Agent based Resource Allocation Mechanism Focusing Cost Optimization in Cloud Computing

Aarti Singh, Maharishi Markendeshwar Institute of Computer Technology and Business Management, Maharishi Markandeshwar University, Haryana, India

Manisha Malhotra, Maharishi Markendeshwar Institute of Computer Technology and Business Management, Maharishi Markandeshwar University, Haryana, India

ABSTRACT

A cloud computing environment offers a simplified, centralized platform or resources for use when needed on low cost. One of the key functionality of this type of computing is to allocate the resources on an individual demand. However, with the expanding requirements of cloud user, the need of efficient resource computing is also emerging. The main role of service provider is to effectively distribute and share the resources which otherwise would result into resource wastage. In addition to the user getting the appropriate service according to request, the cost of respective resource is also optimized. In order to surmount the mentioned shortcomings, this paper proposes a new agent based optimized resource assignment algorithm which is not only responsible for searching comprehensive services but also considers reducing the cost of virtual machines which are consumed by on-demand services only.

Keywords: Cloud Computing, Cloud Mobile Agent, Cost Optimization, Resource Allocation, Virtual Machine

1. INTRODUCTION

Cloud computing is a business model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction. The cloud computing platform guarantees subscriber the functionalities like on demand service, resource pooling, broad network access, rapid elasticity. In Cloud Computing environment, virtual resources are provided by a service or on user demand.

DOI: 10.4018/IJCAC.2015070104
Cloud Computing involves the concept of delivering the resources to the end-user over World Wide Web (www) Wei, H (2010). Due to its rapid elasticity (Capabilities can be elastically provisioned and released according to end user demand) feature a user can acquire the resources from the cloud which can be released at any time Yarmolenko et al. (2006). From the vast number of resources available on the cloud, end-user is required to pay only for services demanded and provided. There are number of virtual machines present at cloud datacenter and each virtual machine handles one resource respectively and since the resources are used on request of end user, therefore the cost increases automatically which becomes a major bottleneck in the deployment of too many virtual machines. With an unlimited number of resources at cloud data center, there allocation and discovering active and most suitable service resource is another major challenge.

The major contribution of this work is to propose a novel algorithm for resource allocation considering cost optimization of virtual machines as primary factor. The work proposes to deploy software agents to achieve such an effective framework. Mobile agents appears to show potential as these execute asynchronously, autonomously, reduce network load and are adaptive in nature, especially in distributed applications.

The paper is structured as follows. Section 2 discusses the related work in this field. Section 3 describes the proposed technique, flowchart and algorithm based on it. Section 4 elaborates on the results and comparisons with the existing techniques. Finally conclusion is depicted in section 5.

2. RELATED WORK

The section throws light on the work of some renowned researchers who had been pillars and founders of the current research work.

Compiere is directing ERP software and CRM system. This organization provides an individual VM for each customer to maintain service level agreement (SLA) requirements such as response time and capacity. This induces the wastage of hardware resources which results in high infrastructure cost. However customers may not use complete VM capacity which is reserved to serve their requests. Schneider et al. Schneider et al. (2009) proposed an adaptive algorithm to adjust the level of parallelism at runtime so that the system can handle collapse data which is based on the current workload on the node. Fu et al. (2002) proposed an SLA-based dynamic scheduling algorithm of distributed resources for streaming. Gaber et al. (2009) adapted the data mining algorithm output on streaming applications according to resource availability and data arrival rate. Moreover, Yarmolenko et al. (2006) evaluated various SLA-based scheduling heuristics on parallel computing resources using resource utilization and income as evaluation metrics. Lee et al. (2010) looked into the profit impelled service request scheduling for workflow. Zheng et al. (2009) proposed binary integer programming method to solve independent optimization problems while changing multiplex strategies of initial optimal solution by minimizing their loss. The proposed algorithms are useful for linear problems not for dynamic and complex problems. From the literature survey this has been observed that there is need to pay more attention on the resource scheduling policy. The main purposes of scheduling algorithms are to minimize the resource starvation and to ensure for providing the effective and fairness resources. Scheduling deals with the problem of decision making to find out the outstanding request and which request is to be allocated resources. Yet none of algorithm exists which is effective and fair for resource scheduling in cloud computing. This work focuses on scheduling enterprise applications on VMs in cloud computing environments.
Comparative Analysis of International Education Systems
[www.igi-global.com/chapter/comparative-analysis-international-education-systems/50226?camid=4v1a](www.igi-global.com/chapter/comparative-analysis-international-education-systems/50226?camid=4v1a)

The Impact of Contract Type on Service Provider Information Requirements
[www.igi-global.com/article/impact-contract-type-service-provider/71945?camid=4v1a](www.igi-global.com/article/impact-contract-type-service-provider/71945?camid=4v1a)