Applied Energy 138 (2015) 432–444

Contents lists available at ScienceDirect

Applied Energy

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

Multi-objective efficiency enhancement using workload spreading in an operational data center

Ali Habibi Khalaj [a,*], Thomas Scherer [b], Jayantha Siriwardana [a], Saman K. Halgamuge [a]

[a] Department of Mechanical Engineering, The University of Melbourne, Parkville, Victoria, Australia
[b] Cloud & Computing Infrastructure Department, IBM Research – Zurich, Rüschlikon, Switzerland

HIGHLIGHTS

- Development of the heat-flow reduced order model (HFROM) for the IBM ZRL data center.
- Verification of the developed HFROM with the experimentally verified CFD model.
- Multi-objective efficiency enhancement of the HFROM using particle swarm optimization.
- Improving the COP of the data center's cooling system by about 17%.
- Increasing the total allocated workload of the servers by about 10%.

ARTICLE INFO

Article history:
Received 6 May 2014
Received in revised form 22 October 2014
Accepted 26 October 2014

Keywords:
Data center
Cooling system
Workload spreading
Heat flow reduced order model
Multi-objective particle swarm optimization
CFD

ABSTRACT

The cooling systems of rapidly growing Data Centers (DCs) consume a considerable amount of energy, which is one of the main concerns in designing and operating DCs. The main source of thermal inefficiency in a typical air-cooled DC is hot air recirculation from outlets of servers into their inlets, causing hot spots and leading to performance reduction of the cooling system. In this study, a thermally aware workload spreading method is proposed for reducing the hot spots while the total allocated server workload is increased. The core of this methodology lies in developing an appropriate thermal DC model for the optimization process. Given the fact that utilizing a high-fidelity thermal model of a DC is highly time consuming in the optimization process, a three dimensional reduced order model of a real DC is developed in this study. This model, whose boundary conditions are determined based on measurement data of an operational DC, is developed based on the potential flow theory updated with the Rankine vortex to account for buoyancy and air recirculation effects inside the DC. Before evaluating the proposed method, this model is verified with a computational fluid dynamic (CFD) model simulated with the same boundary conditions. The efficient load spreading method is achieved by applying a multi-objective particle swarm optimization (MOPSO) algorithm whose objectives are to minimize the hot spot occurrences and to maximize the total workload allocated to servers. In this case study, by applying the proposed method, the Coefficient of Performance (COP) of the cooling system is increased by 17%, and the total allocated workload is increased by 10%. These results demonstrate the effectiveness of the proposed method for energy efficiency enhancement of DCs.

Crown Copyright © 2014 Published by Elsevier Ltd. All rights reserved.

1. Introduction

Data Centers (DCs) are computing infrastructure facilities housing a considerable amount of Information Technology (IT) equipment required for processing, transmitting and storing information. Uninterrupted operation of DCs is one of the most crucial requirements, in which reliable power supply and cooling are two vital factors. Uninterrupted power supply is assured by installing several power backup sources. Provision of reliable cooling, on the other hand, can be a more complex issue, since the temperature and humidity of the DC must be constantly maintained in an appropriate condition [1–5].

Currently, DCs consume 1–2% of the world's electricity production, and due to their growing energy demand this number is increasing at a rate of 12% annually [6]. A significant portion of this power consumption, amounting to approximately 40% for a typical DC, can be ascribed to its cooling system [6,7]. Therefore, DC energy consumption is considered as a growing concern for

* Corresponding author. Tel.: +61 416106135.
E-mail addresses: alih@student.unimelb.edu.au, a.habibikhalaj@gmail.com (A. Habibi Khalaj).

http://dx.doi.org/10.1016/j.apenergy.2014.10.083
0306-2619/Crown Copyright © 2014 Published by Elsevier Ltd. All rights reserved.

business profitability and consequently thermal management of DCs has become one of the main foci in DC optimization.

In a typical raised floor DC, all servers are housed in racks located on a raised floor plenum. These racks are organized into rows that are separated by aisles. These aisles alternate between cold air intake and hot air exhaust. Supplied cold air from computer room air conditioner (CRAC) units is distributed into the DC room through perforated floor tiles in order to cool down the servers. By passing through the servers, the cold air absorbs heat dissipated by the servers and exhausts it into the hot aisles. Finally, the hot air is returned to the CRAC units through ceiling intakes, completing the cycle. This process is illustrated in Fig. 1.

Inefficiency of the cooling system in air-cooled DCs has several sources. The main sources are heat recirculation and bypass air caused by intermixing of hot and cold air streams and escape of refrigerated air from the cold aisle, respectively. As seen in Fig. 1, the heat recirculation creates localized increased temperature hot spots around racks. Therefore, the CRAC units have to over-cool the room to supply sufficient cold air for cooling these localized hot spots. This affects the efficiency of the CRAC units not only through higher utilization, but also through reducing its Coefficient of Performance (COP) [2–5,8].

The knowledge of heat recirculation, CRAC efficiency and the magnitude of over-cooling can be utilized to predict the potential energy savings for upcoming workload. By placing the workload intelligently with respect to its thermal impact, CRAC units can be operated more efficiently by increasing the supply temperature set point. Given the fact that a high fidelity model of a DC requires considerable simulation time, we require an accurate and fast heat flow model of the DC that is capable of predicting the temperature profile efficiently. This highlights the potential of having a heat flow reduced order model (HFROM) of the DC in order to investigate the impact of workload variation on the temperature profile around racks. By coupling this HFROM with a fast and robust optimization algorithm, workload can be intelligently redistributed among servers with the aim of hot spot reduction.

In this article, related work is described in Section 2. Section 3 presents the proposed strategy for minimizing the hot spots when the total server workload is maximized. Section 4 provides the results of our case study, where we applied our optimization strategy to a production DC. Finally, the article concludes with a brief summary in Section 5.

2. Related work

Recently, considerable attention has been paid to increasing the performance of cooling systems in raised floor DCs from chip to room levels. In this section, we provide a brief overview on the relevant methodologies proposed for thermally aware workload spreading for servers. The core of these methodologies lies in the method utilized for predicting the air heat flow behavior inside the DC when design parameters are changed.

2.1. Abstract heat flow model

Tang et al. [9], proposed an abstract heat flow model using a cross-interference coefficient matrix in order to describe the heat transfer ratio of recirculated air around racks. This matrix is obtained through a calibration process for a specific DC and is capable of showing how much the exhaust heat from the outlet of each server contributes to the inlet temperature of other servers. This method is proposed for temperature prediction at inlets and outlets of servers while their power consumption is changing. Since this model assumes a linear behavior for the temperature profile inside the DC, it can only capture the temperature trends at the rack inlets. The assumption is only valid if buoyancy and recirculation effects are negligible. Therefore, the accuracy of this model reduces when the size of the DC increases. Utilizing this method, researchers investigated the effectiveness of various thermal aware workload spreading scenarios [10–15].

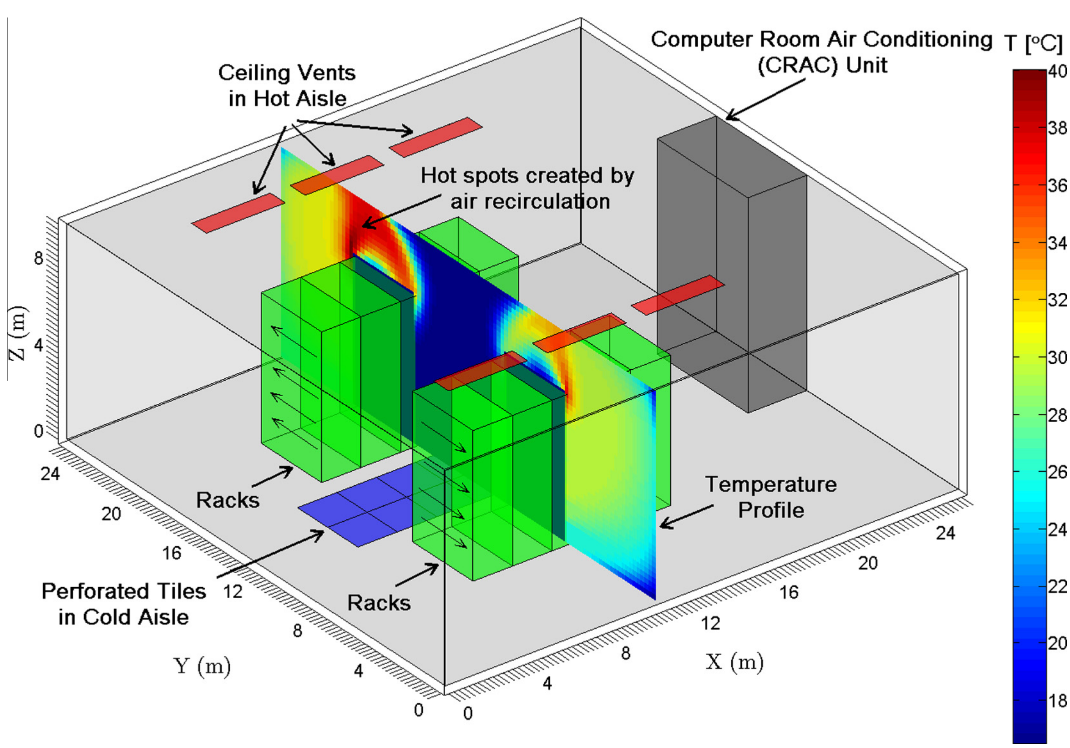

Fig. 1. General layout of a typical raised floor DC.

Download English Version:

https://daneshyari.com/en/article/6688319

Download Persian Version:

https://daneshyari.com/article/6688319

Daneshyari.com