Toward A Performing Resource Provisioning Model for Hybrid Cloud

Mohammed Rebbah, University of Mascara, Mascara, Algeria
Yahya Slimani, Department, ISAMM, University of Manouba, Tunisa, Tunisia
Mohammed Debakla, University of Mustapha Stambouli, Mascara, Algeria
Omar Smail, University of Mustapha Stambouli, Mascara, Algeria

ABSTRACT

This article describes how the idea of a hybrid cloud comes from the coupling of public and private clouds to more efficiently address user requirements. This article addresses the problem of resource provisioning in hybrid cloud. This article is mainly concerned about optimizing the resources provisioning task through the reduction of the tasks completion time together with minimal cost and more reliable services. Two steps are considered in the proposed model, which are brokering and scheduling. In the brokering strategy, this article formalizes the problem as a minimization problem of the completion time as the objective function, under cost and service reliability constraints. The scheduling strategy contains two phases: (i) use the balanced k-means method to classify the submitted tasks and, (ii) perform a minimum assignment using the Hungarian algorithm. The proposed model is evaluated within the simulation framework CloudSim. Experimental results demonstrate that the provisioning model significantly reduces both the response time and the slowdown of user’s requests for different scheduling algorithms.

KEYWORDS

Assignment problem, Balanced k-means, Hybrid cloud computing, Reliability, Resource provisioning

1. INTRODUCTION

The term Cloud Computing usually refers to online delivery and consumption model for business and customer services. Clouds provide a virtualized platform for users to create and manage the software stack from the operating system to the applications. This particular type of cloud is known as an Infrastructure-as-a-Service (IaaS) cloud. The customizability, complete control over the software stack, and on-demand access to IaaS cloud make them an attractive solution to the problem of dynamically extending the resources of a static site to adjust to changes in demand (Marshall, Keahey, Freeman, 2010). In Cloud computing (Buyya, Yeo, Venugopal, Broberg, Brandic, 2009), resources can be either externally owned (public cloud), or internally owned (private cloud), the former being offered by Cloud providers. Public clouds offer access to external users who are typically billed on a pay-as-you-use economic model. Resource size is dynamic, growing by way of on-demand creation of the resources of the desired type (e.g., virtual machines VMs), this kind of dynamic sizing being supported by virtualization technologies that enable dynamic creation, migration, and destruction of resources. IaaS can customize and configure VMs-based on application demands providing massive scalability, high reliability and performance (for example, the Amazon Elastic Compute Cloud (EC2)

DOI: 10.4018/IJGHPC.2018100102

Copyright © 2018, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.
(Amazon, n.d.a)), as well as virtual storage (for example, the Amazon Simple Storage Service (S3) (Amazon, n.d.b)).

Utilizing public Cloud services along with local resources (e.g., private Cloud) to support Hybrid Clouds (Sotomayor, Montero, Llorente, Foster, 2009A), is one of the most widely used Cloud computing model. Hybrid Cloud platforms help scientists and businesses leverage the scalability and cost effectiveness of the public Cloud by paying only for IT resources consumed (server, connectivity, storage) while delivering the levels of performance and control available in private Cloud environments without changing their underlying IT setup (He & Wang, 2015). However, efficient policies to integrate public and private Clouds remain a challenge (Javadi, Abawajy & Sinnott, 2012).

But, to provide the same features found in commercial clouds, hybrid cloud software must meet a set of requirements: provide a uniform and homogeneous view of virtualized resources, regardless of the underlying virtualization platform (e.g., Xen, KVM, VMWare, etc.); manage the full lifecycle of a virtual machine, including setting up networks dynamically for groups of VMs and managing the storage requirements of VMs, such as deployment of VM disk images or on-the-fly creation of software environments; support for configurable resource allocation policies to meet the specific goals of the organization (e.g., high availability, server consolidation to minimize power usage, etc.); and, adaptability to an organization’s changing resource needs, including peaks where local resources are insufficient, and changing resources, including addition or failure of physical resources. Thus, a key component in hybrid cloud will be virtual infrastructure management, the dynamic orchestration of virtual machines on a pool of physical resources, meeting the requirements outlined above (Sotomayor, Montero, Llorente, Foster, 2009B), de Assunção, di Costanzo and Buyya (2009) proposed scheduling strategies to integrate resources from the public Cloud provider and local cluster. In this work, the requests are first instantiated on cluster and in the event more resources are needed to serve user requests, IaaS Cloud provider virtual machines are added to the cluster. This is done to reduce users’ response time. Wang, Chang, Lo and Lee (2013) propose the Adaptive Scheduling with QoS Satisfaction algorithm, namely AsQ, for the hybrid cloud environment to raise the resource utilization rate of the private cloud and to reduce task response time as much as possible. The AsQ tries to maximize the utilization rate of the private cloud and to minimize the renting cost of the public cloud. However, despite all the advantages hybrid Cloud services can provide. Improved security is another main benefit. Users run sensitive applications and store sensitive data in the private cloud platform. There are many security issues such as managing the communication link between both sites during delivery of IT infrastructure from cloud to enterprises, defining firewall with optimal rules to allow only approved traffic from cloud, incompatible network policies (Mather, Kumaraswamy & Latif, 2009). First of all, we can note the security aspect is out of scope of this paper. Nevertheless, security is an important issue if we use hybrid, as in our case. In the context of our work, we have to study firstly, the provisioning resources for public cloud from private one: in this case, the security is not critic because we offer a new property for public resources, namely security and secondly, the provisioning resources for private cloud from public one: in this case, the security is an important issue because public cloud does not guarantee security. This case is not discussed in the current version of our paper. But, as first approach, we can provide resource for private cloud for data that are not very critical from security view point.

In this paper, minimizing both task’s execution time and cost with a reliable service in cloud computing environment is our motivation to explore a resource provisioning model for hybrid cloud. Our proposed model is based on two strategies executed by the InterGrid Gateway (IGG). In the first strategy, called brokering, a model composed of an objective function under a set of constraints is used to solve an optimal resources provisioning problem. The solution of this problem is the role of our strategy, called scheduling, which passes through two phases. In the first one, we apply a classification by the k-means method (MacQueen, 1967), followed with a minimum allocation by the Hungarian algorithm (Kuhn, 1955). This second phase is completed by a distribution of all the tasks grouped in clusters to local IGGs.
Dynamic Reconfigurable NoC (DRNoC) Architecture: Application to Fast NoC Emulation
www.igi-global.com/chapter/dynamic-reconfigurable-noc-drnoc-architecture/44227?camid=4v1a

Adjusting Thread Parallelism Dynamically to Accelerate Dynamic Programming with Irregular Workload Distribution on GPGPUs
www.igi-global.com/article/adjusting-thread-parallelism-dynamically-to-accelerate-dynamic-programming-with-irregular-workload-distribution-on-gpgpus/114710?camid=4v1a