Egyptian Informatics Journal (2014) **15**, 1–12

Cairo University

Egyptian Informatics Journal

www.elsevier.com/locate/eij
www.sciencedirect.com

ORIGINAL ARTICLE

Optimum Resource Allocation of Database in Cloud Computing

Fatma A. Omara *, Sherif M. Khattab, Radhya Sahal

Faculty of Computers and Information, Cairo University, Egypt

Received 19 January 2013; revised 7 January 2014; accepted 14 January 2014
Available online 14 February 2014

KEYWORDS

Virtualization;
Resource allocation;
PSO;
Query optimizer;
Calibration

Abstract Cloud computing is a new generation of computing based on virtualization technology. An important application on the cloud is the Database Management Systems (DBMSs). The work in this paper concerns about the Virtual Design Advisor (VDA). The VDA is considered a solution for the problem of optimizing the performance of DBMS instances running on virtual machines that share a common physical machine pool. It needs to calibrate the tuning parameters of the DBMS's query optimizer in order to operate in a what-if mode to accurately and quickly estimate the cost of database workloads running in virtual machines with varying resource allocation.

The calibration process in the VDA had been done manually. This manual calibration process is considered a complex, time-consuming task because each time a DBMS has to run on a different server infrastructure or to replace with another on the same server, the calibration process potentially has to be repeated. According to the work in this paper, an Automatic Calibration Tool (ACT) has been introduced to automate the calibration process.

Also, a Greedy Particle Swarm Optimization (GPSO) search algorithm has been proposed and implemented in the VDA instead of the existed greedy algorithm to prevent the local optimum states from trapping the search process from reaching global optima. The main function of this algorithm is to minimize the estimated cost and enhance the VMs configurations.

The ACT tool and the GPSO search algorithm have been implemented and evaluated using TPC-H benchmark queries against PostgreSQL instances hosted in Virtual Machines (VMs) on the Xen virtualization environment.

* Corresponding author. Tel.: +20 22734030.
E-mail addresses: f.omara@fci-cu.edu.eg (F.A. Omara), s.khattab@fci-cu.edu.eg (S.M. Khattab), radhya.sahal@grad.fci-cu.edu.eg (R. Sahal).

Peer review under responsibility of Faculty of Computers and Information, Cairo University.

Production and hosting by Elsevier

1. Introduction

Cloud computing is a new generation of computing. It allows users to use computational resources and services of data centers (i.e., machines, network, storage, operating systems, application development environments, application programs) over the network to deploy and develop their applications [1]. The main feature of cloud computing is providing self-service pro-

1110-8665 © 2014 Production and hosting by Elsevier B.V. on behalf of Faculty of Computers and Information, Cairo University.
http://dx.doi.org/10.1016/j.eij.2014.01.002

visioning, which allows the users to deploy their own sets of computing resources [2]. The cloud computing technology is based on virtualization. Virtualization is a technology that separates computation functions from physical hardware. It allows the users to partition and multiplex physical machine infrastructure (e.g., CPU, memory, I/O, storage, and network interface cards) [3]. The applications are running on virtual machines instead of physical ones. The **V**irtual **M**achine (VM) is a software implementation of a computing environment to simulate a physical machine directly executing on physical hardware [4]. The **V**irtual **M**achine **M**onitor (VMM) is used to create and manage the VMs (e.g., Xen, VMware, VirtualBox, and KVM) [5]. The virtual machine configuration or resource allocation controls the sharing of physical resources (CPU, memory, I/O bandwidth) allocated to VMs. The problem of optimizing the performance of the virtualized applications (i.e., the applications that run on VMs) is critical to the success of the cloud computing paradigm, because VM configuration affects the application performance [2,6].

On the other hand, The Database Management System (DBMS) is considered one of the applications deployed on the cloud. Each DBMS instance has its own tuning parameters. The tuning parameters interact with cost model in DBMS' query optimizer to change the performance (e.g., CPU parameters and buffer parameters) [7]. DBMS needs to calibrate its tuning parameters in order to be aware of virtualized environment and produce an accurate estimated cost. Indeed, DBMS faces a challenge of tuning resource allocation because each workload (a set of SQL statements) has different characteristics and needs different resource allocation. In other words, how DBMS instances can get a benefit of resource allocation for each VM in the shared physical pool, this called Virtualization Design Problem (VDP) [7–9]. Virtual Design Advisor (VDA) is a technique that offers a solution for such problem. It gives recommended configurations for multiple VMs running different workloads among shared resources [2,7–9]. It explores the characteristics of workloads to distinguish their intensity (e.g., CPU or I/O intensive, etc.) and makes a decision for best resource allocation for VM which run this workload. The DBMS has a query optimizer tool to choose the best execution plan based on the estimated cost. The cost model is a module in the query optimizer tool which is responsible for the cost estimation. Database cost model expresses the total resources consumption for a given workload. It depends on static assumptions for tuning parameters to generate the execution plan. In fact, the accuracy of the execution of the current resources consumption is considered a problem for database's cost model.

In other words, the query optimizer's cost model is not aware of virtualized environment because it takes the default values of tuning parameters. So, the query optimizer parameters are needed to be calibrated in order to be aware of different resource allocation in virtualized environment. Each time, the DBMS instance moves from one infrastructure to another, or the DBMS instance is replaced by another DBMS instance in the same infrastructure, the calibration process is repeated. Unfortunately, this process had been executed manually. So, the calibration process is needed to be automated in order to save time, money and produce an accurate estimated cost. In this paper, an **A**utomatic **C**alibration **T**ool (ACT) has been introduced to tune parameters of DBMS query optimizer in virtualized environment to solve the manual calibration problem in the VDA.

On the other hand, a **P**article **S**warm **O**ptimization (PSO) is considered a modern evolutionary algorithm which is used to explore the search space of a given problem [10]. It is used to find optimal or near-optimal solutions for maximization/minimization search problems. As stated previously, the VDP is considered a search problem which tries to minimize the allocation cost of virtualized resources for database systems in cloud environment [2,7–9]. In this paper, a search algorithm called **G**reedy **P**article **S**warm **O**ptimization (GPSO) has been proposed to overcome the local optimum problem of the existed greedy algorithm in the VDA. The proposed GPSO algorithm is considered an amalgamation of heuristic greedy search and particle swarm optimization to optimize configurations based on the workload profile in virtualized environments. The GPSO algorithm has been implemented in the VDA enumerator module, which initially makes an equal resource allocation of VMs and adapts these allocations based on the estimated cost obtained by cost models of the database system query optimizer.

To evaluate the ACT tool and the GPSO search algorithm, prototype experiments have been conducted based on the optimal CPU allocation for the different virtual machines. Tests have been performed using PostgreSQL 8.4.8, running TPC-H benchmark queries as workloads [11,12]. The experimental results show that the ACT runtime increases linearly with the number of calibration sampling points, and the GPSO algorithm can provide effective configurations for different types of workloads than that the existed greedy algorithm.

The rest of this paper is organized as follow; the related works are described in Section 2. The calibration problem in the VDA is described in Section 3. The proposed automatic calibration tool for DBMS query optimizer is discussed in Section 4. In Section 5, the optimization problem in the VDA will be handled. In Section 6, the proposed GPSO algorithm will be discussed. In Section 7, the ACT and the GPSO algorithm evaluation results are introduced. In Section 8, the paper is concluded; also a brief outlook into the future work is given.

2. Related work

There are many research papers in the field of performance optimization of applications running in virtualized environments [8,9,13], and resource allocation [14,15]. A related problem to the work of this paper is the virtualization design problem which addresses the question of how to optimally (with respect to application throughput) partition the resources of a physical machine over a number of VMs, each running a potentially different database appliance (i.e., preconfigured DBMS and a set of workload queries) [7–9]. In [8,9], the virtual design advisor has been presented to solve the virtualization design problem by using the query optimizer, which is typically built-in in most DBMSs, as a cost model to evaluate potential resource partitioning configurations. This "*what-if*" usage of the query optimizer has also been used in non-virtualized environments to justify upgrades of resources based on the predictions of the expected improvement in workload performance [16,17]. In [2], the virtual design advisor has been used to optimize the performance of database appliances that had been deployed in the Amazon EC2 cloud. Ideally, the