International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-8 Issue-6, August 2019

Resource Allocation in Cloud Computing for Energy Efficiency

Deepa R, Dheeba J

Abstract: ***Cloud computing is a paradigm in which we have virtualized computer systems that deliver services, processing, storage, network, and other fundamental computing resources. Cloud computing enables low cost, device location independence, high reliability, scalability and sustainability. This paper describes the present state of cloud computing research by examining literature, identifying current study trends. We have analyzed the resource allocation method and concluded. It typically designs for high performance that supports the peak resource requirements. After several analyses the power consumption of data center and cloud systems as increased almost several times. There is a lack of research that addresses challenges of managing multiple resources with objective of allocating enough resources for each work load to optimizing power consumption. These papers survey various types of resource allocation algorithms that improve the cloud Infrastructure.***

Keywords : Cloud computing, Resource Allocation, power consumption, data center, multiple resources, cloud infrastructure.

I. INTRODUCTION

Cloud computing was created for the tent name to describe the type of advanced on-demand computing services. The main requirement as a provider-based service is a contribution to cloud computing, storage, and software[1]. It is now as a fast-moving region in a cloud framework and sector. In the cloud, opportunities are established on how to build an application and how to provide various services to the end user via internet services virtualization. Cloud service providers bringing utilization determined large scale computing infrastructure and providing it with a very flexible approach [2]. The purpose of this cloud computing model is to provide more desirable software as a service, improve the accessibility of resources, and increase efficiency. Buyya et al.[1] described it as follows: ' Cloud is a parallel and distributed processing structure composed of a set of dynamically provided and submitted pcs between associated and virtualized pcs as one or more unified registration resources within the framework of the Service Level Agreement (SLA) formed by transaction between the professional organisation and clients. '

Revised Manuscript Received on August 30, 2019.

* Correspondence Author

Deepa R*, B.sc, (Information System Management) 2011, M.sc., (Computer Science) 2014 and studying M.Tech (By Research) in computer Science and Engineering at Vellore Institute of Technology, Vellore, Tamil Nadu, India.

Dheeba J, Computer science and Engineering from Anna University, Chennai. Associate Professor in School of Computer Science and Engineering at Vellore Institute of Technology, Vellore.

Her research interests include medical image analysis, Machine learning approaches and big data analytics.

Varquero et al.[4] stated that "clouds are a wide variety of easy-to-use and accessible virtualized resources (such as hardware, development platforms or services).

These resources can be dynamically reconfigured to suit a variable load, allowing for optimum resource utilization as well. This resource pool is typically used through a pay-per-use model in which the provider of infrastructure guarantees through customized service-level agreements.

Cloud computing is defined by the National Institute of Standards and Technology[5] as "... a pay-per-use system that allows access to a shared pool of configurable computing resources (e.g. networks, servers, storage, applications, services) that can be supplied and released rapidly with minimal effort or interaction among service suppliers.

II. CLOUD COMPUTING TYPES

Three cloud computing services are split by suppliers based on capacity and model [1].

A. Infrastructure as a service (IAAS)
B. Platform as a service and (PAAS)
C. Software as a service (SAAS)

A. IaaS

It offers virtualized resources that are called as a service infrastructure, such as computing, storage applications, and communication.

The cloud infrastructure has server provisioning specifications and works with custom software stack various operating system alternatives.

An infrastructure supplier has been investigating that bottom layer of cloud systems [1].

B. PaaS

Importance of infrastructure-oriented clouds offering raw computing and storage services and another way is supplied with a higher level of consideration as a cloud that is efficiently programmable, called a service platform. This offers developers with an atmosphere for developing and deploying applications, and no need to know how many processors they would use or how much memory they would use. In addition, different programming models and specific services are given as construction blocks for new applications.

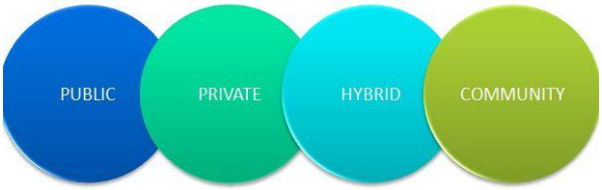
C. SaaS

This layer has provided services that can be accessed by customers through web portals. As a result, clients have increasingly moved from locally installed computer programs to the internet facilities it offers the same functionally. Let's see that it is possible to access traditional desktop applications such as word processing and spreadsheet as a web service and this model is called service software.

Cloud computing has classified according to their deployment models that are [6][7]

Deployment Model

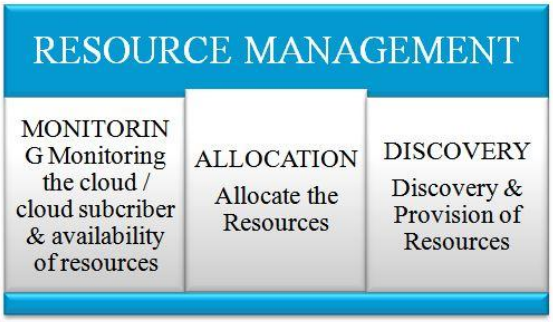
- Public
- Private
- Hybrid
- Community



Public

Here we concentrate primarily as a service model on the infrastructure. It provides virtualized resources such as (computing, storing, and communicating) called as IAAS. Cloud provider resource allocation is a significant challenge in cloud computing, and resource allocation is emerging in studies into the cloud setting. Many researchers work to allocate funds efficiently.

Resource Management Model



III. LITERATURE SURVEY

Resource allocation in cloud computing is discussed through the kinds of problem, algorithm used, advantages, disadvantages and efficiency provided by various proposals are discussed here in this paper. Hadi Goudarzi et al.[9] suggested a technique for minimizing the data center's operating costs. For hierarchical resource management, the benefit of this algorithm is used. It's very scalable and performs better. The disadvantage includes there are smaller number of migrations because of resource limits. The utilization of the suggested management algorithm reduces the data center's operating costs by up to 40% and decreases the algorithm for resource management by up to 86%. Xiaolong Xu et al .[10] This paper is dicussed on energy-conscious VM deployment for the execution of science workflow and proposed that the algorithm using EnReal an Energy Aware Resource Allocation Algorithm is very effective. Here's a switch pm mode that also causes the consumption of energy. Efficient and effective is an effectiveness of this suggested technique. Hongbing Wang et al [11] have investigated the objective of this paper to improve the effectiveness of the cloud computing service and to allocate the suitable or best resources that will truly improve cloud computing facilities. The proposed system is the Cloud Resource Allocation Algorithm through Fitness-enabled Auction (CRAA / FA) and is more efficient for cloud suppliers, but not only for the single renter, but also for the entire cloud computing platform or resources. But due to study funding constraints, it is difficult to experiment with this algorithm with more than one thousand servers or virtual machine. The efficacy of the distribution is more efficient than the continuing dual auction. Dan Li et al.[12] addressed that the data center, the resource allocation model for cloud employment, both fulfill their cloud computing deadlines to guarantee job performance so that data center resources can be used flexibly and efficiently. The algorithm presents that Dcloud accepts more work than the current solution. Dcloud scheme can boost the cloud income provider by 50 percent, but despite lower individual work costs. Lena Mashayekhy et al.[13] researched pay claims in cloud computing as you go model, as customers charged based on internet system involve a user sending a request or some of the funds assigned are permitted and accessible. The proposed scheme is an internet auction-based mechanism for supplying, allocating and pricing VM in clouds, it is consistent with incentives and customers to disclose their demands helping consumers and customers to compete in a good way. The comprehensive experimental outcome is given by this mechanism. Decide and explore fresh internet processes in cloud environments for future job are on this suggested scheme. Rodrigo da et al.[14] suggested scheme is Auto Elastic provides effective energy consumption and resource among the amount of virtual machine users without any user interposition, making use of elasticity management for both cloud administrators and cloud users. The proposed scheme is a model based on auto-elastic paas for automatic elasticity management that this algorithm obtained five feasible outcomes at six observations compared to others.

And this algorithm has had low intrusiveness for AutoElastic when there is no configuration. But the output with AutoElastic manager gets 26 percent of OpenNebula middleware. Michael Pantazoglou et al.[15] researched the decentralized strategy to virtual machine leadership that is scalable and energy effective and that is supplied to big, company clouds. An goal of this article to overcome the intrinsic limitations of centralized or hierarchical models in scalability and efficiency. And the algorithm used decentralized workload management that was energy effective. The decentralized strategy is to load balancing is very scalable and effective here. Live VM migrations with turning on and off compute nodes are becoming increasingly expensive. It is also very elastic, power efficient and scalable when it comes to greater workload. Xiangming Dai et al.[16] studied reducing power usage rates in data centers with the objective of brilliantly putting VM on information center servers and the suggested method is the minimum power virtual machine (VM) scheduling algorithm (MinES) and the minimum virtual machine communication scheduling algorithm (MinCS). It shows that MinES and MiCS have also accomplished excellent energy savings for all desirable systems while maintaining SLA tenants. This can be almost susceptible to error in anticipating contract deadlines so it is necessary to manage the contract deadline dates of the tenant better. The projected income is within 4.3 to 6.1% of the power consumption. Haikun Liu et al.[17] discussed the issue of resource sharing with various tenants and suggested scheme is F2C support association in cloud users to purchase resource amount and share the type of VMs. This algorithm fulfills better efficiency in order to ensure and fit. Present Iaas models close 2.2X and above resource sharing about insignificant runtime. The algorithm gains application performance with various tenants at about 45% and 92% of economically appropriate. Jing Bi et al.[18] studied the main objective of minimizing the total amount of Virtual Machines allocated to a multitier Ambient Assisted Living application and the main objective of this algorithm to improve resource provision while meeting the time limit for applications. The proposed heuristic allocation algorithm for VM (HVMA) is not required for mathematical source, but is more important for other algorithms. It is very difficult to get the job's complexion out of optimization. It can demonstrate that Virtual Machines can be dynamically allocated by hybrid multi-tier queuing model while meeting the time limit for returning applications. Their study will be expanded to examine different heterogeneous apps This paper's research by Hao Wu et al.[19] demonstrates the minimization of resource and VM instance time and the deployment of cloud computing application deadlines. The proposed name of the algorithm is Minimal Slack Time and Minimal Distance (Msmd) Based Scheduling with guaranteed apps and less resource limit compared to other resources. The model is based on the execution times of the tasks, but in the cloud environment it is difficult to predict exactly. Based on the allocated funds, the job can decrease the application average by 3.4 percent of deadlines. This is why the submitted algorithm will be enhanced to study the virtualization overhead in the cloud with tasks implementation time. Qi Qi et al.[20] addressed the question of optimization, which examines price, time sharing and reliability. Particle Swarm Optimization (PSO)-based algorithm with it receives the accurate optimal solution with all the cases examined can achieve nearly 90 percent of the solution in sufficient time and will study security procedures using the suggested resources with multiple home cloud applications. The resource allocation task suggested by Tianyi Chen et al.[21] is described as a solid issue of optimization with minimizing the smallest net price. The proposed Robust Workload and Energy Management system with the algorithm is extensive numerical results and a corroborated approach to initial information. There is a doubt about the ongoing operation of DCs. Mahyar Movahed Nejad et al.[22] discussed designing a reliable impatient system to fix the multi-resource Virtual Machine supply and allocation issue. The proposed algorithm is greedy mechanisms capturing the dynamic demand of the market efficiently, providing the resources to make it and producing the high profit. It shows that greedy mechanism execution time is limited. For cloud provider, this mechanism is achieved high revenue for the cloud providers and it implements a prototype allocation system in cloud computing with experimental solution and examine the performance of that mechanism. Xingwei Wang et al.[23] researched a smart financial strategy for bidding on dynamic resource allocation (IEDA) and asked for a sensible auction round and a full qualified transaction relationship with suppliers and clients. The proposed algorithm is a cost prediction algorithm based on the neural back propagation network (BPNN) and a price matching algorithm, which is superior to market-level economic efficiency and confidence. Encouraging the participant, not only useful in minimizing market excess and strength, but also included in encouraging respondents to be frank and implementing a prototype system strategy.

IV. TABLE WORK

AUTHOR	TECHNIQUES	ADVANTAGE	DISADVANTAGE
Hadi Goudarzi et al.	hierarchical resource management	Proposed algorithm is very scalable and higher performance	There is a smaller number of migration because of resource limits
Xiaolong Xu et al.	EnReal: An Energy-Aware Resource Allocation Algorithm	Enreal method is effective for cloud	Here the pm mode of the switch also causes the energy consumption.

Hongbing Wang et al.	Cloud Resource Allocating Algorithm via Fitness-enabled Auction (CRAA/FA)	It is more efficient for cloud provider	That is impossible to experiment because of limitation of research funding.
Dan Li et al.	Dcloud	Dcloud can accept more job than the existing solution	DCloud's resource allocation algorithm will have limited advantage in the efficient use of the cloud resource
Lena Mashayekhy et al.	auction-based online mechanism for VM provisioning, allocation, and pricing in clouds	It helps for users to reveal their requests helping a healthy competition among users	This method will investigate new online mechanisms in cloud settings.
Rodrigo da et al.	auto elastic paas based model for automatic elasticity management	This could achieve five possible results at six observation	Proposed system got low intrusiveness
Michael Pantazoglou et al.	decentralized and energy efficient workload management	Decentralized management is very scalable and efficient.	It has been increasing cost of live VM migrations.
Xiangming Dai et al.	minimum energy virtual machine (VM) scheduling algorithm (MinES) and minimum communication virtual machine scheduling algorithm (MinCS)	Proposed method can achieved a great energy saving	This method needed for better management of tenants contract
Haikun Liu et al.	F2C	It satisfies better performance and suitability.	Presented IAAS model insignificant of runtime resource sharing.
Jing Bi et al.	heuristic VM allocation algorithm (HVMA)	HVMA is not needed for mathematical source.	Discussed method difficult to get the complexion of the work out optimization
Hao Wu et al.	Minimal Slack Time And Minimal Distance (Msmd) Based Scheduling	MSND can guarantee applications and less resources compared others.	Tough to exactly predict the task execution time in cloud.
Qi Qi et al.	Particle Swarm Optimization (PSO)-based algorithm	The algorithm obtains the exact optimal solution.	It will investigate security mechanism utilize suggest resources at home to home clouds with multiple application.
Tianyi Chen et al.	Robust Workload And Energy Management	Method has obtained comprehensive numerical results	There is a doubt involved in operating continuous DCs
Mahyar Movahed Nejad et al.	greedy mechanisms	Greedy mechanisms is very effectively and generating the high profit	But the execution time is limited.
Xingwei Wang et al.	Back propagation neural network (BPNN) based price prediction algorithm and a price matching algorithm.	It's superiority in economic efficiency and confidence.	That algorithm will implement approach in a prototype system.

TABLE WORK

V. CONCLUSION

In this paper we have discussed about resource allocation on cloud computing, there is lack of research that address challenges of managing multiple resources with objective of allocating enough resources for each work load to optimizing power consumption. This paper has surveyed various types of resource allocation algorithms that improve the cloud Infrastructure. The proposed system can achieve good performance, efficient, accepts more jobs, helps to participate to be honest, possible results and compared to others that is the guaranteed application etc,. There are some of drawbacks also included that are resources limits, modes causes energy consumption, impossible to experiment results due to research funds, live migration, and execution time limits ect., Therefore it enhances for future work to method investigate new online mechanisms in cloud settings, implement in open source OS and prototype systems.

REFERENCES

1. William Voorsluys, James Broberg, And Rajkumar Buyya, "Cloud Computing Principles and Paradigms", John Wiley & Sons, Inc., 2011.
2. Aswathy B Namboothiri1, Dr. R. Joshua Samuel Raj 2, "A Comparative Study on Job Scheduling Algorithm Augmenting Load Balancing in Cloud" Second International Conference on Science Technology Engineering And Management(ICONSTEM), pp.582-588, 2016.
3. Kshitiza Vasudewa, Punit Gupta, "A Survey on Elastic Resource Allocation Algorithm for Cloud Infrastructure", 1st International Conference on Innovation and Challenges in Cyber Security, pp.199-203, 2016.

4. J. van Bon and A. van der Veen, "Foundations of IT Service Management based on ITIL", Vol. 3, Van Haren Publishing, Zaltbommel, September 2007.
5. R. Buyya, C. S. Yeo, and S Venugopal, "Market-oriented cloud computing: Vision, hype, and reality for delivering IT services as computing utilities", in Proceedings of the 10th IEEE International Conference on High Performance Computing and Communications, September 25-27, 2008.
6. Ahmed Shawish and Maria Salama, "Cloud Computing: Paradigms and Technologies" Springer, pp.67-39, 2014.
7. Chander Diwaker, Priyanka Dembla, "Survey on Cloud Computing" International Journal of Advanced Research in Computer Science, vol 5, no. 5, May-June 2014.
8. Brendan Jennings Rolf Stadler, "Resource Management in Clouds: Survey and Research Challenges", Springer J Netw Syst Manage, no.23, pp. 567-619, March 2014.
9. Hadi Goudarzi and Massoud Pedram, "Hierarchical SLA-Driven Resource Management for Peak Power-Aware and Energy-Efficient Operation of a Cloud Datacenter", IEEE Transactions On Cloud Computing, vol.4 no.2, pp. 222-236, April-June 2016.
10. Xiaolong Xu, Wanchun Dou, Xuyun Zhang, and Jinjun Chen, "EnReal: An Energy-Aware Resource Allocation Method for Scientific Workflow Executions in Cloud Environment", IEEE Transactions On Cloud Computing, vol. 4, no. 2, pp. 166-179, April-June 2016.
11. Hongbing Wang, Zuling Kang, and Lei Wang, "Performance-Aware Cloud Resource Allocation via Fitness-Enabled Auction", IEEE Transactions On Parallel And Distributed Systems, vol. 27, no. 4, pp. 1160-1173, Apr-16.
12. Dan Li, Congjie Chen, Junjie Guan, Ying Zhang, Jing Zhu, and Ruozhou Yu, "DCloud: Deadline-Aware Resource Allocation for Cloud Computing Jobs", IEEE Transactions On Parallel And Distributed Systems, vol. 27, no. 8, pp. 2248-2260, Aug-16.
13. Lena Mashayekhy, Mahyar Movahed Nejad, Daniel Grosu, and Athanasios V. Vasilakos, "An Online Mechanism for Resource Allocation and Pricing in Clouds", IEEE Transactions On Computers, vol. 65, no. 4, pp. 1172-1184, April 2016.
14. Rodrigo da Rosa Righi, Vinicius Facco Rodrigues, Cristiano Andr e da Costa, Guilherme Galante, Luis Carlos Erpen de Bona, and Tiago Ferreto, "AutoElastic: Automatic Resource Elasticity for High Performance Applications in the Cloud", IEEE Transactions On Cloud Computing, vol. 4, no. 1, pp. 06 to 19, January-March 2016.
15. Michael Pantazoglou, Gavriil Tzortzakis, and Alex Delis, "Decentralized and Energy-Efficient Workload Management in Enterprise Clouds", IEEE Transactions On Cloud Computing, vol. 4, no. 2, pp. 196-209, April-June 2016.
16. Xiangming Dai, Jason Min Wang, and Brahim Bensaou, "Energy-Efficient Virtual Machines Scheduling in Multi-Tenant Data Centers", IEEE Transactions On Cloud Computing , vol. 4, NO. 2, pp. 210-221, April-June 2016.
17. Haikun Liu, and Bingsheng He, "F2C: Enabling Fair and Fine-Grained Resource Sharing in Multi-Tenant IaaS Clouds", IEEE Transactions On Parallel And Distributed Systems, vol. 27, no. 9, pp. 2589-2602, Sep-16.
18. Jing Bi 1, 2, Haitao Yuan 3*, Ming Tie 4, 5, Xiao Song6, "Heuristic Virtual Machine Allocation for Multi-Tier Ambient Assisted Living Applications in a Cloud Data Center", China Communications, pp. 56-65, May 2016.
19. Hao Wu, IEEE, Xiayu Hua, Zheng Li, and Shangping Ren, "Resource and Instance Hour Minimization for Deadline Constrained DAG Applications Using Computer Clouds", IEEE Transactions On Parallel And Distributed Systems, vol. 27, no. 3, pp. 885-899, Mar-16.
20. Qi Qi, Jingyu Wang, Qi Li, Tonghong Li, and Yufei Cao, "Resource Orchestration for Multi-task Application in Home-to-home Cloud ", IEEE Transactions on Consumer Electronics, vol. 62, no. 2, pp. 191-199, May-16.
21. Tianyi Chen, Yu Zhang, Xin Wang, and Georgios B. Giannakis, "Robust Workload and Energy Management for Sustainable Data Centers ", IEEE Journal On Selected Areas In Communications, vol. 34, no. 3, pp. 651-664, March 2016.
22. Mahyar Movahed Nejad, Lena Mashayekhy, Daniel Grosu, "Truthful Greedy Mechanisms for Dynamic Virtual Machine Provisioning and Allocation in Clouds", IEEE Transactions On Parallel And Distributed Systems, vol. 26, no. 2, pp. 594-603, Feb-15.
23. Xingwei Wang, Xueyi Wang, Hao Che, Min Huang, and Chengxi Gao, "An Intelligent Economic Approach for Dynamic Resource Allocation in Cloud Services", IEEE Transactions On Cloud Computing, vol. 3, no. 3, pp. 275-289, July-September 2015 .

AUTHORS PROFILE

Deepa R received B.sc, (Information System Management) 2011, M.sc., (Computer Science) 2014 and studying M.Tech (By Research) in computer Science and Engineering at Vellore Institute of Technology, Vellore, Tamil Nadu, India. She is presently working as Teaching cum Research Assistant in School of Computer Science and Engineering at VIT University. Her research interests include Cloud Computing, Internet of Things and Machine Learning approaches (deepa.ruth91@gmail.com).

Dheeba J received her B.E (2005), M.E (2008) and Ph.D (2013) in Computer science and Engineering from Anna University, Chennai. She is presently working as Associate Professor in School of Computer Science and Engineering at Vellore Institute of Technology, Vellore. Her research interests include medical image analysis, Machine learning approaches and big data analytics.