Optimization of Workload Distribution of Data Centers Based on a Self-Learning In Situ Adaptive Tabulation Model

Xu Han[1], Wei Tian[2], Wangda Zuo[1], James W. VanGilder[2]
[1]University of Colorado Boulder, Boulder, CO, USA
[2]Schneider Electric, Andover, MA, USA

Abstract

Data center cooling typically involves non-uniform airflow and temperature distributions, which are affected by the IT workload distribution. It is helpful to simulate the airflow and temperature to optimize the workload distribution. Traditional computational fluid dynamics (CFD) simulation is usually time-consuming while conventional reduced order models (ROMs), though computationally fast, may generate inaccurate results even after being fully trained. In Situ Adaptive Tabulation (ISAT), contracting to conventional ROM, can make prediction with error lower than a user-specified tolerance. To demonstrate using of ISAT for optimal workload distribution in data center, this paper presents a preliminary study of an ISAT-based genetic algorithm optimization platform. The ISAT is trained offline by using the results from CFD simulations using a hypothetical simple data center. The optimal workload distribution determined by the platform leads to approximately 6.8% of energy savings when compared to the benchmark with a uniform workload distribution. We note that the time cost for the entire optimization process, including the training of ISAT is about 4 hours, which is acceptable in the design phase.

Introduction

ASHRAE (2015) recommends 27 ºC as the maximum inlet temperature for the IT equipment. In reality, computer room air conditioners (CRACs) in many data centers are operated at a significantly lower supply temperature to avoid hot spots. While providing redundant cooling, this leads to reduced cooling efficiency (Tang et al. 2008). Therefore, it has been a research topic on how to improve the energy efficiency while still providing enough cooling to the data center whitespace, where the IT equipment is placed. One of the approaches is to optimize the workload distribution such that CRACs can operate at a higher supply temperature or lower airflow rate (Banerjee et al. 2010).

Moore et al. (2005) proposed several algorithms for workload placement to maximize the supply air temperature while ensuring the server inlet temperature not exceeding the threshold. Tang et al. (2007) investigated several load placement strategies to minimize heat recirculation. Among these research, the CFD simulations were conducted to obtain the thermal map of the data center with regard to workload distribution, which required huge computational efforts to cover a large number of scenarios.

To reduce the computational cost, simplified airflow models were proposed for fast prediction of indoor environment. Zhang et al. (2009) developed a lumped model to predict rack-inlet temperatures under various supply airflow and rack-load scenarios. Since only one effective aggregated rack and cooling unit were considered based on a well-mixed assumption, the lumped model may fail to capture local airflow pattern and temperature distribution. Potential flow models were used to predict the airflow and temperature distribution (Toulouse et al. 2009). However, the potential flow model is an approximate and simplified model, which may not be as accurate as CFD, particularly when simulating the jet flows from the perforated tiles.

Other than using physic-based models to generate simplified models, statistics-based approaches, sometimes referred to as reduced order models (ROMs) were also employed. Proper orthogonal decomposition (POD) was used to predict the velocity and temperature distribution in an office (Elhadidi and Khalifa 2005) and rack-inlet temperature distribution in a raise-floor data center (Demetriou and Khalifa 2013). Artificial neural network (ANN) models were used to predict thermal map of data centers (Moore et al. 2006), perforated-tile flowrates and rack-inlet temperatures (Song et al. 2011). However, the predictions of such statistics-based models may become inaccurate when the queries lie beyond the training domain.

To overcome the limitations of these simplified models, an online self-learning ROM called in situ adaptive tabulation (ISAT) was proposed. ISAT is a storage-and-retrieval algorithm that was originally developed by Pope (1997) to accelerate turbulent combustion simulations. Tian et al. (2018) applied it in the indoor environment simulation and presented promising results in terms of its training and prediction. ISAT, which allows user to specify an error tolerance, retrieves the output of a query using linear interpolation if the error of such retrieval is estimated within the error-tolerance. Otherwise, as opposed to the conventional ROMs that would continue the retrieving, ISAT will call a full-scale simulator (such as CFD) to resolve the query. We will cover the detail of the ISAT in the coming section.

In this paper, we present an optimization platform based on ISAT and genetic algorithm to optimize workload

distribution in a hypothetical data center. To accelerate the training of ISAT, we employ a fast fluid dynamics (FFD) model running on graphics processing unit (GPU) (Tian et al. 2017; Tian, VanGilder, Han, et al. 2019; Tian, VanGilder, Condor, et al. 2019) to simulate the airflow and temperature distributions under different workload distributions. With the optimal workload distribution determined by the platform, CRAC supply-air temperature can be increased approximately by 0.8 ℃, and this leads to 6.8% cooling energy savings.

Methodology

Introduction of ISAT and FFD

The idea of ISAT is to dynamically establish a lookup table, based on which a new query can be quickly evaluated by local linear regressions (Pope 1997).

In both training and evaluation stages, to answer a new query, ISAT will first determine the nearest record to the query point. If a query point is within the Ellipsoid of Accuracy (EOA) of the nearest record in the ISAT table, ISAT will estimate outputs using linear approximation (*retrieve*). If the query point is outside the EOA of the nearest record, ISAT will simultaneously perform a linear approximation (*retrieve*) and a direct evaluation by calling CFD simulations to determine the outputs.

Afterwards, the error between the solutions from *retrieve* and direct evaluation will be compared. If the difference is less than the total error tolerance, a *grow* action will be executed to enlarge the current EOA of the record to include the query point; Otherwise, an *add* action will be performed to store the query point and its related data as a new record in the ISAT table. The detailed description of the workflow of the ISAT algorithm can be found in (Tian et al. 2018).

To accelerate the training process, we employ FFD running on graphics processing unit to perform the airflow and thermal simulation. FFD is a faster alternative to traditional CFD. FFD solves the same governing equations as CFD does, but employs a different solving technique. FFD was reported to be 50 times faster than CFD and an additional factor of 30-1000 times speedup can be achieved using parallel computing (Zuo and Chen 2010; Tian et al. 2017). Tian, VanGilder, Han, et al. (2019) adopted FFD to simulate data center airflow and found that FFD could achieve comparable accuracy to CFD with the potential of significantly reduced solution time. The readers may refer to (Zuo and Chen 2009; Tian, VanGilder, Han, et al. 2019) for detailed description of FFD.

Optimization platform based on ISAT-FFD

The idea of ISAT-FFD is that the prediction can be retrieved based on an existing data record if the estimated error is within a pre-defined tolerance. Otherwise, the prediction will be obtained by calling FFD to conduct an airflow and thermal simulation. We have conducted some preliminary studies to evaluate using of ISAT-FFD for indoor airflow simulations (Tian et al. 2018).
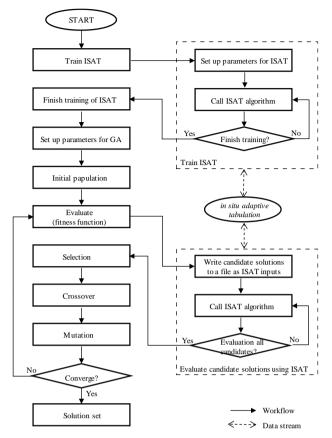

Figure 1: Framework of the ISAT-based genetic algorithm (GA) optimization platform

In this study, the optimization is performed based on ISAT after it is fully trained offline using FFD simulations. As shown in Figure 1, an optimization platform that links ISAT-FFD with genetic algorithm (Mitchell 1998) is developed to automate the whole process from training/evaluation of ISAT to genetic algorithm optimization.

The conventional airflow optimization evaluates the fitness function (also called objective function, which may be, for example, best thermal comfort) by directly performing airflow simulations (e.g. CFD or FFD) to predict the indoor velocity and temperature distribution and then calculate the predicted mean vote (PMV) to evaluate the thermal comfort. Different from conventional method, the idea of this optimization platform is that we first train a reduced order model ISAT by a series of data set extracted from airflow simulations. For example, we may use a series of data set including the boundary conditions (as inputs) and the corresponding average velocity and temperature at the occupant zone (as outputs) to train the ISAT and store data in a database. Then the ISAT should be able to predict the average velocity and temperature at the occupant zone for a new boundary condition based on existing neighboring points in the database.

After the training finishes, the genetic algorithm is called to perform optimization as shown in Figure 1. The genetic algorithm first generates a series of candidate solutions (*initial population*), which are then evaluated and sorted with fitness function predicted by ISAT. Through

selection, *crossover* and *mutation*, better solutions will be inherited to the next generation with a higher probability. After iterations of the above procedure until the genetic algorithm is converged, the optimal solution will be exported. The advantage of this platform is that a fully-trained ISAT should be able to evaluate the performance of candidate solutions using the *retrieve* action (linear regression) for most cases, which is much faster than full airflow simulations.

Case Study

In the data center airflow management area, the rack-inlet temperature is a critical parameter to evaluate the cooling performance, which can be influenced by multiple factors, such as, CRAC supply air temperature, and airflow pattern near the racks (which may lead to mixture of hot air with cool air). The maximum rack-inlet temperature should not exceed 27 ℃ per ASHRAE specs (ASHRAE 2015). As a result, if the maximum rack-inlet temperature is raised due to local hot spots caused by mixture of hot air with cold air in front of the racks, the cooling system has to lower supply air temperature or increase supply flowrate, which leads to a lower efficiency of the cooling system.

The workload distribution, which this paper is focusing on, has significant effect on the airflow pattern near the racks and further influences the rack-inlet temperature. Therefore, we could eliminate the hot spots by optimizing workload distribution and increase the supply air temperature to improve the cooling.

FFD Modeling of the hypothetical data center

The whitespace (data center room where the IT equipment is placed) is 6.7m, 5.5m, 3.7m in length, width, and height, respectively. The total whitespace floor area of the reference data center is 36.8 m^2 (396 ft^2) and the layout is shown in Figure 2. The data center includes 10 racks with a total power of 60 kW. CRAC pumps cool airflow to the raised-floor plenum, and the supply airflows is distributed to the inlet of racks through 10 25%-open-area perforated tiles. The cool airflow is then drawn into the racks, heated, and extracted from hot aisles back to the CRAC directly through the room without a ceiling plenum. We note that this hypothetical data center is proposed to reflect a typical data-center layout.

The total workload is evenly distributed over 10 racks in the benchmark. The IT airflow is estimated as rack power multiplying by 212 m^3/h/kW (125 cfm/kW). The total supply airflow rate is 12,750 m^3/h, which leads to an air-ratio of 1.

For modeling simplicity/convenience, we made the following simplifications. First, only whitespace is modeled, and the underfloor plenum is excluded. Since all the perforated tiles in this case are 25% open-area-ratio, to model the whitespace and under-floor plenum in a detached way is justifiable, as the airflows in the two spaces are fairly separated (Tian, VanGilder, Han, et al.

2019). Second, we employed a black-box rack model and therefore workload is shifted from rack-rack (as opposed to IT-to-IT). For the detailed description of the rack model, refer to literature (Tian, VanGilder, Condor, et al. 2019).
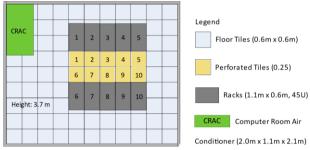

Figure 2: Plan view of the hypothetical simple data center

Settings for FFD Simulations

As recommended by VanGilder and Zhang (2008), FFD simulation adopted an uniform 6 inches (0.15 m) grid, resulting into a total of 38,016 cells. Since FFD simulation is inherently transient; we use a time step size of 0.1 s and simulation time of 200 s, which is adequately long to reach steady-state conditions. FFD employs a zero-equation turbulence model (Dhoot et al. 2016). The FFD simulation was performed on a workstation with 4 Intel(R) Xeon(R) E5-1603 CPUs and an AMD FirePro W8100 GPU and the results of initial tests showed that the time cost of a FFD simulation is about 61 s.

Training of ISAT-FFD

From sensitivity analysis, we found that the IT power of the racks located at the ends of the rows have relatively larger influence on the overall maximum rack-inlet temperature. We assume that the powers of *Rack-1* and *Rack-6* are always the same (represented by *Input 1*) and the powers of *Rack-5* and *Rack-10* are always the same (represented by *Input 2*). For other racks, we assume that the remaining IT power is uniformly distributed over the 6 racks in the middle of the rows.

Table 1: Results of training ISAT-FFD models for the three cases

Items\Cases	Case 1-1	Case 1-2	Case 1-3
No. of Queries	255,497	15,630	4,190
No. of *Retrieve*	253,673	15,433	4,103
No. of *Grow*	1,542	160	63
No. of *Add*	282	37	24
Training Time (hr.)	34.7	3.8	1.8

We proposed three cases with predefined total error tolerances: 0.2 ℃, 0.4 ℃, and 0.6 ℃ for *Case 1-1, 1-2*, and *1-3*, respectively. The two inputs range from 1.0 kW to 10.0 kW. The output is the maximum rack-inlet temperature of the 10 racks. The results of the three cases are shown in Table 1*Table 1*. In *Case 1-1*, the total number of queries used in the training is 255,497, in which 99.3% are *retrieve* actions, 0.6% are *grow* actions and 0.1% are *add* actions. In *Case 1-2* and *Case 1-3*, the total numbers of queries drastically dropped to 15,630 and

4,190. Consequently, the numbers of the *retrieve*, *grow* and *add* actions are also significantly decreased. As a result, the training process takes 34.7 hours for *Case 1-1*, and only 3.8 hours and 1.8 hours for *Case 1-2* and *Case 1-3*.
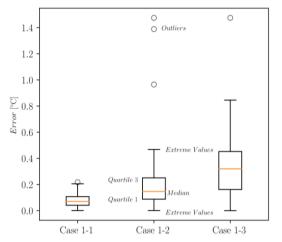

Figure 3: Error distributions of evaluation points for the three cases

After the training was completed, we evaluated the accuracy of the ISAT table using 50 randomly-generated data points (the same 50 data points for each case) and correlated the prediction accuracy with the user-defined total tolerance in the training. As shown in Figure 3, the errors of ISAT predictions for the evaluated points are generally within 0.2 ℃, 0.45 ℃ and 0.85 ℃ for *Case 1-1*, *Case 1-2* and *Case 1-3*, respectively. Moreover, 75% of the predictions of the queries are within 0.12 ℃, 0.25 ℃ and 0.45 ℃ for the three cases, respectively. Generally, we conclude that the ISAT prediction error is at the same order of the total error tolerance specified at training stage. To achieve a balance between prediction accuracy and training cost, we think that 0.4 ℃ is a proper error tolerance for this specific case.

Optimization of Workload Distribution

The optimization of the rack workload distribution was performed based on the trained ISAT. Here we note that although we chose 0.4 ℃ error tolerance in *Case 1-2* for training of ISAT, we also include the optimization results based on *Case 1-1* and *1-3* for comparison.

We set in GA the population size to be 30 and the number of generations to be 100. The *crossover* and *mutation* probabilities are set to be 0.6 and 0.1, respectively. The fitness function of the optimization study is to minimize the maximum rack-inlet temperature when the total supply flow rate of the CRAC remains unchanged.

The optimization results are shown in Table 2. Generally, the three cases obtain similar optimal solutions. The optimal values of *Input 1* and *Input 2* are in the range of about 1.0-1.3 kW and 9.6-9.9 kW, respectively. The maximum rack-inlet temperatures with optimal workload distributions are 15.4 ℃ from ISAT predictions for the three cases. To evaluate the accuracy of the ISAT-based optimization, we performed a FFD-based optimization. The results show that there is only 0.1 ℃ difference

between the ISAT predictions and the results from FFD simulations for the three cases.

Table 2: Optimization results of the three cases

Items\Cases	Case 1-1	Case 1-2	Case 1-3
Fitness Function	Minimize *maximum rack-inlet temperature*		
Input 1 (kW)	1.22	1.01	1.27
Input 2 (kW)	9.57	9.9	9.88
ISAT Output (℃)	15.4	15.4	15.4
Direct FFD Simulation (℃)	15.5	15.5	15.5

Figure 4 compares the time costs for performing optimization in different cases. The time cost of FFD-based optimization (*FFD-GA*), which directly calls FFD simulations to perform optimization, is roughly estimated by the time cost of each FFD simulation (61s) and the total number of FFD calls (3000) during the optimization. *Case 1-1*, *Case 1-2*, and *Case 1-3*, which first train ISAT using FFD and then perform optimization by calling ISAT, take only 69.1%, 7.7% and 7.1% of the time of the FFD-GA.
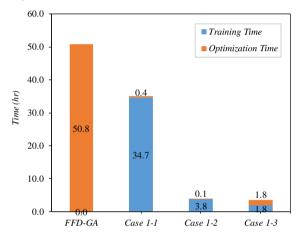

Figure 4: Comparison of time costs for FFD-based optimization (FFD-GA) and ISAT-based optimization (Case 1-1, Case 1-2 and Case 1-3)

Evaluation of Energy-Saving Potential

The CRAC supply air temperature can be scaled up until the rack-inlet temperature reaches maximum threshold recommended by ASHRAE to improve cooling efficiency. In the present case, the maximum rack-inlet temperature is decreased by 0.8 ℃ from the benchmark by optimizing workload. Therefore, the supply air temperature of CRACs can be raised from 20.7 ℃ to 21.5 ℃. Based on the empirical formula describe in (Moore et al. 2005), the coefficient of performance (COP) is estimated to be increased from 3.39 to 3.63, which results in an energy saving of 1.1 kW. The cooling energy saving percentage is estimated to be 6.8% by optimizing workload distribution compared to the benchmark with a uniform workload distribution.

Conclusion

This paper proposes an optimization platform that integrates ISAT with a global optimization engine to seek

the optimal IT workload distribution. The platform is demonstrated using a hypothetical data center, in which ISAT is trained by using FFD simulations. With an error tolerance of 0.4 ºC, the training cost of ISAT is approximately 3.8 hours and the optimization takes an additional 0.1 hours. Compared to the benchmark having a uniform workload distribution over racks, the optimal workload distribution results into an increase of supply air temperature by 0.8 ºC, and this is estimated to save 6.8% of cooling energy. In the future, we can further this study by exploring advanced ISAT training method to reduce training time cost and perform more case studies using real data center layouts.

Acknowledgement

Disclaimer

References

ASHRAE. 2015. *Thermal Guidelines for Data Processing Environments*. Fourth Edition ed.

Banerjee, A., T. Mukherjee, G. Varsamopoulos, and S. K. Gupta. 2010. Cooling-Aware and Thermal-Aware Workload Placement for Green Hpc Data Centers. *Proceedings of the Green Computing Conference, 2010 International*.

Demetriou, D. W., and H. E. Khalifa. 2013. Thermally Aware, Energy-Based Load Placement in Open-Aisle, Air-Cooled Data Centers. *Journal of Electronic Packaging,* 135 (3):030906.

Dhoot, P., J. W. VanGilder, Z. Pardey, and C. M. Healey. 2016. "Zero-Equation Turbulence Models for Large Electrical and Electronics Enclosure Applications." In.: Schneider Eletric.

Elhadidi, B., and H. E. Khalifa. 2005. Application of Proper Orthogonal Decomposition to Indoor Airflows. *Ashrae Transactions,* 111 (1).

Mitchell, M. 1998. *An Introduction to Genetic Algorithms*: MIT press.

Moore, J., J. S. Chase, and P. Ranganathan. 2006. Weatherman: Automated, Online and Predictive Thermal Mapping and Management for Data Centers. *Proceedings of the 2006 IEEE International Conference on Autonomic Computing*.

Moore, J. D., J. S. Chase, P. Ranganathan, and R. K. Sharma. 2005. Making Scheduling" Cool": Temperature-Aware Workload Placement in Data Centers. *Proceedings of the USENIX annual technical conference, General Track*.

Pope, S. B. 1997. Computationally Efficient Implementation of Combustion Chemistry Using in Situ Adaptive Tabulation.

Song, Z., B. T. Murray, and B. Sammakia. 2011. Multivariate Prediction of Airflow and Temperature Distributions Using Artificial Neural Networks. *Proceedings of the ASME 2011 Pacific Rim Technical Conference and Exhibition on Packaging and Integration of Electronic and Photonic Systems*.

Tang, Q., S. K. Gupta, and G. Varsamopoulos. 2007. Thermal-Aware Task Scheduling for Data Centers through Minimizing Heat Recirculation. *Proceedings of the Cluster Computing, 2007 IEEE International Conference on*.

Tang, Q., S. K. S. Gupta, and G. Varsamopoulos. 2008. Energy-Efficient Thermal-Aware Task Scheduling for Homogeneous High-Performance Computing Data Centers: A Cyber-Physical Approach. *IEEE Transactions on Parallel and Distributed Systems,* 19 (11):1458-72.

Tian, W., T. A. Sevilla, D. Li, W. Zuo, and M. Wetter. 2018. Fast and Self-Learning Indoor Airflow Simulation Based on in Situ Adaptive Tabulation. *Journal of Building Performance Simulation,* 11 (1):99-112.

Tian, W., T. A. Sevilla, and W. Zuo. 2017. A Systematic Evaluation of Accelerating Indoor Airflow Simulations Using Cross-Platform Parallel Computing. *Journal of Building Performance Simulation,* 10 (3):243-55.

Tian, W., J. W. VanGilder, M. Condor, X. Han, and W. Zuo. 2019. "An Accurate Fast Fluid Dynamics Model for Data Center Applications." In *IEEE ITherm Conference*. Las Vegas, NV.

Tian, W., J. W. VanGilder, X. Han, C. M. Healey, M. Condor, and W. Zuo. 2019. "Fast Fluid Dynamics Modeling of Data-Center Floor Plenums." In *2019 ASHRAE Winter Conference*. Atlanta, GA.

Toulouse, M. M., G. Doljac, V. P. Carey, and C. Bash. 2009. Exploration of a Potential-Flow-Based

Compact Model of Air-Flow Transport in Data Centers. *Proceedings of the ASME 2009 International Mechanical Engineering Congress and Exposition*.

VanGilder, J. W., and X. S. Zhang. 2008. Coarse-Grid CFD: The Effect of Grid Size on Data Center Modeling. *Ashrae Transactions,* 114 (2):166-82.

Zhang, X. S., J. W. VanGilder, and C. M. Healey. 2009. A Real-Time Data Center Airflow and Energy Assessment Tool. *Proceedings of the ASME 2009 InterPACK Conference collocated with the ASME 2009 Summer Heat Transfer Conference and the ASME 2009 3rd International Conference on Energy Sustainability*.

Zuo, W., and Q. Chen. 2009. Real-Time or Faster-Than-Real-Time Simulation of Airflow in Buildings. *Indoor air,* 19 1:33-44.

Zuo, W., and Q. Chen. 2010. Fast and Informative Flow Simulations in a Building by Using Fast Fluid Dynamics Model on Graphics Processing Unit. *Building and Environment,* 45 (3):747-57.