ARTICLE IN PRESS

Future Generation Computer Systems ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

Future Generation Computer Systems

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

Energy cost minimization with job security guarantee in Internet data center

Zhongjin Li [a,b], Jidong Ge [a,b,*], Chuanyi Li [a], Hongji Yang [c], Haiyang Hu [a,b,d], Bin Luo [a], Victor Chang [e]

[a] State Key Laboratory for Novel Software Technology, Software Institute, Nanjing University, Nanjing 210093, China

[b] State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing 100876, China

[c] Centre for Creative Computing (CCC), Bath Spa University, England, UK

[d] School of Computer, Hangzhou Dianzi University, Hangzhou 310018, China

[e] International Business School Suzhou, Xi'an Jiaotong Liverpool University, Suzhou, China

HIGHLIGHTS

- The energy cost optimization architecture is proposed for IDC operator.
- A heuristic algorithm is devised to select security services to guarantee the job security.
- The temporal diversity of electricity price is considered in minimizing the energy cost.
- The energy cost minimization algorithm is based on Lyapunov optimization technique.
- Extensive evaluation experiments demonstrate the effectiveness of our algorithms.

ARTICLE INFO

Article history:
Received 31 August 2016
Received in revised form
10 December 2016
Accepted 12 December 2016
Available online xxxx

Keywords:
Internet data center
Security service
Risk probability constraint
Energy cost minimization
Deregulated electricity markets

ABSTRACT

With the proliferation of various big data applications and resource demand from Internet data centers (IDCs), the energy cost has been skyrocketing, and it attracts a great deal of attention and brings many energy optimization management issues. However, the security problem for a wide range of applications, which has been overlooked, is another critical concern and even ranked as the greatest challenge in IDC. In this paper, we propose an energy cost minimization (ECM) algorithm with job security guarantee for IDC in deregulated electricity markets. Randomly arriving jobs are routed to a FIFO queue, and a heuristic algorithm is devised to select security levels for guaranteeing job risk probability constraint. Then, the energy optimization problem is formulated by taking the temporal diversity of electricity price into account. Finally, an online energy cost minimization algorithm is designed to solve the problem by Lyapunov optimization framework which offers provable energy cost optimization and delay guarantee. This algorithm can aggressively and adaptively seize the timing of low electricity price to process workloads and defer delay-tolerant workloads execution when the price is high. Based on the real-life electricity price, simulation results prove the feasibility and effectiveness of proposed algorithm.

1. Introduction

Cloud computing, comprises of infrastructure platforms called Internet data center (IDC), is a large-scale distributed computing to meet the skyrocketing demand of big data applications and services. As an IDC typically consists of tens of thousands of servers, the energy consumption or energy cost is one of the critical problems. For example, many IDCs (e.g., Microsoft, Google, Akamai and INTEL) spend millions of dollars on electricity costs every year, which result in a large portion of operation expense [1,2]. Hence, a considerable cost can be saved even reducing a few percent energy cost.

In cloud computing environment, the jobs or application requests from the cloud users can be submitted to IDC, which are also considered as virtual network (VN) requests [3,4]. These jobs or applications may be delay tolerant big data, such as scientific computing and data intensive MapReduce

* Corresponding author at: State Key Laboratory for Novel Software Technology, Software Institute, Nanjing University, Nanjing 210093, China.
E-mail address: gjdnju@163.com (J. Ge).

http://dx.doi.org/10.1016/j.future.2016.12.017
0167-739X/© 2016 Elsevier B.V. All rights reserved.

applications [2]. Generally, jobs or applications are arrived at IDC randomly. Therefore, the scheduling problem, which is also can be considered as the energy cost problem, is key issue for IDC to ensure the QoS of jobs and reduce the energy overhead [5,6].

Recently, a great attention has been paid to IDC energy management by both academia and industry. Extensive research has been developed to minimize the energy cost by utilizing the electricity price dynamics across geographically distributed regions [7,8], and apply VM migration to achieve the goals of saving energy [9,10]. Especially, the electricity price manifests spatial and temporal diversity in the real life. For instance, in North America, owing to the different power generation profiles and electricity markets have been deregulated, the electricity prices are not constant but vary on the basis of an hour or 15-min [6]. To consider the temporal diversity of electricity price, the energy storage for energy cost saving is studied [11,12], and both service delay and energy cost are taken into account in geographically distributed data center [2].

Security is another critical concern and even ranked as the greatest challenge in cloud computing environment. A survey from international data corporation shows that security is one of the greatest concerns in cloud computing [13]. Many works tackle the security problem on clusters [14], grid computing [15], heterogeneous distributed system [16,17], cloud computing [18–21] and real-time embedded systems [22]. Unfortunately, because cloud computing environment is used to execute various applications of users, applications and users all may be the sources of malicious attack [23]. Furthermore, security mechanism is overlooked and has not been employed to counter any security threats [24,25]. Therefore, it is necessary to deploy security services to protect various applications running in the IDC. However, security workload is incurred by adding security services to applications. Hence, it is a big challenge to tradeoff energy cost and service quality.

In this paper, we propose an energy cost minimization (ECM) algorithm with job security guarantee for IDCs where the electricity price exhibits temporal diversity. These jobs may be delay tolerant big data applications that take from several minutes to more than many hours. Our targets can be described as follows: (1) guaranteeing the risk probability constraint of each arriving job; (2) exploiting the temporal diversity of electricity price to minimize energy cost. First, a heuristic algorithm is devised to select security levels for workload shaping to guarantee the job security. Then, the energy optimization problem is formulated by taking the temporal diversity of electricity price into account. An online ECM algorithm, based on the Lyapunov optimization framework, is applied to solve that optimization problem. Our purpose is to minimize energy cost by deciding: (1) how to select security services to guarantee the risk probability constraint; (2) how many workloads should be processed in each time slot; and (3) how many resources should be provided by the IDC.

The main contributions of this paper can be summarized as follows:

- We propose, design and evaluate an energy cost optimization architecture, which mainly consists of Cloud Users, Job FIFO Queue, IDC Operator (includes Job Analyzer, Workload Shaping with Security Guarantee, Energy Cost Minimization and Server Management) and Servers (see Section 3.1).
- We devise a heuristic algorithm for arriving jobs to select appropriate security services to guarantee the job security. Based on it, the security workload shaping can be finished. In our architecture, the total workload consists of task execution workload and security workload (see Sections 3.5 and 4.1).
- We exploit the temporal diversity of electricity price to minimize the energy cost in deregulated electricity markets. The ECM algorithm is based on Lyapunov optimization technique which can facilitate energy cost versus delay trade-off for IDC operator (see Sections 3.6 and 4.2).
- Based on real-life electricity price data sets, the simulation results show that our approach can achieve energy cost saving and security guaranteeing simultaneously (see Section 5).

The rest of this paper is organized as follows. Section 2 summarizes the related work. In Section 3, we describe the system architecture, models and problem formulation. Section 4 introduces the algorithm design. The performance evaluation approaches and results, comparisons with similar work, research contributions and limitations are conducted in Section 5. Conclusions and envisages our future work are given in Section 6.

2. Related work

Security is one of the critical problems in distributed computing environment. However, most existing well-known scheduling studies neglect the security problems, and only few groups of researchers consider the security-driven scheduling policy for applications. Azzedin and Maheswaran [26] presented a trust brokering system which implicated the security meaning and was applied to the public resource grids. Song et al. [15] proposed six risk-resilient scheduling strategies for job security-assured under different risky conditions in grid environment. Xie and Qin [14] built three security overhead models for measuring execution time incurred by the security-critical tasks in clusters. Also the performance evaluations of security heterogeneity scheduling algorithm were studied in distributed computing systems [16]. Tang et al. [17] used the differential equation to build system node trust model and proposed a security-driven scheduling architecture for directed acyclic graph (DAG) applications. As for the workflow applications in cloud, Zeng et al. [18] introduced a security-aware and budget-aware (SABA) scheduling strategy to minimize the makespan with budget constraint. Then, Li et al. [27] proposed a security and cost aware scheduling (SCAS) algorithm for workflow application to optimize the execution cost with deadline and risk probability guarantee in clouds. Due to financial sector confronts the problems of inaccurate and inadequate assessment, Chang [28] deployed complex models in cloud to improve accuracy on risk analysis and prediction. The balance between benefits and risks should be considered for the projects of organization. Hence, based on organizational sustainability modeling (OSM) [29], Chang et al. [30] proposed a new technique, capital asset price modeling (CAPM), to evaluate the risks and benefits of commercial projects.

Energy consumption or energy cost problem of cloud data center has been attracted many attentions [21,31–34]. Qureshi et al. [1] proved that electricity prices exhibit both temporal and spatial variations in deregulated electricity markets. According to the feature of electricity price, Rao et al. [7] proposed an energy cost minimization algorithm with guaranteeing quality of service under multiple electricity markets environment. Liu et al. [35] derived three distributed algorithms to achieve optimal geographical load balancing and also proved that geographical load balancing can significantly reduce brown energy use under special conditions. Shao et al. [6] used the mixed-integer nonlinear programming (MINLP) technique to achieve the optimal load balancing and energy cost management for IDCs. Luo et al. [36] proposed an energy cost optimization-IDC (eco-IDC) algorithm to minimize energy cost with service delay guarantee for data center.

In the light of risk preferences of IDC operators, Yu et al. [37] studied the problem of achieving the optimal tradeoff between