Engineering and Technology Journal
Journal homepage: engtechjournal.org

Data Center Enhancement by Server Resources Utilization

Haider A. Ghanem [a*], Rana F. Ghani [b], Mohammed N. Fadhil [c]

[a] Department of Computer Science, University of Technology, Baghdad, Iraq, Haider8889@Gmail.com

[b] Department of Computer Science, University of Technology, Baghdad, Iraq, 110016@uotechnology.edu.iq

[c] Department of Computer Science, University of Technology, Baghdad, Iraq, mohammad.n.fadhil@uotechnology.edu.iq

* Corresponding author.

Submitted: 07/04/2019 Accepted: 23/08/2019 Published:25/03/2021

K E Y W O R D S

Data Center, Horizontal Scaling, Vertical Scaling, Virtual Machine.

A B S T R A C T

Data centers are the main nerve of the Internet because of its hosting, storage, cloud computing and other services. All these services require a lot of work and resources, such as energy and cooling. The main problem is how to improve the work of data centers through increased resource utilization by using virtual host simulations and exploiting all server resources. In this paper, we have considered memory resources, where Virtual machines were distributed to hosts after comparing the virtual machines with the host from where the memory and putting the virtual machine on the appropriate host, this will reduce the host machines in the data centers and this will improve the performance of the data centers, in terms of power consumption and the number of servers used and cost.

How to cite this article: H.A. Ghanem , R.F. Ghani and, M.N. Fadhil "Data Center Enhancement by Server Resources Utilization" Engineering and Technology Journal, Vol. 39, Part B, No. 01, pp. 203-208, 2021.
DOI: https://doi.org/10.30684/etj.v39i1B.174

1. INTRODUCTION

At the beginning computers were demanding great possibilities to start them like cooling and electricity and most of them were for military uses, so that security issues were very important to secure them, and they were hosted in special chambers known as Computer Rooms.

In 1980s servers appeared which started contention of the central systems and became more complicated than them and demanding electricity and cooling too. At data centers boom has come through Internet. Companies that provide Internet access services must quickly equip their facilities to connect to the Internet and operate their devices non-stop to deploy their services online, where installing such equipment was very expensive for many of small companies, lots of companies started building very large facilities and called them Data Centers [1].

Recently, Data Centers and facilities that followed occupy large areas led to increased costs incurred in many companies so that Unified Data Centers have appeared for cost savings, huge development in computer networks led to increased uses whereas become a practical fact in many

sectors like business, universities, companies, banks and many others. With spreading of these networks many of separate servers have appeared to serve users in a great way without exposing the network to problems or interrupted service so that Data Centers have designed in appropriate environment like determining temperature, regulating electricity, fire extinguishers and high security features which contain a great number of servers of basic and backup power supplies which provide high speed internet to give great services 24 hours per day. With the great development of using devices and technical systems in different fields led to increasing the need to provide great spaces for Data Centers with their facilities and that is not easy and needs a great deal of money, so from that Default Data Center concept has appeared which is a group of servers and other default devices that connected to virtual network existed on less cost devices aim to serve all network users in a great way and with the same competence that performed at work separated from less financial costs [2].

This paper addresses how to improve the performance of data centers by increasing the use of data center resources where virtual servers are the main technology that can be used to increase resource utilization.

2. DATA CENTER SERVICES

Data center services is a collective for the supporting components necessary for the proper operation of a repository the storage, management and dissemination of data organized around a body of knowledge or pertaining to an enterprise. As such, data center services can involve hardware, software, processes, and personnel.

1) Install and maintain devices.
2) Energy distribution management.
3) Data backup and recovery.
4) Balancing load.
5) Internet access control.
6) Manage email and messages.
7) Antivirus and firewalls.
8) Provide technical support.

3. DATA CENTER FEATURES

1) Protect infrastructure from all physical risks.
2) Observation service 24 hours per day to ensure rapid tracking, responding, and processing.
3) Physical and security barriers are used to guarantee access authorized persons the data center.
4) Continuation of climate control system at temperature, humidity and sensors connected with alarm devices to ensure do not hardware crashes and to avoid the risk of sudden fire.
5) Building up-to-date Data Centers.

4. RELATED WORK

This work proposes how to distribute resources of server to virtual machines into consideration different previous works, Hyser and others. [3] proposed designing infrastructure system to put independent VM for the best use of computing resources and cost saving. Data centers uploads are distributed among all available servers and are balanced using resources as much as possible across all types of resources. [4] Compare the impact of vertical and horizontal gradient techniques on cloud resource management. Vertical scaling technology is effective while horizontal measurements improve the overall availability of applications. [5] Isci et.al presents direct access memory based on the deportation VM technology which significantly reduce the overhead costs of migration. With all these suggestions and solution, the contribution of resources on virtual devices is still very important and exploit all server resources for implementation of application, which helps to start the server and do not idle mode to take advantage of data centers operation in terms of sources, power, and cooling.

5. DATA CENTER

Data center are buildings and central storages which are real or default to save, manage and serve big quantities of important client’s data at client's architecture/server. They contain of large number of servers and devices that connected to internet with high speed and work all day. These buildings must provide with special equipment and techniques to save appropriate temperature and electricity without interrupting. Usually, Data Centers demand redundant extensive systems (repeated) or

backup copies from power supply, cooling systems and duplicate network connections. They contain management of Data Centers to ensure the reliability of communication with important data and information stored in data centers stores. Data Centers have different names like: Data Centers, Servers Chambers and Technical Information Chambers, but Data Centers is the most meaningful precisely named. Designing Data Centers building is different from other buildings, so that designer must know everything about measurements, different technical of data centers and different types of wires in one system considering the capacity and regulation of passing and making sure of everything is working in a correct way to do all the services. Data centers architectures and demanding are different very much for example: data centers that designed to provide cloud service like Amazon EC2 which meet requirements of infrastructure and security that completely different from the special data centers. All facilities must choose the suitable data centers to decrease costs and increase gains at the same time [6].

Data centers performance can be improved by increasing resource utilization and virtual server simulation are the main technology that can be used to increase resources where applications are supplied with a set of virtual machines (VMs) that share physical resources and consist of virtual machines operating system (VMs), programs and applications. It gives a division of the potential resources inherent in the physical machine (PM). VM must be able to increase resources to support the growing load. At the same time VM must be able to reduce resources at light load. Excessive provision of resources leads to a lower profit margin, while lack of credit leads to customer dissatisfaction. The solution is that resources are measured dynamically based on the demand for the workload without any interruption. When the resources allocated to the virtual machine are unable to meet user request, additional resources are allocated by means of a measurement technique based on VM. As scalability is the ability of the application to optimize the use of resources at different levels of workload by avoiding over-allocation, lack of use and under provisioning. There are two types of scaling, where the document measurement can be applied to VM either by changing the resource section (like CPU, memory, storage) inside VM device and that is called Vertical Scale which gets more computing power by adding more resources into the vertical machine or adjusting the number of instances of VM and that is called Horizontal Scale which gets more computing power by adding more virtual devices, shown in Figure 1[7].

However, the use of horizontal measurement uses migration VM, subsequently, the VM relay eliminates the processing time and consumes a large amount of IO and network passing. So that the dynamic measurement uses to control the demanding resources amount in VM and the structure proposed in this report is a combination of vertical and horizontal measurement techniques that consider the current workload, the resources allocated to the application and the resources available in the host servers as shown in Figure 2 [8].

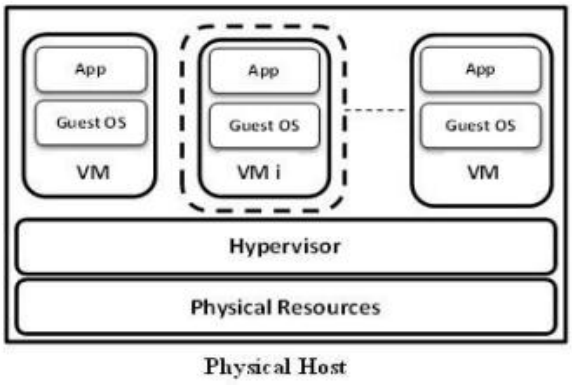

Figure 1: Vertical Scaling

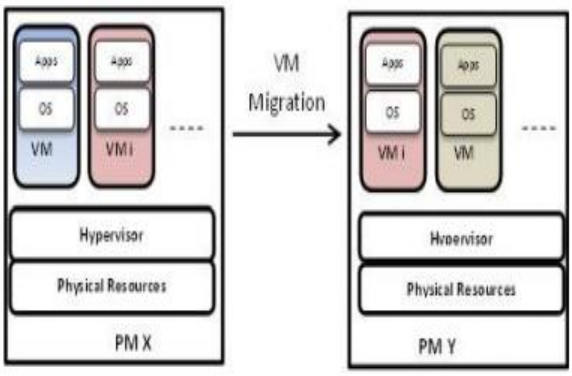

Figure 2: Horizontal Scaling

6. PROPOSED ARCHITECTURE

The contribution of server resources in data centers on the virtual devices leads to improve data centers in terms of the quality of the service provided. The algorithm provided exploits all server memory is taken and through which server memory can be compared with virtual machine memory, where place the virtual machine on the appropriate server to execute the applications and therefore use all the server memory to execute the application, considering that server contains more than one virtual machine according to server memory. When the system needs to other virtual machine to execute applications, so it uses another new server because the old server has been completely used. Figure 3 describes the proposed architecture of algorithm.

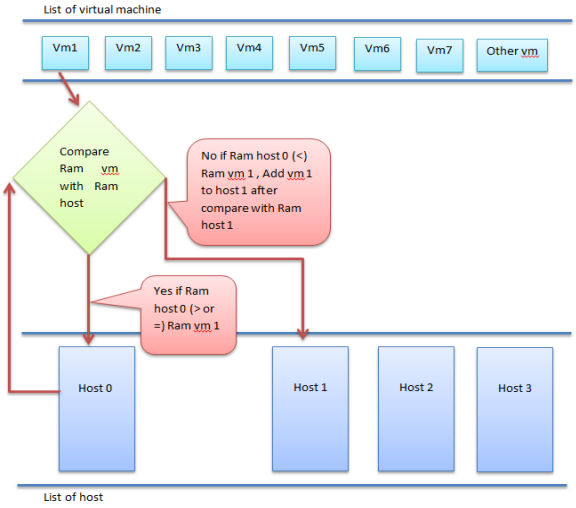

Figure 3: Proposed architecture of algorithm

Algorithm 1:

```
LH = List of  Host
LV = List of  VM
For i=1 To LH
    For j=1 To LV
        If (LH(i). Ram > LV(j). Ram)
          LH(i) = LV(j)
      Else
          LV(j) is failed by Ram in this LH
For j=1 To LV
    Submit task to the LV(j)
End
```

7. RESULTS

The algorithm relies on the distribution of virtual machines on the server by comparing server memory with virtual machine memory and it is necessary to evaluate the virtual data center structure. However, it is very difficult to perform scaling experiments on a real underlying structure in a cloud data center. So that to test the suggested algorithm. (Cloud-Sim) a simulation model has been used to simulate data centers system and uses four hosts in different memory sizes. Where the (host 0) contains memory size (15 GB), (host 1) contains (20 GB), (host 2) contains (25 GB), and (host 3) contains (30 GB), with virtual machines (VM 40), which have been built and contributed inside the server by using algorithm.

Initially, all VMs are randomly placed without using the process of comparing the virtual machine and the server in terms of memory where virtual machines are distributed on the appropriate host by considering the resources needed for the VM and the resources available in the host. Therefore, the default hardware distribution results are displayed in Figure 4, where all four servers are used at the same time. Virtual machines are distributed to all servers without taking advantage of server resources and then moving to another server. This causes a great number of servers to run, high power consumption and excessive cost to increase the number of servers used.

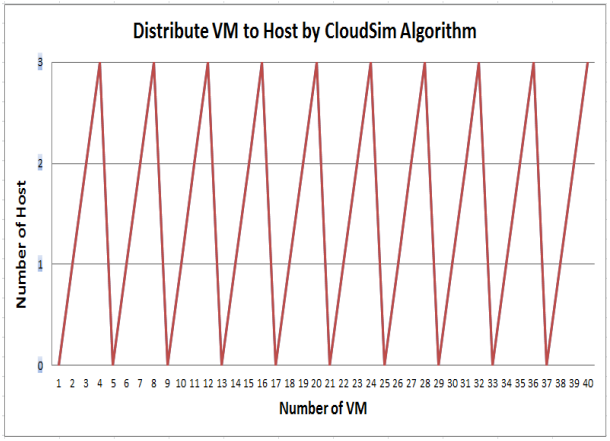

Figure 4: Automatic distribution of VM

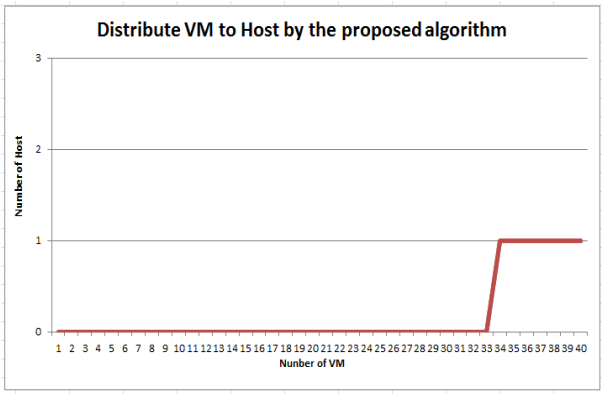

Figure 5: Distribution of VM by proposed algorithm

For the proposed algorithm, the virtual machines will be distributed to the servers after comparing the virtual machines memory with the host device memory see Figure 5, where only (host 0 and 1) are used to distribute virtual machines and this approach will reduces number of servers, power consumption due to the number of servers used, and as well as too reduces the cost because of using fewer servers.

8. CONCLUSION

In data centers, allocating host virtual machine and all host resources is important to improve data center performance in terms of running servers with maximum power and their resources. In this paper, we have identified data centers, their importance, their applications, and their problems, we have worked to exploit one resource to contribute the virtual machines on the host by comparing

memory. In the future we will exploit all the host resources (central processing unit, frequency, storage) to improve Data Centers.

References

[1] Bo Li, Jianxin Li, Jinpeng Huai, Tianyu Wo, Qin Li, Liang Zhong, "Enacloud: An energy-saving application live placement approach for cloud computing environments," IEEE International Conference on, India, 2009.

[2] Haikun Liu, Hai Jin, Xiaofei Liao, Chen Yu, and Cheng-Zhong Xu. "Live virtual machine migration via asynchronous replication and state synchronization." parallel and distributed Systems, IEEE Transactions Vol. 22, no. 12, pp. 1986-1999, USA, 2011.

[3] Sijin He, Li Guo, and Yike Guo. "Real time elastic cloud management for limited resources" In Cloud Computing (CLOUD), International Conference on, pp.622-629. IEEE, USA, 2011.

[4] Wenting Wang, Haopeng Chen, and Xi Chen. "An availability-aware virtual machine placement approach for dynamic scaling of cloud applications." In Ubiquitous Intelligence & Computing and 9th International Conference on Autonomic & Trusted Computing (UIC/ATC), pp. 509-516. IEEE, Japan, 2012..

[5] Canturk Isci, J. Liu, B. Abali, J.O. Kephart, and J. Kouloheris, "Improving server utilization using fast virtual machine migration," IBM Journal of Research and Development, Vol. 55, No. 6, pp.1-4, USA, 2011.

[6] Chien-Yu Liu, Meng-Ru Shie, Yi-Fang Lee, Yu-Chun Lin, and Kuan-Chou Lai, "Vertical/Horizontal Resource Scaling Mechanism for Federated Clouds", In Information Science and Applications (ICISA), 2014 International Conference on, Vol. 1, pp.1-4, IEEE, Seoul, South Korea, 2014.

[7] Trieu C.Chieu, and Hoi Chan. "Dynamic resource allocation via distributed decisions in cloud environment." In e-Business Engineering (ICEBE), IEEE 8th International Conference on, pp. 125-130, Beijing, China, 2011.

[8] Nguyen Van Hien, Frederic Dang Tran, and Jean-Marc Menaud. "Autonomic virtual resource management for service hosting platforms." In Proceedings of the ICSE Workshop on Software Engineering Challenges of Cloud Computing, pp.1-8. IEEE Computer Society, Canada, 2009.