Future Generation Computer Systems 78 (2018) 89–101

Contents lists available at ScienceDirect

Future Generation Computer Systems

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

Greening cloud data centers in an economical way by energy trading with power grid

Chonglin Gu, Longxiang Fan, Wenbin Wu, Hejiao Huang *, Xiaohua Jia

School of Computer Science and Technology, Harbin Institute of Technology Shenzhen Graduate School, Shenzhen 518055, Guangdong, China

HIGHLIGHTS

- A novel scheduling architecture has been designed to green cloud data centers with lower cost through using ESDs and energy trading.
- Our green scheduler considered the requests with different resources demand such as CPU, memory, disk, bandwidth and execution time.
- The data centers can be powered by the self-generated wind or solar energy, or that purchased from the renewable power plant.
- Experiment result shows that the total energy cost and carbon emissions can further be reduced by energy trading with the power grid.

ARTICLE INFO

Article history:
Received 8 April 2016
Received in revised form
28 September 2016
Accepted 21 December 2016
Available online 28 December 2016

Keywords:
Green
Cloud
Data Center
Renewable Energy
Energy Trading

ABSTRACT

In this paper, we mainly study the issues of green scheduling for cloud data centers in an economical way by energy trading with the power grid. The data centers can be powered by the self-generated wind or solar energy, or by the energy purchased from the renewable power plant. The renewable energy can be used to power the data centers directly, or stored into ESDs (Energy Storage Devices) for later use, or sell back to the power grid to finance part of the high energy expenditure of the data centers. It is hard to make decisions on the usage of each type of energy considering dynamic resource demand of different types of requests, time-varying and location-varying electricity prices, and intermittent supply of renewable energy in each time slot. We focus on two optimization problems: (1) Minimizing total energy cost through scheduling of requests, servers, and the usage of different energy sources. (2) Minimizing total carbon emissions within the budget of energy cost. We formulate each problem as an optimization problem during the whole period of time. Our simulation is based on traces from real world. Experiment results show that our scheduling methods can significantly reduce the carbon emissions for cloud data centers.

1. Introduction

Cloud data centers consume a large amount of energy, which aroused people's concern on both energy cost and environmental implications such as global warming. It has been reported that Google can consume over 1120 GWh with 67 million dollars per year, and Microsoft consumes 600 GWh with 36 million dollars annually [1]. In the United States, generating 100 MWh energy will emit 50 tons of carbon dioxide [2]. In view of this, huge amount of efforts have been devoted to greening cloud data centers. Some are trying to enhance energy efficiency by consolidating VMs onto fewer servers with idle servers shut down, or DVFS (dynamic voltage and frequency scaling) techniques; while others are trying to power data centers using renewable energy. For example, Google has purchased 20 years' wind energy from Iowa wind farm to power its data centers in Oklahoma [3]. Facebook and Green House are trying to power their data centers using the self-generated solar energy or wind energy [4]. However, it is the high generating cost and infrastructure investment that hinder the extensive usage of renewable energy for cloud data centers.

In this paper, we study the issues of green scheduling for cloud data centers in an economical way by energy trading with the power grid. Each data center has its own wind turbines and solar panels to generate renewable energy, which can be used to power the data centers directly, or stored into ESDs (Energy Storage Devices) for latter use, or sold back to the power grid to

* Corresponding author.
E-mail addresses: guchonglin-6@163.com (C. Gu), lxfan@gmail.com, 784625011@qq.com (L. Fan), wbwu15@hitsz.edu.cn, 438314711@qq.com (W. Wu), hjhuang@hitsz.edu.cn (H. Huang), csjia@cityu.edu.hk (X. Jia).

http://dx.doi.org/10.1016/j.future.2016.12.029

finance part of the high energy expenditure of the data centers. An alternative option for greening the data centers is to buy wind or solar energy from the renewable power plant when the self-generated energy is insufficient. To make the best use of renewable energy, we use ESDs to store renewable energy when its supply is abundant, or when the outside electricity price is low. The energy stored in ESDs can be discharged to power data centers when outside electricity price is high, or sold back to the power grid to lower down the overall energy cost. We focus on two optimization problems: (1) Minimizing total energy cost through scheduling of users' requests, servers, and the usage of different energy sources. (2) Minimizing total carbon emissions within the budget of energy cost. However, it is hard to make decisions on the usage of each type of energy considering the dynamic incoming requests of users, time-varying and location-varying electricity prices, and fluctuating energy supply in each time slot. Besides, the requests of different types may have different resources demand such as CPU, memory, disk, and bandwidth, so we need to take all those factors into account when scheduling the requests to the data centers. We formulate each problem as a complex MILP (Mixed Integer Linear Programming) problem with millions of decision variables. It is almost infeasible to solve the problem with such large size using existing methods. We solve each problem in three steps: (1) rank the data centers according to different optimization goals. (2) allocate the incoming requests onto as few servers as possible for the data centers one by one through multi-dimension bin packing method. (3) use *Cplex* solver to schedule the usage of different types of energy. Experiments show that our scheduling can significantly reduce both energy cost and carbon emissions for cloud data centers.

The rest of this paper is organized as follows. Section 2 reviews related work. Section 3 introduces the architecture of our green scheduler. In Section 4, the formulation of the problems will be given in detail. In Section 5, our algorithms will be introduced. In Section 6, we setup experiment and make evaluations. Finally, Section 7 concludes this paper.

2. Related work

Greening cloud data centers is becoming increasingly important considering the economical and environmental pressures. In general, the greening techniques for cloud data centers can be classified into two categories: enhancing energy efficiency and utilizing renewable energy.

DVFS (dynamic voltage and frequency scaling) and DPM (dynamic power management) are the two most important approaches in reducing the energy consumption of servers. The latter is more energy-saving, since all components in the servers can be powered down rather than only powering down the CPU in DVFS. Elnozahy et al. in [5] try to reduce the total power consumption of data centers by combining DVFS and power on/off techniques. Johan et al. in [6] propose an energy priority scheduling policy on a variable-voltage platform to reduce the power consumption of processors. Gandhi et al. in [7] exploit the optimal power allocation of the server farms by CPU frequency scaling through DVFS. Lin et al. in [8] propose a DPM scheme called dynamic right-sizing method to decide the number of active servers to reduce the power consumption of data center as much as possible. Besides, dynamic switching of servers between sleeping state and active state can save a large amount of energy. Niyato et al. in [9] design a scheme to switch the operation mode (active or sleep) of the server to minimize the power consumption while meeting the performance requirements based on constrained Markov decision process. As for the consolidation of VMs, Buyya et al. in [10] propose to allocate the VMs with heterogeneous applications onto fewer servers using bin packing method. Srikantaiah et al. in [11] try to minimize the energy consumption of data centers by consolidating heterogeneous workload onto fewer servers. However, none of those work considered utilizing renewable energy to power the data centers to further reduce carbon emissions to the atmosphere.

For utilizing renewable energy, Liu et al. in [12] try to lower down the price of brown energy by using renewable resources in a specific market. Brown et al. in [13] propose a simulation infrastructure called Rerack, which can be used to evaluate the cost of data centers using renewable energy. Zhang et al. in [14] propose a scheduler called GreenWare trying to maximize the usage of renewable energy under a certain budget. Gu et al. in [15,16] exploit to minimize carbon emissions using wind and solar energy. Li et al. in [17] try to coordinate the workload power with the supply of the renewable energy to reduce the carbon footprint. Authors in [18,19] build a demo system with solar panels to power a rack of servers. Based on this system, they designed two schedulers (Green Hadoop and Green Slot) to maximize the usage of renewable energy within deadlines of the tasks. However, none of those work considered lowering down the high energy cost when greening cloud data centers.

In smart grid, there has been studies regarding selling the self-generated solar energy back to the power grid to make profit for the end-users [20,21]. Therefore, we consider lowering down the electricity cost by energy trading with the power grid by leveraging the varying electricity prices and energy storage in our paper. As far as we know, our work is the first to explore the problem of greening cloud data centers from economic perspective.

3. Architecture of green scheduler

Cloud service providers usually have multiple data centers in geographically distributed locations. We assume each data center is powered by three types of energy: brown energy from the power grid, wind energy, as well as solar energy. The renewable energy can be self-generated, or purchased from the renewable power plant. Considering the energy trading with the power grid, the architecture of our scheduling system is shown in Fig. 1.

In this system, we achieve green scheduling through dispatching the incoming requests to different data centers. Due to the intermittent nature of renewable energy, we use ESDs to store energy when its supply is abundant, or when the energy price is low. We focus on two optimization problems: (1) minimizing the total energy cost, and (2) minimizing total carbon emissions under a given budget of energy cost. The constraints include the computing resources of the servers in each data center, and the energy supply of different types.

4. Problem formulation

A. Workload model

The main task of our scheduler is to dispatch the incoming requests to the servers of each data center. Suppose there is a cloud service provider having N geographically distributed data centers, each having M_j servers, where $j \in \{1, 2, \ldots, N\}$. Our scheduling period is T time slots. For each time slot t, there are λ^t incoming requests. Since the resource demand of the requests are different from each other, we denote each request as $R_i^t = \{CPU, Mem, Disk, Band\}$, where $i \in \{1, 2, \ldots, \lambda^t\}$. We divide the dispatching process into two levels:

Level 1: Dispatch the requests from scheduler to the data centers. Let the boolean variable x_{ij}^t denote whether the requests i is dispatched into data center j. During transmission, the total