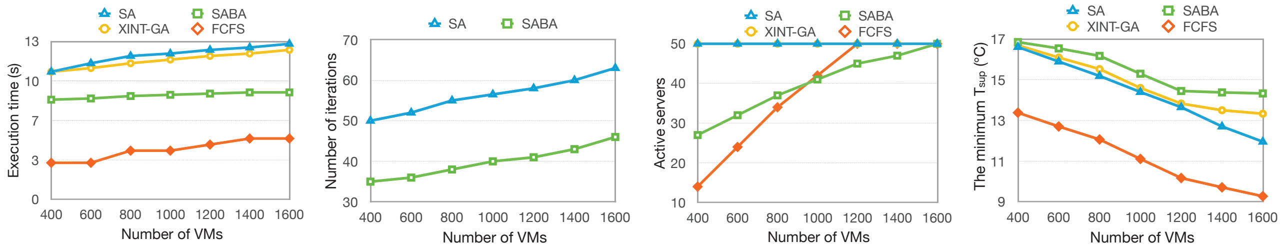
We compare SABA algorithm with existing algorithms, namely, SA, FCFS and XINT-GA [14], to evaluate the energy efficiency and performance of our VMs placement strategy through SABA. First Come First Served algorithm (FCFS) has been widely used in the field of VM scheduling. FCFS schedules VMs according to the order of the arrival times of the VMs. Thus, VM that enters the system first is selected first. Unlike FCFS, XINT-GA algorithm is designed for distributing arrival jobs among servers in order to maximize CRAC cooling efficiency by keeping the supply temperature as high as possible. To resemble servers adopted in a real-world data-center environment, we also simulate jobs utilized in algorithms(e.g., XINT-GA).

To make a fair comparison, we set sizes of VMs from 200 to 1600; And these VMs arrive in the same period. Because the number of VMs needed in tasks can not exceed the maximum number that the simulated data center can handle (i.e.,2000), scheduling algorithms is employed by VMTS to address such resource scarcity problem by scheduling tasks in a short time period. Concerned with above scheduling issue, we also make a reasonable assumption that all VMs enter into the wait queue in the same period, then scheduled by VMTS. Moreover, we perform more than 20 calculations to obtain average values of measurements in each of this group experiments.

Fig. 2(a) shows the execution time comparison between algorithms SA, SABA, XINT-GA and FCFS. The results plotted in Fig. 2(a) indicate that the execution time of these four algorithms increases as the number of VMs increases. The execution time of algorithms SA, SABA and XINT-GA is higher than that of FCFS algorithm. This trend is expected, because SA, SABA and XINT-GA are intelligent iterative algorithms, and such algorithms take a longer time to obtain the optimal solution.

We observe from Fig. 2(b) that algorithm SABA is superior to algorithms SA and XINT-GA in terms of iterations. This result is contributed by two facts: First, SABA algorithm optimizes both the initial solution and conditions of generating a new solution. Second, since the initial solution produced by SABA is close to the final solution, SABA performs a much smaller number of iterations than SA and XINT-GA do to obtain the final solution.

Fig. 2(c) demonstrates the comparison of active servers between SABA, SA, XINT-GA and FCFS. In this set of experiments, we set the utilization of server where FCFS algorithm is adopted to 60%. Fig. 2(c) indicates that when the number of VMs is small (i.e., 400 VMs), the number of servers activated by the FCFS algorithm is the smallest among these four algorithms. However, as the number of VMs increases, SABA launches fewest servers to handle tasks when the number of VMs exceeds 1000. Unlike SABA and FCFS, SA and XINT-GA start a large number of servers, even when the number of tasks is small (i.e.,400). This changeless active servers trend is attributed to the fact that SA and XINT-GA algorithms only focus on the result of the



(a) Comparison of Execution time (b) Comparison of Frequency of Iterations (c) Comparison of Active servers (d) Comparison of Maximum T_{sup}

Fig. 2. Experiments under VMs arriving in the same period

optimization, ignoring the factor of number of active servers. In the meantime, only a meager amount of tasks are running on each server (even as low as 5%). These two facts combined together give rise to a enormous waste of energy consumption.

Lastly and most importantly, we investigate the impact of maximum CRAC temperature T_{sup} on the energy savings of a date center under various algorithms (i.e., SA, SABA, XINT-GA and FCFS). We argue that a higher temperature T_{sup} provided by CRAC implies a higer CRAC energy saving. Therefore, Fig. 2(d) suggests that a data center employed our SABA yields the highest energy savings.

In particular, SABA algorithm provides an average of 4 degrees higher than FCFS algorithm does. Since increasing the temperature by just one degree can save up to 2%-5% of the energy consumption [16], SABA saves the energy consumption of FCFS by approximately to 8%-20%. Although SA and XINT-GA algorithms also show good performances, SABA outperform the other three algorithms.

VI. Conclusions

In this study, we propose a heat-recirculation-aware VM placement strategy for data centers. This strategy is built to minimize the energy consumption and avoid hotspots within a data center. We conducted extensive experiments to evaluate the effectiveness of our heat-recirculation-aware VM placement strategy. Experimental results demonstrate that our SABA outperforms other three approaches (SA, XINT-GA and FCFS) in terms of the minimum temperature air supplied by CRAC, algorithm execution time, algorithm iterations and the number of active servers.

Acknowledgment

This work is supported by the NSFC (no.61572232), in part by the Science and Technology Planning Project of Guangzhou (no.201802010028, and no.201802010058), and in part by the Open Research Fund of Key Laboratory of Computer System and Architecture, Institute of Computing Technology, Chinese Academy of Sciences under Grant(CARCH201705).

References

[1] Y. Gong, B. He, and L. Dan, “Network performance aware optimizations on iaas clouds,” *IEEE Transactions on Computers*, vol. 66, no. 4, pp. 672–687, 2017.

[2] X. Li, P. Garraghan, X. Jiang, Z. Wu, and J. Xu, “Holistic virtual machine scheduling in cloud datacenters towards minimizing total energy,” *IEEE Transactions on Parallel and Distributed Systems*, vol. 29, no. 6, pp. 1–1, 2018.

[3] Y. Deng, “What is the future of disk drives, death or rebirth?” *ACM Computing Surveys (CSUR)*, vol. 43, no. 3, p. 23, 2011.

[4] H. Feng, Y. Deng, and Y. Zhao, “Criso: An incremental scalable and cost-effective data center interconnection by using 2-port servers and low-end switches,” in *2018 IEEE 26th International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems (MASCOTS)*. IEEE, 2018, pp. 123–130.

[5] Z. Hu, B. Li, and J. Luo, “Time- and cost- efficient task scheduling across geo-distributed data centers,” *IEEE Transactions on Parallel and Distributed Systems*, vol. PP, no. 99, pp. 1–1, 2018.

[6] W. Shi, C. Wu, and Z. Li, “A shapley-value mechanism for bandwidth on demand between datacenters,” *IEEE Transactions on Cloud Computing*, vol. 6, no. 1, pp. 19–32, 2018.

[7] S. Sotiriadis, N. Bessis, and R. Buyya, “Self managed virtual machine scheduling in cloud systems,” *Information Sciences*, 2017.

[8] W. Lloyd, S. Pallickara, O. David, M. Arabi, and K. Rojas, “Dynamic scaling for service oriented applications: Implications of virtual machine placement on iaas clouds,” in *IEEE International Conference on Cloud Engineering*, 2014, pp. 271–276.

[9] W. Fang, X. Liang, S. Li, L. Chiaraviglio, and N. Xiong, “Vmplanner: Optimizing virtual machine placement and traffic flow routing to reduce network power costs in cloud data centers,” *Computer Networks*, vol. 57, no. 1, pp. 179–196, 2013.

[10] A. Hameed, A. Khoshkbarforoushha, R. Ranjan, P. P. Jayaraman, J. Kolodziej, P. Balaji, S. Zeadally, Q. M. Malluhi, N. Tziritas, and A. Vishnu, “A survey and taxonomy on energy efficient resource allocation techniques for cloud computing systems,” *Computing*, vol. 98, no. 7, pp. 751–774, 2016.

[11] Q. Fang, J. Wang, Q. Gong, and M. Song, “Thermal-aware energy management of hpc data center via two-time-scale control,” *IEEE Transactions on Industrial Informatics*, vol. PP, no. 99, pp. 1–1, 2017.

[12] Q. Tang, S. K. S. Gupta, D. Stanzione, and P. Cayton, “Thermal-aware task scheduling to minimize energy usage of blade server based datacenters,” in *IEEE International Symposium on Dependable, Autonomic and Secure Computing*, 2015, pp. 195–202.

[13] H. Feng, Y. Deng, and L. Yu, “Modeling the failures of power-aware data centers by leveraging heat recirculation,” *Concurrency and Computation: Practice and Experience*, no. 4, 2019.

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

[15] N. Wang, J. Zhang, and X. Xia, “Energy consumption of air conditioners at different temperature set points,” *Energy and Buildings*, vol. 65, no. 65, pp. 412–418, 2013.

[16] N. El-Sayed, I. A. Stefanovici, G. Amvrosiadis, A. A. Hwang, and B. Schroeder, “Temperature management in data centers: why some (might) like it hot,” *ACM SIGMETRICS Performance Evaluation Review*, vol. 40, no. 1, pp. 163–174, 2012.