International Journal of Research
Available at https://edupediapublications.org/journals
e-ISSN: 2348-6848
p-ISSN: 2348-795X
Volume 05 Issue 16
June 2018

A Literature Review on Resource Allocation in Cloud Environment

Nimanshi & Ms Mamta

[1]Department of Computer Science South Point Institute of Technology and Management (SITM)
Deenbandhu Chhotu Ram University of Science & Technology (DCRUST), Sonepat
[1]nimanshisehgal07@gmail.com
[2]Department of Computer Science South Point Institute of Technology and Management (SITM)
Deenbandhu Chhotu Ram University of Science & Technology (DCRUST), Sonepat
[2] mamtakalra21@gmail.com

Abstract— *Cloud Computing provides us a means by which we can access the applications as utilities, over the Internet. It allows us to create, configure, and customize applications online. It offers online data storage, infrastructure and application. The term Cloud refers to a Network or Internet. Applications such as e-mail, web conferencing, customer relationship management (CRM), all run in cloud. Cloud computing is highly promising technology because of its unlimited resource provisioning and data storage services which help us in managing the data as per requirements. Resource allocation is process of assigning the available resources in an economic way and efficient and effective way. This paper provides review of different resource allocation strategies in cloud environment.*

Keywords— Cloud Computing, Resource Allocation, Resource Provisioning

I. INTRODUCTION

Cloud computing [1] is a technology that uses the internet and central remote servers to maintain data and applications. Cloud computing allows consumers and businesses to use applications without installation and access their personal files at any computer with internet access. Cloud computing is a comprehensive solution that delivers Information Technology as a service (figure 1).

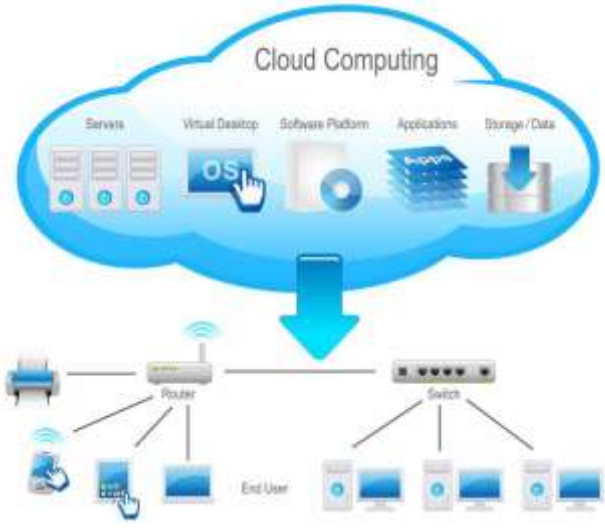

Figure 1: Cloud Computing

Cloud Computing provides us a means by which we can access the applications as utilities, over the Internet. It allows us to create, configure, and customize applications online. It offers online data storage, infrastructure and application. The term **Cloud** refers to a **Network** or **Internet**. In other words, we can say that Cloud is something, which is present at remote location. Cloud can provide services over network, i.e., on public networks or on private networks, i.e., WAN, LAN or VPN. Applications such as **e-mail, web conferencing, customer relationship management (CRM),** all run in cloud. This technology allows for much more efficient computing by centralizing storage, memory, processing and bandwidth.

The cloud makes it possible for users to access information from anywhere anytime. It removes the need for users to be in the same location as the hardware that stores data. Once the internet connection is established either with wireless or broadband, user can access services of cloud computing through various hardware. This hardware could be a desktop, laptop, tablet or phone.

Cloud computing comprises of 2 components —the front end and the back end.

The front end includes client's devices and applications that are required to access cloud. And the back end refers to the cloud itself. The whole cloud is administered by a central server that is used to monitor client's demands.

One of the main services of cloud environment is resource allocation. Resource allocation is process of assigning the available resources in an economic way and efficient and effective way Resource allocation is the scheduling of the available resources and available activities required by those activities while taking into consideration both the resource availability and the project time. Resource provisioning and allocation solves that problem by allowing the service providers to manage the resources for each individual request of resource.

This paper provides review of different resource allocation strategies in cloud environment.

II. APPLICATIONS OF CLOUD COMPUTING

Cloud Computing has its applications in almost all the fields such as **business, entertainment, data storage, social networking, management, entertainment, education, art** and **global positioning system**, etc. Some of the widely famous cloud computing applications are explained below:

Business Applications

Cloud computing has made businesses more collaborative and easy by incorporating various apps such as MailChimp, Google Apps for business, and Quickbooks.

MailChimp: It offers an e-mail publishing platform. It is widely employed by the businesses to design and send their e-mail campaigns.
Google Apps for Business: Google offers creating text documents, spreadsheets, presentations, etc., on Google Docs which allows the business users to share them in collaborating manner.
Quickbooks: It offers online accounting solutions for a business. It helps in monitoring cash flow, creating VAT returns and creating business reports.

Data Storage and Backup

Box.com, Mozy, Joukuu are the applications offering data storage and backup services in cloud.

Box.com: Box.com offers drag and drop service for files. It just required to drop the files into Box and access from anywhere.
Mozy: Mozy offers online backup service for files during a data loss.
Joukuu: is a web-based interface. It allows displaying a single list of contents for files stored inGoogle Docs, Box.net and Dropbox.

Management Applications

There are apps available for management task such as time tracking, organizing notes. Applications performing such tasks are discussed below:

Toggl: It helps in tracking time period assigned to a particular project.
Evernote: Evernote is an application that organizes the sticky notes and even can read the text from images which helps the user to locate the notes easily.
Outright: It is an accounting app. It helps to track income, expenses, profits and losses in real time.

Social Applications

There are several social networking services providing websites such as Facebook, Twitter, etc.

Facebook: Facebook offers social networking service. One can share photos, videos, files, status and much more.
Twitter: Twitter helps to interact directly with the public. One can follow any celebrity, organization and any person, who is on twitter and can have latest updates regarding the same.

Entertainment Applications

Audiobox.fm: It offers streaming service, i.e., music can be stored online and can be played from cloud using service's own media player.

III. RESOURCE ALLOCATION STRATEGY

Resource Allocation Strategy (RAS) [10] is all about integrating cloud provider activities for utilizing and allocating scarce resources within the limit of cloud environment so as to meet the needs of the cloud application. It requires the type and amount of resources needed by each application in order to complete a user job. The order and time of allocation of resources are also an input for an optimal RAS. An optimal RAS should avoid the following criteria as follows:

a) Resource contention situation arises when two applications try to access the same resource at the same time.
b) Scarcity of resources arises when there are limited resources.
c) Resource fragmentation situation arises when the resources are isolated.
d) Over-provisioning of resources arises when the application gets surplus resources than the demanded one.
e) Under-provisioning of resources occurs when the application is assigned with fewer numbers of resources than the demand.

From the perspective of a cloud provider, predicting the dynamic nature of users, user demands, and application demands are impractical. For the cloud users, the job should be completed on time with minimal cost. Hence due to limited resources, resource heterogeneity, locality restrictions, environmental necessities and dynamic nature of resource demand, we need an efficient resource allocation system that suits cloud environments.

Cloud resources consist of physical and virtual resources. The physical resources are shared across multiple compute requests through virtualization and provisioning. The request for virtualized resources is described through a set of parameters detailing the processing, memory and disk needs. Provisioning satisfies the request by mapping

virtualized resources to physical ones. The hardware and software resources are allocated to the cloud applications on-demand basis. For scalable computing, Virtual Machines are rented. The complexity of finding an optimum resource allocation is exponential in huge systems like big clusters, data centres or Grids. Since resource demand & supply can be dynamic and uncertain.

Advantages:

▪ The first major benefit of resource allocation is that user neither has to install software nor hardware to access the applications, to develop the application and to host the application over the internet.
▪ The next major benefit is that there is no limitation of place and medium. We can reach our applications and data anywhere in the world, on any system.
▪ The user does not need to expend on hardware and software systems.
▪ Cloud providers can share their resources over the internet during resource scarcity.

Limitations:

▪ Since users rent resources from remote servers for their purpose, they don't have control over their resources.
▪ Migration problem occurs, when the users wants to switch to some other provider for the better storage of their data. It's not easy to transfer huge data from one provider to the other.
▪ In public cloud, the clients' data can be susceptible to hacking or phishing attacks. Since the servers on cloud are interconnected, it is easy for malware to spread.
▪ Peripheral devices like printers or scanners might not work with cloud. Many of them require software to be installed locally. Networked peripherals have lesser problems.
▪ More and deeper knowledge is required for allocating and managing resources in cloud, since all knowledge about the working of the cloud mainly depends upon the cloud service provider.

IV. History of Cloud Computing

Cloud computing is achieving more and more acceptability day by day. You have likely been using the cloud for some years now, with such things as Google-apps, MSN Messenger, Skype and Flickr. The idea started in the 1960s when John McCarthy thought of computation as a public utility [4, 5, 6]. Distributed computing appeared with organizations and universities offering dialup in the late 1970s [7]. Grid computing in the early 1990s aimed at providing easy access to computer power like an electric power grid.

In various contexts the term "cloud" has been used to describe large ATM networks in the 1990s [10]. A major shift in the 1990s was observed due to the rise of the Internet and the increase of speed for cheap Internet connections. The idea of Virtual Private Networks (VPNs) was discovered after the need for a safe and secure data transfer among the communication between branches [8]. These solutions required load balancing to optimize resource utilization. VPN is more secure than simple dialup, but connectivity to the outside world requires additional security measures.

On the other hand, Web 2.0 shifted the Web to a more interactive and collaborative manner, assured peers' social interaction and collective intelligence, and introduced new opportunities for influencing the Web and attracting its users more efficiently. Enterprises were rapidly adopting Web 2.0, which is the second phase in the Web's evolution [9].

Various computing paradigms were presented during the 21st century. The popular ones between them are cluster, grid, and cloud computing [10]. Among the popular names that are linked to cloud computing are Sales force with the idea of supplying enterprise applications through a website, Amazon with its Amazon Web Services (AWS) and Amazon's Elastic Computing Cloud (EC2), Microsoft and its famous Windows Azure, Google with its several services such as Google Docs which gave cloud computing a great push and public visibility. Eucalyptus, OpenNebula and Nimbus were introduced as the first open source platforms for deploying private, as well as hybrid, clouds [11, 12, 13]. These were designed around different core uses of cloud computing parallel processing, distributed computing and creation of virtual frameworks in order to provide Virtual Machines (VMs) to users on demand. Other famous organizations such as IBM, Oracle, Dell, Fujitsu, Teradata, HP, Yahoo, and a number of other important names introduced cloud computing after that.

V. Literature Review

Several works related to our work, which presents the efficient resource allocation in cloud computing as describe below:

A. Meera and S. Swamynathan [4] proposed an approach for allocating resources based on the analyzed data that is being analyzed by a monitoring agent. The monitoring agent will collect the resources usage information that is currently being used by a virtual machine and will display it in a dashboard. Statistical report that is being displayed on a

dashboard provides information for cloud administrator for better optimization of resources.

K. C. Gouda, Radhika T. V., and Akshatha M. [5] proposed a priority based resource allocation approach with minimum wastage and a maximum profit. Priority is being considered in terms of different parameters like time, cost, numbers of processor requests etc. priority can be used for better resource allocation in cloud environment.
Diptangshu Pandit, Matangini Chattopadhyay, and Nabendu Chaki [6] proposed an efficient resource allocation algorithm with the use of simulated annealing. In this approach authors had introduced the concept of bin, soft computing and simulated annealing. In this approach, problem of resource allocation is being solved with the help of bin packing problem. In this approach temperature is being considered as a control parameter but no formal procedure of selecting the temperature has been described in this approach.

Nguyen Trung Hieu, Mario Di Francesco, Antti YlaJaaski [7] wrote a paper "A Virtual Machine Placement Algorithm for Balanced Resource Utilization in Cloud Data Centers". In this article, they proposed an algorithm, called Max-BRU, that maximizes the resource utilization and balances the usage of resources across multiple dimensions. Their algorithm is based on multiple resource-constraint metrics that help to find the most suitable server for deploying VMs in large cloud data centers. The proposed Max-BRU algorithm is evaluated by simulations based on synthetic datasets. Experimental results show two major improvements over the existing approaches for VM placement. First, Max-BRU increases the resource utilization by minimizing the amount of physical servers used. Second, Max-BRU effectively balances the utilization of multiple types of resources.

Sumita Bose, Jitender Kumar [8] "An Energy Aware Cloud Load Balancing Technique using Dynamic Placement of Virtualized Resources" in their papers provide provision of migration of resources from one cloud to another for efficient load balancing in cloud environment. This migration concept provides extra overhead for cloud processing.

Pratik P. Pandya, Hitesh A. Bheda [9] "Dynamic Resource Allocation Techniques in Cloud Computing" in their papers provide provision of dynamic allocation of resources for the client. They also provide provision of migration from one cloud to another for efficient load balancing in cloud environment.

VI. CONCLUSION

A cloud environment is one of the most shareable environments where multiple clients are connected to the common environment to access the services and the products. A cloud environment can be public or the private cloud. In such environment, all the resources are available on an integrated environment where multiple users can perform the request at same time. In such case, some approach is required to perform the effective scheduling and the resource allocation. This paper provided review of different resource allocation strategies in cloud environment.

REFERENCES

[1] Anthony T.Velte, Toby J.Velte, Robert Elsenpeter, "Cloud Computing, A Practical approach"

[2] B. Hayes, "Cloud Computing," *Commun. ACM*, vol. 51, no. 7, pp. 9–11, Jul. 2008.

[3] P. Mell and T. Grance, "The NIST definition of cloud computing (draft)," *NIST special publication*, vol. 800, no. 145, p. 7, 2011.

[4]. A. Meera, S. Swamynathan, "Agent based Resource Monitoring system in IaaS Cloud Environment", International Conference on Computational Intelligence: Modeling Techniques and Applications (CIMTA), 2013.

[5] K C Gouda, Radhika T V, Akshatha M, "Priority based resource allocation model for cloud computing", International Journal of Science, Engineering and Technology Research (IJSETR) Volume 2, Issue 1, January 2013.

[6]. Diptangshu Pandit, Matangini Chattopadhyay, and Nabendu Chaki, "Resource Allocation in Cloud using Simulated Annealing", Applications and Innovations in Mobile Computing (AIMoC), Feb. 27 2014 – March 1 2014, Pages 21-27.
[7] Nguyen Trung Hieu, Mario Di Francesco, and Antti YlaJaaski, "A Virtual Machine Placement Algorithm for Balanced Resource Utilization in Cloud Data Centers", 2014 IEEE International Conference on Cloud Computing © 2014 IEEE.

[8] Sumita Bose, Jitender Kumar "An Energy Aware Cloud Load Balancing Technique using Dynamic Placement of Virtualized Resources" Advances in Computer Science and Information Technology (ACSIT) Volume 2, Number 7; April – June, 2015 pp 81 – 86

[9] Pratik P. Pandya, Hitesh A. Bheda "Dynamic Resource Allocation Techniques in Cloud Computing", International Journal of Advance Research in Computer Science and Management Studies Volume 2, Issue 1, January 2014

[10] V. Vinothina, Dr. R. Sridaran, Dr. Padmavathi Ganapathi, "A Survey on Resource Allocation Strategies in Cloud Computing", International Journal of Advanced Computer Science and Applications, Vol. 3, No.6, 2012.