Chapter 11

Heuristic Task Consolidation Techniques for Energy Efficient Cloud Computing

Dilip Kumar
National Institute of Technology, India

Bibhudatta Sahoo
National Institute of Technology, India

Tarni Mandal
National Institute of Technology, India

ABSTRACT

The energy consumption in the cloud is proportional to the resource utilization and data centers are almost the world’s highest consumers of electricity. The complexity of the resource allocation problem increases with the size of cloud infrastructure and becomes difficult to solve effectively. The exponential solution space for the resource allocation problem can be searched using heuristic techniques to obtain a sub-optimal solution at the acceptable time. This chapter presents the resource allocation problem in cloud computing as a linear programming problem, with the objective to minimize energy consumed in computation. This resource allocation problem has been treated using heuristic approaches. In particular, we have used two phase selection algorithm ‘FcfsRand’, ‘FcfsRr’, ‘FcfsMin’, ‘FcfsMax’, ‘MinMin’, ‘MedianMin’, ‘MaxMin’, ‘MinMax’, ‘MedianMax’, and ‘MaxMax’. The simulation results indicate in the favor of MaxMax.

1. INTRODUCTION

Cloud computing infrastructures are designed to support the accessibility and deployment of various service oriented applications by the users (Hwang, Fox, & Dongarra, 2012; Mell & Grance, 2011). Cloud computing services are made available through the server firms or data centers. The concept of cloud computing has been emerging from the concept of heterogeneous distributed computing, grid computing, utility computing and autonomic computing (Buyya, Broberg,
A cloud computing is the convergence of 3 major trends, these trends are Virtualization, Utility components, and Software as a service. To meet the growing demand for computations and large volume of data, the cloud computing environment provides high performance servers and high speed mass storage devices (Beloglazov, Abawajy, & Buyya, 2012). These resources are the major source of the power consumption in data centers along with air conditioning and cooling equipment (Rodero et al., 2010). Moreover the energy consumption in the cloud is proportional to the resource utilization and data centers are almost the world’s highest consumers of electricity (Buyya, Beloglazov, & Abawajy, 2010a). Due to the high energy consumption by data centers, it requires efficient technology to design green data center (Liu et al., 2009). Cloud data center, on the other hand, can reduce the energy consumed through server consolidation, whereby different workloads can share the same server or physical host using virtualization and hence un-used servers or physical host can be switched off.

Power management represents a collection of IT processes and supporting technologies geared toward optimizing data center performance against cost and structural constraints. This includes increasing the deploy-able number of servers per rack, when the racks are subject to power or thermal limitations, and making power consumption more predictable and easier to plan for. Power manager comes in two categories: static and dynamic. Static power management deals with fixed power caps to manage aggregate power, while the policies under dynamic power management take advantage of additional degrees of freedom inherent through virtualization, as well as the dynamic behaviors supported by advanced platform power management technologies (ITU, 2012).

Generally, clouds are deployed to customers giving them three levels of access: Software-as-a-Service (SaaS), Platform-as-a-Service (PaaS), and Infrastructure-as-a-Service (IaaS). Clouds use virtualization technology in distributed data centers to allocate resources to customers as they need them. The task originated by the customer can differ greatly from customer to the customer. Entities in the Cloud are autonomous and self-interested; however, they are willing to share their resources and services to achieve their individual and collective goals. In such an open environment, the scheduling decision is a challenge given the decentralized nature of the environment. Each entity has specific requirements and objectives that need to achieve. Server consolidation is allowing the multiple servers running on a single physical server simultaneously to minimize the energy consumed in a data center (Ye, Huang, Jiang, Chen, & Wu, 2010). Running the multiple servers on a single physical server are realized through virtual machine concept. The task consolidation is also known as server/workload consolidation problem (Lee & Zomaya, 2012). Resource allocation problem discussed in this chapter is the task consolidation problem on cloud data center. Task consolidation problem addressed in this chapter is to assign a task to a set of resources in a cloud computing environment. This energy efficient load management maintains the utilization of all computing nodes and distributes virtual machines in a way that is energy efficient. The goal of this algorithm is to maintain availability to compute nodes while reducing the total power consumed by the cloud.

Cloud computing resources are managed through the centralized resource manager. The centralized resource manager assigned the tasks to the appropriate VMs. The resources of cloud data center are available to the users/applications through Virtual Machines (VMs). Virtual Machines are used to meet the resource requirement and run time support for the applications. In particular executing an application for required resource can be made available through two steps: creating an instance of the virtual machine as required by the application (VMs provisioning) and scheduling the request to the physical resources.
Related Content

Location-Based Advertising Using Location-Aware Data Mining
www.igi-global.com/chapter/location-based-advertising-using-location-aware-data-mining/206596?camid=4v1a

Predictive Modeling for Imbalanced Big Data in SAS Enterprise Miner and R
www.igi-global.com/article/predictive-modeling-for-imbalanced-big-data-in-sas-enterprise-miner-and-r/210567?camid=4v1a

Information Dispersal Algorithms and Their Applications in Cloud Computing
www.igi-global.com/chapter/information-dispersal-algorithms-and-their-applications-in-cloud-computing/195343?camid=4v1a

Novel Taxonomy to Select Fog Products and Challenges Faced in Fog Environments