

SCOPUS

Improving Overall usage of Servers by Measuring Uneven Utilization of a Server and allocating the Applications in the Face of Multidimensional Resource Constraints

S.K. Sonkar^{1*}, M. U. Kharat²

¹Dept. of computer engineering, KKWagh IEE&R Nashik, SP Pune University, India

²Dept. of computer engineering, METIOE Nashik, SP Pune University, India

*Corresponding Author: Sonkar83@gmail.com, Tel.: +919970061302

Available online at: www.ijceonline.org

Accepted: 20/Sept/2018, Published: 30/Sept/2018

Abstract—Major objective of cloud provider is to maximize the resource utilization of cloud servers as well as to reduce the energy consumption and operative cost of the datacenter. However, the servers in many existing datacenters are underutilized in practice due to over-provisioning of peak demand. Many times, the datacenter come across situations wherein large number of application requests simultaneously demand multidimensional resources such as CPU, memory, bandwidth. In such situations it is highly impractical for the cloud service provider to satisfy the application requests of all the users within stipulated time, especially when sufficient resources are not available with them. In order to address this problem, we designed a system which measures the resource utilization of all servers before allocating the application requests to server and then dynamically allocate the application requests to the server which is underutilized. This yields in improving the overall utilization of servers. Our system initially checks the server utilization in terms of CPU, Memory and Bandwidth resource utilization against predefined threshold value. If resource of any server goes beyond its threshold value, then application request will not be allocated to that server to avoid the server overloading. That means our system redirect the application request to the underutilized server so as to improve the server resource utilization in the face of multidimensional resource constraints. The experimental results demonstrate that our system improves the overall server resource utilization by 10%.

Keywords— Cloud Service Provider, User Request, Resource utilization, Resource constraints.

I. INTRODUCTION

The cloud access information globally, it is available anytime and anywhere on demand [1]. Cloud server is created to manage and store data and runs different applications & deliver the services

The cloud provides mainly three types of services. (SaaS) Software as a service, in this application runs on cloud Infrastructure and accessible from different client devices through a program interface or web browser. (PaaS) Platform as a service, in this get a core OS and other building block services which allow to run third-party or your own application. Application created using programming languages, services, libraries and tools supported by the provider. (IaaS) Infrastructure as a Service, this provide different services such a Storage, virtualization, Load balancers and Networking etc.

There are four types of cloud deployment models on the basis of size and access. Public Cloud Infrastructure is open to use, flexible, low Cost and Reliable. Private Cloud is owned by particular organization, enterprise or institution. Private cloud provides better security, better control than the public cloud. Community cloud is a cloud shared manually by specific community from different organization or firms and use for ventures, joint businesses, and research organization. Hybrid cloud is a combination between two or more clouds (public, private or community) [2].

Virtualization is the technology is to create virtual version of actual machines such as desktop, server, operating system, and network or storage device. Due to the virtualization multiple applications and OS can run simultaneously on the same machine. This minimizes the cost and increases flexibility, utilization and efficiency of exiting computer hardware [3].

Resource Management is a main issue in cloud computing. Providing all requested resource is the challenging task to cloud provider. Resource management and allocation is very