Computers & Operations Research 75 (2016) 103–117

Contents lists available at ScienceDirect

Computers & Operations Research

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

A multi-objective co-evolutionary algorithm for energy-efficient scheduling on a green data center

Hongtao Lei [a,*], Rui Wang [a], Tao Zhang [a,b], Yajie Liu [a], Yabing Zha [a,b]

[a] College of Information Systems and Management, National University of Defense Technology, Changsha, Hunan 410073, PR China
[b] State Key Laboratory of High Performance Computing, National University of Defense Technology, Changsha, Hunan 410073, PR China

ARTICLE INFO

Article history:
Received 27 December 2015
Received in revised form
25 April 2016
Accepted 22 May 2016
Available online 24 May 2016

Keywords:
Scheduling
Energy-efficient
Green data center
Multi-objective optimization

ABSTRACT

Nowadays, the environment protection and the energy crisis prompt more computing centers and data centers to use the green renewable energy in their power supply. To improve the efficiency of the renewable energy utilization and the task implementation, the computational tasks of data center should match the renewable energy supply. This paper considers a multi-objective energy-efficient task scheduling problem on a green data center partially powered by the renewable energy, where the computing nodes of the data center are DVFS-enabled. An enhanced multi-objective co-evolutionary algorithm, called OL-PICEA-g, is proposed for solving the problem, where the PICEA-g algorithm with the generalized opposition based learning is applied to search the suitable computing node, supply voltage and clock frequency for the task computation, and the smart time scheduling strategy is employed to determine the start and finish time of the task on the chosen node. In the experiments, the proposed OL-PICEA-g algorithm is compared with the PICEA-g algorithm, the smart time scheduling strategy is compared with two other scheduling strategies, i.e., Green-Oriented Scheduling Strategy and Time-Oriented Scheduling Strategy, different parameters are also tested on the randomly generated instances. Experimental results confirm the superiority and effectiveness of the proposed algorithm.

1. Introduction

Modern computing centers and data centers consume an enormous amount of energy, and the consumed energy mainly comes from the conventional energy which is generated using fossil fuels. This results in high electricity bills, and more importantly it leads to high carbon emissions and huge amount of pollution. For example, the Tianhe-2 system in Guangzhou, the fastest supercomputing system in world, has a power consumption of 17,808 kW [43]. It costs more than 23 million USD per year assuming the electricity charge is 0.15 USD per kW h [27]. Also, the four data centers of Facebook consumed about 678 million kW h of energy in 2012, which costs about more than 25 million USD per data center [40].

In order to mitigate the negative environmental implication and the energy crisis caused by the rapidly increased energy consumption, many companies, e.g., Google, Microsoft, Yahoo! and Apple, have started to power their data centers with the renewable energy, reducing the dependence of power supply by conventional energies. And the solar and wind generations are two of the most promising renewable energy technologies because of their clean and wide uses. This motivates some data center operators to either generate their own solar (or wind) energy or draw power directly from a nearby solar (or wind) farm. Many data centers which are partially or fully powered by solar and (or) wind energy are being built all over the world [8,4]. This trend will continue, as these technologies' capital costs keep going down and governments continue to provide generous incentives for green power generation.

However, unlike the conventional energy drawn from the gird, the renewable energies (e.g., wind or solar energy) have the intermittent nature. Their outputs are often effected by the period of a day, the weather, the season and so on. For the efficient utilization of the renewable energy, it is an effective method to predict the output of renewable energy before using it, and balance the energy supply and the workload in the data center based on predictions.

In this paper, we investigate how to schedule a data center's workload (computational tasks), and select the computing nodes and their supply voltages and clock frequencies to match the renewable energy supply based on the prediction of renewable energy. Our focus is that we use the prediction information of the renewable energy supply as input to schedule the tasks of data center, and choose suitable computing nodes and their voltages

* Corresponding author.
E-mail addresses: leihongtao@nudt.edu.cn (H. Lei), wangrui@nudt.edu.cn (R. Wang), zhangtao@nudt.edu.cn (T. Zhang), liuyajie@nudt.edu.cn (Y. Liu), zhayabing@nudt.edu.cn (Y. Zha).

http://dx.doi.org/10.1016/j.cor.2016.05.014
0305-0548/

and frequencies adapting the renewable energy supply. Our objectives are to maximize the utilization of renewable energy and the satisfaction rate of tasks, and simultaneously minimize the makespan of tasks and the least total energy consumption. Main contributions of the paper are as follows:

- To the best of our knowledge, this paper is first to consider the multi-objective energy-efficient scheduling for the tasks on a green data center partially powered by the renewable energy, where the computing nodes of the data center is Dynamic Voltage Frequency Scaling enabled (DVFS-enabled).
- Three mathematical models are defined for the proposed problem including a task model, an energy model for the consumption of renewable energy and conventional energy, and a multi-objective scheduling model with the aim of maximizing the utilization of renewable energy and the satisfaction rate of tasks and minimizing the makespan of tasks and the total energy consumption.
- Based on the models, this paper proposes an enhanced co-evolutionary algorithm with the generalized opposition based learning strategy and the smart time scheduling strategy for the multi-objective energy-efficient scheduling of tasks on the green data center.
- The superiority of the enhanced co-evolutionary algorithm is demonstrated in comparison with the state-of-the-art MOEA, i.e., PICEA-g, and two other time scheduling strategies, i.e., Green-Oriented Scheduling Strategy (GOSS) and Time-Oriented Scheduling Strategy (TOSS), on a set of test instances.

The remainder of the paper is structured as follows. Section 2 reviews related works. Section 3 describes the proposed problem. Section 4 presents the mathematical models—the task model, the energy model and the multi-objective scheduling model. In Section 5, an enhanced PICEA-g algorithm for the problem solving is provided. Results of evaluation experiments are discussed in Section 6. Section 7 concludes this paper.

2. Literature review

A rich literature has been investigated on the task scheduling of distributed systems over the last decade. Mishra et al. [31] proposed static power management scheme which allocates global static slack according to the degree of parallelism in a given static schedule generated from any list scheduling heuristic algorithm. Zhu et al. [61] proposed two novel power-aware scheduling algorithms for task sets with and without precedence constraints executing on multiprocessor systems, based on the concept of slack sharing among processors. Aydin et al. [1] addressed the power-aware scheduling of periodic tasks to reduce CPU energy consumption in hard real-time systems through dynamic voltage scaling, and proposed a static solution to compute the optimal speed at the task level on the base of the worst-case workload for each arrival. Ge et al. [13] provided several distributed performance-directed DVFS scheduling strategies for use in scalable power-aware HPC clusters so as to obtain significant energy savings without increasing execution time. Tavares et al. [41] studied a pre-runtime scheduling method which considered the DVFS technique for reducing energy consumption and took the inter-task relations and runtime overhead into account. Also a formal model based on time Petri net was employed as a mathematical basis for precise pre-runtime schedule generation. Von Laszewski et al. [44] focused on scheduling virtual machines in a compute cluster to reduce power consumption by the DVFS technique, and presented an efficient scheduling algorithm to allocate virtual machines in a DVFS-enabled cluster by dynamically scaling the supplied voltages. Liu et al. [29] developed an energy-performance balanced task duplication based clustering scheduling algorithm in homogenous clusters, which can significantly save energy by judiciously shrinking communication energy consumption when assigning parallel tasks to computing nodes. Zong et al. [66] proposed two energy-aware duplication-based scheduling algorithms, namely, energy-aware duplication algorithm, EAD, and performance-energy balanced duplication algorithm, PEBD, to achieve the goal of optimizing both performance and energy efficiency in clusters. He et al. [19] developed a rolling-horizon scheduling strategy for the energy constrained distributed real-time embedded systems. Zhu et al. [62] presented a fault-tolerant scheduling algorithm called QAFT for real-time tasks with QoS needs on heterogeneous clusters. Zhu et al. [63] proposed an adaptive energy-efficient scheduling, AEES, for aperiodic and independent real-time tasks on heterogeneous clusters with dynamic voltage scaling. Later, Zhu et al. [64] developed an energy-efficient elastic scheduling for aperiodic, independent and non-real-time tasks with user expected finish times on DVFS-enabled heterogeneous computing systems. Wang et al. [46] concentrated on minimizing energy for precedence-constrained parallel task execution in a cluster, proposed two scheduling algorithms in DVFS-enabled clusters for executing parallel tasks: the PATC and PALS, and developed a green SLA-based mechanism to reduce energy consumption by returning users' increased tolerant scheduling makespan. Zhang and Guo [60] addressed the issue of minimizing overall energy consumption of a real-time sporadic task system, considering a generalized power model. A static sporadic tasks scheduling algorithm, SSTLPSA, and a dynamic sporadic tasks scheduling algorithm, DSTLPSA, have been proposed. Guzek et al. [18] investigated and solved the multi-objective precedence constrained application scheduling problem on a distributed computing system. Multi-objective evolutionary algorithms are developed for the problem solving. Moschakis and Karatza [32] developed the simulated annealing and thermodynamic simulated annealing in the multi-criteria scheduling of a dynamic multi-cloud system with virtual machines of heterogeneous performance serving Bag-of-Tasks applications. Iturriaga et al. [21] and Banković and López-García [5] proposed multi-objective energy efficient scheduling algorithms for the task scheduling, respectively in heterogeneous computing systems and in multicore environments. Different from our work, they do not consider the utilization of the renewable energy for the scheduling, and our paper emphasizes to use the renewable energy for the multi-objective energy-efficient scheduling on the green data center where the computing nodes is dynamic voltage frequency scaling enabled.

The researches on exploiting renewable energy in data centers have become recently active. Stewart and Shen [39] studied how to maximize the green energy used in data centers with the request distribution in multi-datacenter interactive services. Zhang et al. [59] provided the GreenWare, a novel middleware system that conducts dynamic request dispatching to maximize the percentage of renewable energy used to power a network of distributed data centers, subject to the desired cost budget of the Internet service operators. Goiri et al. [15] proposed the GreenSlot, a parallel job scheduler for data centers partially powered by solar energy. The GreenSlot predicts solar energy availability two days into the future, and schedules jobs to maximize the use of green energy and limits brown energy costs using the predictions and avoiding deadline violations. Krioukov et al. [23,24] suggested green energy-aware job schedulers for a single data center, and explored the scheduling of workload to match the available renewable energy supply so as to make more intermittent renewable energy available for use. Liu et al. [30] investigated whether a balance of geographical load can encourage the use of green renewable energy and reduce the use of brown fossil fuel energy. In