Future Generation Computer Systems 54 (2016) 95–122

Contents lists available at ScienceDirect

Future Generation Computer Systems

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

Virtual machine consolidated placement based on multi-objective biogeography-based optimization

Qinghua Zheng [a,b], Rui Li [a,b], Xiuqi Li [c], Nazaraf Shah [d], Jianke Zhang [a,e], Feng Tian [a,f,*], Kuo-Ming Chao [d], Jia Li [a,b]

[a] *MOE Key Lab for Intelligent Networks and Network Security, Xi'an Jiaotong University, Xi'an, China*

[b] *Department of Computer Science and Technology, Xi'an Jiaotong University, Xi'an, China*

[c] *Department of Computer Science and Mathematics, University of North Carolina at Pembroke, Pembroke, NC 28372, USA*

[d] *Faculty of Engineering and Computing, Coventry University, UK*

[e] *School of Science, Xi'an University of Posts and Telecommunications, Xi'an 710121, China*

[f] *Systems Engineering Institute, Xi'an Jiaotong University, Xi'an, China*

HIGHLIGHTS

- Clarify problems of incremental placement and consolidated placement of virtual machine.
- Build a optimization model of power consumption, resource wastage, server loads, inter-VM and storage network traffic.
- Firstly apply the BBO meta heuristic to virtual machine consolidated placement problem.
- Adopt a new strategy about migration rate generation, which beats original and other three strategies.
- Experimental results verified the robustness, adaptability and extensibility of the proposed method.

ARTICLE INFO

Article history:
Received 1 September 2014
Received in revised form
27 January 2015
Accepted 18 February 2015
Available online 30 March 2015

Keywords:
Virtual machine placement
Multi-objective optimization
Resource utilization
Biogeography-based optimization
Cloud computing

ABSTRACT

Virtual machine placement (VMP) is an important issue in selecting most suitable set of physical machines (PMs) for a set of virtual machines (VMs) in cloud computing environment. VMP problem consists of two sub problems: incremental placement (VMiP) problem and consolidated placement (VMcP) problem. The goal of VMcP is to consolidate the VMs to more suitable PMs. The challenge in VMcP problem is how to find optimal solution effectively and efficiently especially when VMcP is a kind of NP-hard problem. In this paper, we present a novel solution to the VMcP problem called VMPMBBO. The proposed VMPMBBO treats VMcP problem as a complex system and utilizes the biogeography-based optimization (BBO) technique to optimize the virtual machine placement that minimizes both the resource wastage and the power consumption at the same time. Extensive experiments have been conducted using synthetic data from related literature and data from two real datasets. First of all, the necessity of VMcP has been proved by experimental results obtained by applying VMPMBBO. Then, the proposed method is compared with two existing multi-objective VMcP optimization algorithms and it is shown that VMPMBBO has better convergence characteristics and is more computationally efficient as well as robust. And then, the issue of parameter setting of the proposed method has been discussed. Finally, adaptability and extensibility of VMPMBBO have also been proved through experimental results. To the best of our knowledge, this work is the first approach that applies biogeography-based optimization (BBO) to virtual machine placement.

1. Introduction

Cloud computing has been a popular computing paradigm in the IT industry since 2008. It delivers computing infrastructures, computing platforms, and software as hosted services on demand over the Internet. Cloud users can access computing resources

* Corresponding author at: Systems Engineering Institute, Xi'an Jiaotong University, Xi'an, China.
E-mail address: phd.fengtian@gmail.com (F. Tian).

http://dx.doi.org/10.1016/j.future.2015.02.010
0167-739X/

without having to own, manage, and maintain them. There are three common cloud computing models known as Infrastructure-as-a-Service (IaaS), Platform-as-a-Service (PaaS), and Software-as-a-Service (SaaS) [1–3]. In the most basic model IaaS, like Amazon Web Services, cloud users are provided with physical or, more frequently, virtual machines and additional resources like raw block storage, which are allocated from massive computing resource pools in large-scale data centers. The focus of this paper is on IaaS model.

From the perspective of a cloud provider, to reduce the operating cost, the use of computing resources in cloud needs to be maximized. In addition, the power consumption needs to be minimized as it has become a significant contributor to the operating cost [4]. The total electricity used by data centers in the US increased about 56% from 2005 to 2010 [5].

The core technology in cloud computing is virtualization [6], which separates resources and services from the underlying physical delivery environment. The resources of a single physical machine (PM) are sliced into multiple isolated execution environments for multiple virtual machines (VMs). Virtual Machine Placement (VMP) is an important topic in cloud environment virtualization, in particular in IaaS model. VMP maps a set of virtual machines to a set of physical machines. For the cloud providers, a good VMP solution should maximize resource utilization and minimize power consumption.

The problem of VMP consists of two sub problems: incremental placement (VMiP) problem and consolidated placement (VMcP) problem [7], as shown in Fig. 1. VMiP deals with continuous arrival of VM deployment and removal requests at runtime. Quick response is a crucial metric for ensuring a high service quality of the VMiP. The problem of VMiP has received much attention [7–10]. However, as VMs are being continuously removed over time, this may leads to a situation where infrastructure may be brought to a poor VMs distribution state over a long period of time. Therefore, there is an urgent need for periodically redeploying existing VMs on more suitable servers [7,11]. The terms "physical machine" and "server" will be used interchangeably in this paper.

VMcP is used by data centers to increase resource utilization and reduce electric power consumption costs [11]. VMcP is particularly important when users' workloads are unpredictable and VMcP need to be revisited periodically [11]. Whenever VM instances change, VMs can be relocated and migrated to different physical servers if necessary [11]. VMcP can be invoked periodically, or be triggered based on some preset conditions. The problem of VMcP is NP-hard [12–17]. The challenge in VMcP problem is how to find optimal solution effectively and efficiently.

The existing research in VMcP can be classified into five categories based on the techniques being used: heuristic bin packing [18–25,12,26], biology-based optimization [27–29,15,30,31], linear programming [32,33], constraint programming [34], and simulated annealing optimization [35]. Another type of classification is single-objective or multi-objective, based on the number of objectives to be optimized during the placement. Recent research [5,30,36] focuses on multi-objective solutions. Both [5,30] optimize resource utilization and power consumption. The thermal dissipation cost is also considered in some research efforts [30]. The work in [36] optimizes CPU utilization, network throughput, and disk I/O rate. Existing biology-based optimization algorithms for VMcP include genetic algorithms [29,5], particle swarm optimization [28] and ant colony optimization [27,30].

In this paper, we propose a novel multi-objective VMcP solution named VMPMBBO. It employs a state-of-the-art evolutionary algorithm, biogeography-based optimization (BBO) [37,38], to find the optimal VM placements that simultaneously minimizes both the resource wastage and the power consumption. Compared with two existing multi-objective evolutionary algorithms [5,30],

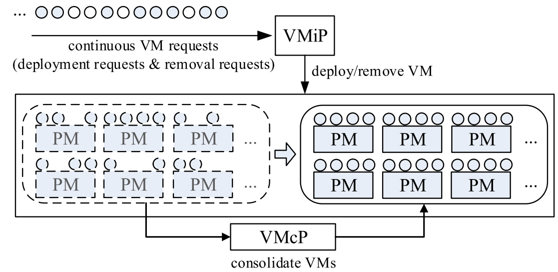

Fig. 1. An example of VMiP and VMcP.

VMPMBBO has better convergence characteristics and is more computationally efficient. Extensive simulation results confirm the effectiveness, efficiency and robustness of the proposed approach. Adaptability and extensibility of VMPMBBO have also been proved by experimental results. To the best of our knowledge, this work is the first approach that applies biogeography-based optimization (BBO) and complex system optimization to VMP problem.

The paper is organized as follows. In Section 2, the existing VMcP solutions are reviewed. VM placement problem is formulated in Section 3. Section 4 presents background knowledge of the proposed VMPMBBO approach. The simulation results are presented in Section 4. Section 5 evaluates the effectiveness of the proposed approach. Section 6 proves the adaptability and extensibility of VMPMBBO and finally Section 7 concludes the paper.

2. Related work

VMcP is one the well research area in cloud computing [4,18–20,27–29,15,30]. These research efforts can be classified into five categories based on their underlying techniques: heuristic bin packing [18–20,32,21–25,12,26], biology-based optimization [27–29,15,30,31], linear programming [32,33], constraint programming [34], and simulated annealing optimization [35].

Heuristic bin packing. Many studies formulated the VM placement as vector bin packing, a well-known NP-hard optimization problem [19]. Simple heuristics like greedy algorithms are utilized to approximate the optimal solution of this NP-hard problem. These include worst fit and best fit in [18], first fit decreasing (FFD) and best fit decreasing (BFD) [32,22]. Improvement attempts in further research efforts resulted in an extended first fit decreasing (FFD) heuristics in the pMapper system [20,38], first fit and best fit modified using node utility and power consumption [21], worst fit based on the profiling data [23], first fit modified using application migration [24], best fit used for VM migration in [12], and a modified best fit decreasing heuristics in [26].

Biology-based optimization. In [27], ant colony optimization method (ACO) is used to pack the VMs into the least number of physical machines necessary for the current workload. The SAPSO approach [28,31] is a self-adaptive particle swarm optimization (PSO) algorithm. SAPSO has been applied to automatically adjust VM placement in response to changing resource pools in a dynamic cloud environment. The GABA approach [29] is a genetic algorithm (GA) based algorithm that dynamically reconfigures the VM mappings according to estimated future workload in a dynamic cloud environment.

Linear programming. In [32], server consolidation problem is considered as a bin packing, for which FFD and BFD are suggested. In addition, the server consolidation problem is formulated as a linear programming that is extended with a number of constraint types elicited from practical applications. An LP-relaxation-based heuristic is designed to minimize the cost of linear programming. The study in [33] formulates the QoS-aware VMcP problem as