Energy and Buildings 136 (2017) 86–99

Contents lists available at ScienceDirect

Energy and Buildings

journal homepage: www.elsevier.com/locate/enbuild

Reduced order modeling of a data center model with multi-Parameters

Long Phan, Cheng-Xian Lin*

Department of Mechanical and Materials Engineering, Florida International University, 10555W. Flagler St., EC 3445, Miami, FL 33174, United States

ARTICLE INFO

Article history:
Received 19 May 2016
Received in revised form 29 October 2016
Accepted 23 November 2016
Available online 24 November 2016

Keywords:
Data center
Proper orthogonal decomposition
Computational fluid dynamics
Thermal management

ABSTRACT

Although full-field simulations using computational fluid dynamics and heat transfer (CFD/HT) tools can be applied to predict the flow and temperature fields inside data centers, their running time remain the biggest challenge to most modelers. In this paper, a reduced order modeling method is used to drastically reduce the running time (up to 600 times faster) while still acceptably preserving the accuracy of the model. The results obtained from Proper Orthogonal Decomposition (POD) method are in good agreement with the results obtained from CFD/HT model simulation. The sensitivity analysis of some of the design parameters in the POD model is evaluated. In addition, a 3-D temperature profile of the data center model constructed from 2D slices are generated with a linear interpolation technique. Tradeoff between accuracy and running time is observed and discussed.

1. Introduction

Ever since the booming era of information technology (IT), computers and electronic devices have become a very essential part of each and every one of us on a daily basis. Since information is stored and transferred from several servers from many locations in the world, the need for server housing or data centers is of paramount importance. Thermal management in data centers today is still an underlying problem in terms of energy efficiency. Over the years, a great deal of efforts have been put in to manipulate and improve the energy consumption. Nonetheless, the enhancement is still very limited, though a number of methods and strategies have been proposed by several researchers of many institutions.

A data center is a facility where the server racks or computer systems are hosted to use for telecommunications. Its infrastructure includes many components such as uninterruptible power supply (UPS) system, environment control system (e.g., chiller plant, hot and cold lines, computer room air conditioner, fire suppression), and security and alarm devices (e.g., sensors, cameras). Data centers are constantly powered by redundant or backup power supply. Therefore, its energy consumption is typically huge compared to any other building types. According to the US Congress [1], the Environment Protection Agency (EPA) has estimated that over 100 billion kWh at a cost of $7.4 billion would be consumed by 2011 on data centers. Beside the demanding electric energy to power up IT equipment (e.g. server hardware, storage, UPS, etc.) in these data centers, the major factor that accounts for these costs is the cooling electric energy consumption. Nowadays, power usage effectiveness (PUE) is a common method to calculate a data center's energy usage. It is defined as total electric power/server power. The total electric power consists of server power, cooling power, and power distribution losses. Typically, the value for PUE is close to 2, which also means that the cooling power takes up to 50% of the total electric energy consumption in a modern data center [2].

Computational fluid dynamics (CFD) is regarded as a powerful tool for analyzing the effectiveness of cooling within racks and aisles. CFD simulation with primarily the Reynolds Averaged Navier-Stokes (RANS) turbulence models such as k-epsilon model was used to study the layout design of data centers from individual components such as server racks [3–6] and perforated tiles [7,8] to large scale designs such as raised floor plenums [9] and whole server rooms [10–13]. It provides a detailed thermal field of the entire data center room, so that potential hot spots or unnecessary cold spots can be identified. Also, air velocity, thermal comfort, as well as indoor air quality can be assessed using CFD. Furthermore, CFD can serve as a prediction and optimization tool in the early designing stages of new-built data centers to avoid unwanted facility reconstruction due to ineffective utilization of energy. As a result, money can be saved, and more new data center projects, with the assurance and confidence of business owners, become possible.

* Corresponding author.
E-mail address: lincx@fiu.edu (C.-X. Lin).

http://dx.doi.org/10.1016/j.enbuild.2016.11.050
0378-7788/

Within the last decade, the load density in data centers have been growing rapidly due to increasing server demand for IT services. This creates problems for cooling servers in existing data centers. These data centers are designed and built for initial IT loads and do not expect the rise in future server load demand. Moreover, the issues of seasonal weather and climate change also cause an impact on thermal management in data centers as well. Finally, during a normal service day, a data center can undergo a surge in server demand due to various reasons, and the HVAC provision may not adequately fulfill the cooling requirement. As far as these existing issues are concerned, the need for a more responsive cooling scheme, such as a real-time system, is significant. In addition, smart HVAC control strategies are vital on the effort of using energy efficiently in data centers.

Although CFD is an excellent analyzing tool for data centers, there are limitations for real-time applications. A real-time system requires a level of responsiveness within a specified time constraint. At the time of a triggered event, the system has to receive, process, and return the results sufficiently quickly to reflect the environment at that time without causing any significant delay. This is almost unachievable with traditional CFD when the computational time and computing resource requirements are very large. Therefore, in order to overcome this issue, new methodologies are needed to improve the speed of large-scale data analysis.

In recent years, many studies have focused on reduced order modeling or simplified modeling to speed up the simulation process while trying to preserve the accuracy of the results. Some of these techniques include ad hoc methods [14], response surface methodology [15], and potential flow methods [16–18]. In general, these methods increase the computational efficiency relative to traditional CFD modeling methods, but on a tradeoff of modeling accuracy. The ad hoc method is simple but needs many empirical coefficients to be determined by experiments. The response surface method requires less data points but requires considerable efforts in model training or calibration. Potential flow method is relatively fast as compared to full Navier-Stokes equation based CFD but always has the concerns of over-simplified flow physics. In another effort, coupled inviscid-viscous CFD technique was also investigated [19]. The viscous region is simulated with RANS CFD models, while the inviscid region is simulated using the Euler equations. Therefore, the technique can reduce the computational time and resource compared to full RANS CFD simulation with turbulence models. However, the inviscid region does not necessarily capture all of the physics in a complex, turbulent flow field in the data center, when the Eulerian equation is used to describe the flow. In addition, the process for dividing the two regions depends on case by case and is actually complicated and time-consuming.

Although there are many appropriate methods that can be used for the thermal modeling of data centers as previously discussed, their weaknesses lie on the fact that they can only achieve either accuracy or efficiency. In this regards, Proper Orthogonal Decomposition (POD), widely used in many fields [21–27], is considered one of the most desired candidates for data center modeling because of not only its quick running time, but also its decent accuracy. POD allows to process large amounts of high-dimensional data with the aim of obtaining low-dimensional descriptions that capture much of the phenomena of interest [28]. Due to its rapid running time, POD can serve as an analyzing tool for real-time HVAC control of a whole data center with the use of sensor integration. In practice, three dimensional temperature and air flow data obtained from sensors within the conditioned space will be constantly fed into a so-called control box where it will be quickly analyzed by a POD solver. The control box will then send commands to the actuator to allow the HVAC system to be adjusted accordingly depending on the thermal load. POD method finds itself in many building applications including data centers. In recent articles, Li et al. [29,30] has applied POD method to optimize the ventilation system operation in office environment and control indoor thermal environment of buildings. Elhadidi and Khalifa [31] applied POD method to quickly and efficiently predict the indoor velocity and temperature distributions inside an office. Sempey et al. [32] used POD method to study the temperature distribution in air conditioned rooms. Samadiani and Joshi [33–35] extensively studied available reduced order modeling approaches including POD in the literature for predicting the flow and specially temperature fields inside a data center (L × W × H = 29.91′ × 33.35′ × 9.38′) with respect to two design parameters, i.e., inlet air velocity of computer room air conditioner (CRAC) unit (1.15 kg/s–31.17 kg/s) and rack heat load (1 kW–30 kW per rack).

In this paper, thermal modeling of data center is carried out using a POD method based on CFD data. Compared to existing works, the scope includes: (1) A model data center with different configurations is used for the computational study. This model data center (L × W × H = 45.9′ × 32.8′ × 11.5′) is a well-representative of a real data center, which has three hot aisles and two cold aisles, and four CRAC units placed at four opposite walls. It is different from the study of Samadiani and Joshi [33–35] where they have two CRAC units each placed at the end walls of the rack rows. (2) The effects of the number of parameters used for the POD basis function on the accuracy of the thermal modeling of data centers are investigated. The key differences between the current study and previous works also lie upon the number of input parameter used to construct the POD basis function. In the current paper, we apply the POD method to obtain the temperature profile of a new representative data center model using three parameters, i.e., the rack heat load, mass flow rate, and an additional parameter which is the inlet temperature. (3) An interpolation technique of constructing 3-D picture of the temperature profiles using 2-D POD simulation is introduced for data center modeling at different conditions. Most of the previous studies that use POD method on data centers focus on 2-D temperature and/or velocity distribution POD representation with the variation of server load and velocity at the inlet for their studied cases. The 2D-based modeling method for a 3D data center model construction is carried out to study the effects of multiple design parameters including inlet air mass flow rate (5.5 kg/s–7.3 kg/s), rack heat load (9 kW–18 kW per rack), and also a new parameter which is also important in data center thermal management, i.e., rack inlet air temperature (12.0 °C–13.2 °C). Case studies are generated from these parameters combinations. Sensitivity analysis of the effects of each design parameter on the temperature is done to assess the fidelity of the POD model as compared to the CFD model.

2. Methodologies

2.1. Fundamentals of proper orthogonal decomposition

The Proper Orthogonal Decomposition (POD), depending on different fields and applications, is also called the discrete Karhunen-Loève transform in signal processing, the Hotelling transform in multivariate quality control, singular value decomposition of $\mathbf{X}$, eigenvalue decomposition of $\mathbf{X}^T\mathbf{X}$ in linear algebra, just to name a few. This is a reduced order modeling technique that finds applications in computationally processing large amount of high-dimensional data with the aim of obtaining low-dimensioning descriptions that capture much of the phenomena of interest. It is a powerful tool that can be used to numerically predict the temperature and flow fields much faster as compared to conventional full-field CFD/HT modeling.

In the application of thermal/fluid system, the method of snapshots introduced by Sirovich [36] is usually used to reduce the calculation effort. The fundamental idea of using this method is to