International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

An AVL Tree Based Algorithm for Virtual Machine Placement Problem

Rose Rani John, E. Grace Mary Kanaga, Sandeep S. G.

Submitted: 13/10/2023 **Revised**: 12/12/2023 **Accepted**: 23/12/2023

Abstract—Virtual machine placement problem is an np-hard problem which plays a major role in providing services over cloud with optimal service level agreement violations and min- imizing the power consumption in data centers. Approximation algorithms, evolutionary algorithms and machine learning based approaches are available in the literature. Best Fit algorithm is one of the best approximation algorithms for virtual machine placement problem. However, the time complexity of best fit algorithm can be reduced to O(logn) by using a self-balancing binary search tree such as AVL Tree. This paper proposes a modified best-fit algorithm by using AVL Tree data structure and analyses performance of that approach in virtual machine placement problem. AVL Tree based algorithm increases the performance in terms of time complexity as the search, insert and delete operations guarantees O(logn) time. Tested in a homogeneous host environment, the AVL Tree based algorithm gives on an average 0.4% better performance than the Next Fit algorithm and its variants.

Index Terms—*Cloud Computing, Virtual machine placement, approximation algorithms, AVL tree, Best fit algorithm*

Introduction

Cloud computing has emerged as a transformative technology that revolutionizes the way businesses and individuals store, process, and access data and applications. It offers unprecedented flexibility, scalability, and cost-effectiveness, empowering users to access a wide range of computing capabilities without the need for extensive on-premises infrastructure. Cloud computing is defined as the delivery of on-demand computing services over the internet, encompassing various resources such as servers, storage, databases, and software. [4] describe cloud computing as ”a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources that can be rapidly provisioned and released with minimal management effort.”

Cloud computing offers various deployment models, each tailored to specific needs and requirements. The most common models include public, private, hybrid, and community clouds. [5] describe these deployment models, highlighting their distinct characteristics, advantages, and challenges. Virtualization, distributed computing, and network technologies play crucial roles in realizing the benefits of cloud computing. [6] provide a comprehensive overview of the technologies underpinning cloud computing, emphasizing their contributions to resource management, scalability, and fault tolerance.

Virtualization is a fundamental technology that underpins cloud computing, allowing for the efficient utilization and management of computing resources. Virtualization involves the abstraction and encapsulation of physical computing resources, such as servers, storage, and networks, into virtual entities that can be dynamically allocated and managed. [7] discuss the core concepts and principles of virtualization, including hardware virtualization, operating system-level virtualization, and application-level virtualization. They emphasize the role of virtualization in decoupling software and applications from the underlying hardware, enabling greater flexibility and resource efficiency.

Dept. of CSE KITS
Coimbatore, India roseranijohn@gmail.com
Dept. of CSE KITS
Coimbatore, India grace@karunya.edu
Coimbatore, India sandeepsg@karunya.edu.in

By consolidating multiple virtual machines (VMs) onto a single physical server, virtualization enables better resource utilization, reducing the need for extensive physical infrastructure. Virtualization plays a significant role in optimizing resource utilization and reducing costs in cloud computing environments. [8] discuss the impact of virtualization on resource allocation, workload consolidation, and energy efficiency. They highlight the ability of virtualization to dynamically allocate and consolidate resources based on workload demands, leading to improved resource utilization and cost savings. Virtual machine (VM) consolidation is a crucial technique in cloud computing for optimizing resource utilization and improving energy efficiency.

VM consolidation involves the efficient allocation and placement of multiple VMs onto a reduced number of physical servers. By consolidating VMs, organizations can achieve better resource utilization, improved performance, and reduced energy consumption. [9] discuss the core concepts and principles of VM consolidation, including workload characterization, resource allocation policies, and migration techniques. They emphasize the importance of workload analysis and optimization algorithms to achieve efficient VM consolidation. Various techniques and algorithms have been proposed for efficient VM consolidation, such as utilization- based consolidation, load balancing, and predictive modeling. [1], [2], [3] [10] discuss these techniques, highlighting their benefits, limitations, and applicability in different cloud computing scenarios. They emphasize the importance of considering workload characteristics and system constraints when selecting an appropriate VM consolidation technique. [25], [27], [28], [29] elaborates more on the energy efficiency concerns in the context of VM consolidation.

Virtual machine (VM) placement is a critical problem in cloud computing that involves determining the optimal allocation of VMs onto physical servers. By strategically placing VMs onto physical servers, organizations can achieve improved resource utilization, reduced energy consumption, enhanced fault tolerance, and better workload distribution. [13] discuss the significance of VM placement, emphasizing its role in meeting service-level agreements (SLAs), minimizing latency, and optimizing resource allocation. Physical servers in cloud environments often have different configurations and capacities, making it challenging to allocate VMs optimally. Workload demands can vary over time, requiring the VM placement algorithm to adapt to changing conditions and effectively handle workload spikes or fluctuations.

Various optimization techniques and algorithms have been proposed to address the VM placement problem, such as approximation algorithms, integer linear programming, genetic algorithms, ant colony optimization, and machine learning-based approaches. [11] discuss these optimization techniques, highlighting their strengths, limitations, and applicability in different cloud computing scenarios. They emphasize the importance of considering system constraints, workload characteristics, and performance objectives when selecting an appropriate VM placement approach. To assess the effectiveness of VM placement algorithms, several evaluation metrics are commonly used, including resource utilization, energy consumption, migration cost, and SLA violations. [12], [26] discuss these evaluation metrics, providing insights into their significance and the trade-offs involved in VM placement optimization.

Related Works

The Virtual Machine Placement Problem (VMPP) is a challenging optimization problem in cloud computing that involves determining the optimal assignment of virtual machines (VMs) to physical servers. *Approximation Algorithms for VM Placement* There are many approximation algorithms which can be used to find solution for virtual machine placement problem. [30], [31] discuss various possibilities in vm placement. Some of the major approaches are listed below.

A. *First Fit Algorithm*

The First Fit algorithm follows a straightforward principle for assigning virtual machines (VMs) to physical servers. It starts by examining the available servers in the order they are encountered and assigns a VM to the first server with sufficient capacity to accommodate it. This approach aims to maximize resource utilization and minimize fragmentation. [20] discuss the principles of the

First Fit algorithm, emphasizing its simplicity and effectiveness in VM placement. The algorithm's simplicity allows for fast and straightforward VM placement, making it a practical choice for real-time deployments. By assigning VMs to the first available server, the algorithm ensures rapid resource allocation without the need for extensive searching or sorting operations. The First Fit algorithm is scalable and can efficiently handle large-scale VM placement scenarios. But this approach may reduce overall resource utilization. If servers are not uniformly provisioned or the VM order is not optimized, the First Fit algorithm may result in load imbalance across servers.

B. *First-Fit Decreasing (FFD) Algorithm*

The FFD algorithm is a well-known approximation algorithm for VM placement. The First Fit Decreasing algorithm combines the principles of the First Fit algorithm with a sorting technique. It sorts the virtual machines (VMs) in descending order of their sizes before assigning them to physical servers. The algorithm assigns each VM to the first server with sufficient capacity, aiming to minimize fragmentation and improve resource utilization.

[21] discuss the principles of the First Fit Decreasing algorithm, highlighting its effectiveness in VM placement.

C. *Best Fit Algorithm*

The Best Fit algorithm is based on the principle of selecting the physical server that has the least available capacity to accommodate a virtual machine (VM). It aims to minimize resource fragmentation and maximize resource utilization by placing VMs in the most suitable servers. [19] discuss the principles of the Best Fit algorithm, emphasizing its efficiency and effectiveness in VM placement. By selecting the server with the least available capacity, the Best Fit algorithm maximizes resource utilization and minimizes resource fragmentation. The algorithm tends to distribute VMs evenly across physical servers, achieving load balancing and preventing resource overloading. The Best Fit algorithm is scalable and can handle large-scale VM placement scenarios efficiently. However, compared to simpler algorithms like First Fit, the Best Fit algorithm requires more computational resources due to the iterative placement process.

D. *Best-Fit Decreasing (BFD) Algorithm*

The BFD algorithm is an improvement over the FFD algorithm. It sorts the VMs in descending order and selects the server with the least available capacity to accommodate each VM. BFD achieves better packing efficiency than FFD but requires higher computational complexity. [14] discuss the principles of the Best Fit Decreasing algorithm, highlighting its effectiveness in VM placement. BFD algorithm requires additional computational resources due to the sorting and iterative placement steps.

E. *Next Fit Algorithm*

The Next Fit algorithm follows a simple principle for assigning virtual machines (VMs) to physical servers. It sequentially scans the available servers and assigns a VM to the next server with sufficient capacity to accommodate it. This approach aims to maximize resource utilization and minimize fragmentation.

F. *Next-Fit Decreasing (NFD) Algorithm*

The NFD algorithm is a variation of the FFD algorithm that assigns a VM to the next server if the current server cannot accommodate it. It reduces the fragmentation of server resources but may lead to suboptimal packing. The Next Fit Decreasing algorithm combines the principles of the Next Fit algorithm with a sorting technique. The algorithm first sorts the virtual machines (VMs) in descending order of their sizes and assigns them to physical servers. It assigns each VM to the next available server with sufficient capacity, aiming to balance efficiency and resource utilization. [22]

, [23] discusses the principles of the Next Fit Decreasing algorithm, highlighting its effectiveness in VM placement.

G. *Worst-Fit Algorithm*

The Worst Fit algorithm searches for the server with the maximum available capacity and assigns the VM to it. This approach aims to maximize resource utilization by allocating VMs to servers with the most remaining resources.

H. *Worst-Fit Decreasing Algorithm*

The Worst Fit Decreasing algorithm builds upon the Worst Fit algorithm by considering a decreasing order of VM sizes. It follows the principle of assigning VMs to servers with the maximum available capacity, while also considering the decreasing order of VM sizes. This approach aims to maximize resource utilization and improve load balancing. In this algorithm, sorting VMs in decreasing order of size adds an additional computational overhead during placement, potentially impacting placement time.

I. PROPOSED WORK

Among the existing approaches in the category of approximation algorithms for virtual machine placement, Best fit algorithm is considered as the best approach for an online problem scenario and Best fit decreasing is considered as the best approach in an offline scenario. Cloudsim toolkit make use of a power aware version of these algorithms for virtual machine placement. However, these algorithms lack in time complexity. The performance of the algorithm can be improved in terms of time complexity by using self-balancing binary search tree. So this paper proposes a modified Best fit algorithm by using AVL Tree.

The AVL Tree is a self-balancing binary search tree that maintains efficient data structure operations by ensuring balanced height. The AVL Tree is a fundamental data structure that provides efficient searching, insertion, and deletion operations with a guaranteed time complexity of O(log n) while maintaining a balanced height. It employs the concept of self-balancing through rotation operations to ensure that the difference in height between its left and right subtrees remains within a specified range. [**?**] first introduced and described the principles of the AVL Tree, highlighting its self-balancing property. The AVL Tree has a binary tree structure where each node contains a key and pointers to its left and right child nodes. In this approach, the available host pool is maintained as an AVL Tree. And the tree is updated as virtual machines are allocated to the hosts. Algorithm for virtual machine allocation using AVL Tree based best fit algorithm is given below. A Hashmap is used to keep track of the virtual machine allocations.

As explained in the algorithm, initially a host pool is created using the AVLTree. Whenever a new virtual machine comes and it needs to be allocated, the AVLHostPool is searched to find the best suited host to accommodate the incoming virtual machine. The search function will return a host that has least available CPU utilization to satisfy the CPU requirement of the virtual machine. Since AVLHostPool is a binary search tree, the search function return the best possible host to accommodate the virtual machine. During the virtual machine allocation process, entries are made into hashmap called VMStore which will keep track of the virtual machine allocations. Taking the count of the unique servers in this hashmap will give the number of active hosts used in this allocation. Also, VMStore will give the active alloation of the virtual machines to the hosts. Once a de-allocation happens, that particular entry is removed from the VMStore as well.

Algorithm 1: VM Placement Algorithm using AVL Tree based Best Fit Algorithm

---

Data: List of available hosts with CPU Utilization, List of VMs to be allocated with CPU requirements.
Result: Number of hosts used, VM allocation
Set up and initialize the available server pool called “AVLHostPool” using AVLTree. ;
When a new virtual machine needs to be allocated to a host
Search the AVLHostPool ;
if *best fit host node is found* **then** bestNodeCpuUtilization -= vmCpuUtilization ; delete the bestNode ;
re-insert the bestNode with updated CpuUtilization ;
update the vm allocation details in VMStore hashmap ;

III and type IV virtual machines. On an average the proposed algorithm is approximately 0.5% better than the Next Fit and Next Fit Decreasing algorithms. Figure 2 shows the performance of the approximation algorithms when G5 servers with 2660MHz is used with four variants of virtual machines as mentioned in table 2.

else
no suitable server found;
end
activeHosts=count of the unique hosts in the VMStore return activeHosts, VMStore

Experimental Setup

The simulation of the algorithm is performed on Cloudsim and with the use of PlanetLab dataset. The simulation is performed on a homogeneous host environment. The server configuration used is given in table 1. In each experiment, around 800 hosts of same type are used.

1	HP ProLiant ML110 G4 Intel Xeon 3040, 2 Cores x 1860MHz, 4GB
2	HP ProLiant ML110 G5 Intel Xeon 3075, 2 Cores x 2660MHz, 4GB

Table I
THE HOST CONFIGURATIONS

Four different types of virtual machines are considered for this simulation and their configurations are listed in Table II.

Type	CPU(MIPS)	Cores	Memory
I	2.5 GHz	1	0.85 GB
II	2.0 GHz	1	3.75 GB
III	1.0 GHz	1	1.7 GB
IV	0.5 GHz	1	613

Table II
THE VIRTUAL MACHINE TYPES

Result Analysis

The performance of the AVLTree based Best fit algorithm is compared with the other existing approximation algorithms for virtual machine placement in a homogeneous host environment. Figure 1 shows the performance of the algorithms when G4 servers with 1860MHz with type

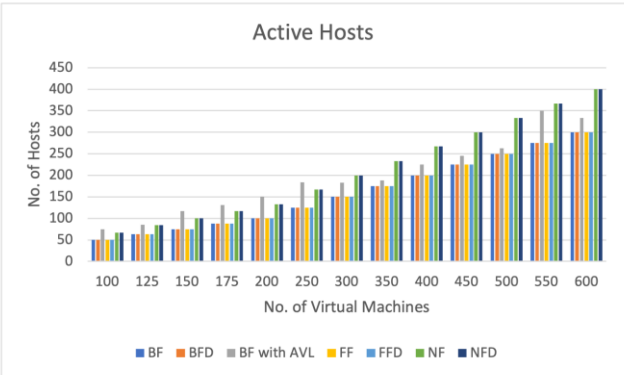
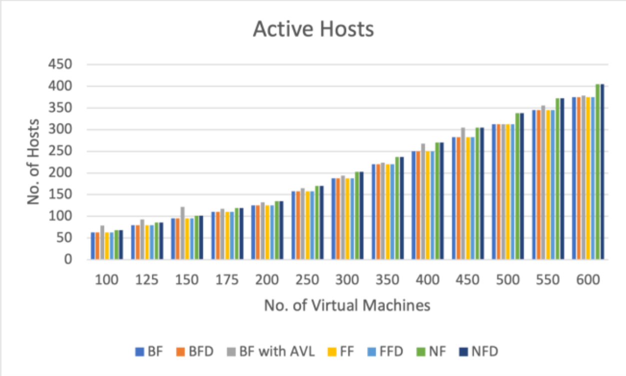


Fig. 1. No. of Active Hosts when the G4 servers are in use

In figure1 algorithms are compared in an environment where homogeneous servers are used. In this environment the proposed algorithms seems to outperform the Next Fit algorithm and its variants.

According to figure 2 the best fit algorithm using AVL Tree gives 0.3% better results than the Next Fit and Next Fit Decreasing algorithms



.

Fig. 2. No. of Active Hosts when the G5 servers are in use

Conclusion

In this paper a self balancing binary search tree based best fit algorithm is proposed and implemented using AVL Tree and the performance of the algorithm is compared with the other existing approximation algorithm with standard benchmark data. The algorithm guarantees $O(logn)$ time complexity for search, insert and delete operations which are required while performing virtual machine placement using this algorithm. The proposed algorithm also gives on anaverage 0.4% better performance than the Next Fit algorithm and its variants. However, due to the need for more insertion operations, more rotations might need to be performed on the tree and that could be a costly segment of this algorithm. So this algorithm can be improved further by using some other strategy by which the number of rotations need to be performed on the tree could be minimized.

References

[1] Beloglazov, A., Abawajy, J., Buyya, R. (2012). Energy-aware resource allocation heuristics for efficient management of data centers for cloud computing. Future generation computer systems, 28(5), 755-768.

[2] Beloglazov, A., Buyya, R. (2012). Optimal online deterministic algo- rithms and adaptive heuristics for energy and performance efficient dynamic consolidation of virtual machines in cloud data centers. Con- currency and Computation: Practice and Experience, 24(13), 1397-1420.

[3] Beloglazov, A., Buyya, R. (2012). Managing overloaded hosts for dynamic consolidation of virtual machines in cloud data centers under quality of service constraints. IEEE transactions on parallel and dis- tributed systems, 24(7), 1366-1379.

[4] Mell, P., Grance, T. (2011). The NIST definition of cloud computing. Communications of the ACM, 53(6), 50-56.

[5] Goyal, P., Mikkilineni, R. (2009, September). Policy-based event-driven services-oriented architecture for cloud services operation and manage- ment. In 2009 IEEE International Conference on Cloud Computing (pp. 135-138). IEEE.

[6] Armbrust, M., Fox, A., Griffith, R., Joseph, A. D., Katz, R., Konwinski, A., ... Zaharia, M. (2010). A view of cloud computing. Communications of the ACM, 53(4), 50-58.

[7] Smith, J. E., Nair, R., Shetty, S. (2005). The architecture of virtual machines. Computer, 38(5), 32-38.

[8] Sabahi, F. (2012). Secure virtualization for cloud environment using hypervisor-based technology. International Journal of Machine Learning and Computing, 2(1), 39.

[9] Ahmad, R. W., Gani, A., Hamid, S. H. A., Shiraz, M., Yousafzai, A., Xia, F. (2015). A survey on virtual machine migration and server consolidation frameworks for cloud data centers. Journal of network and computer applications, 52, 11-25.

[10] Beloglazov, A., Abawajy, J., Buyya, R. (2012). Energy-aware resource allocation heuristics for efficient management of data centers for cloud computing. Future Generation Computer Systems, 28(5), 755-768.

[11] Usmani, Z., Singh, S. (2016). A survey of virtual machine placement techniques in a cloud data center. Procedia Computer Science, 78, 491- 498.

[12] Masdari, M., Nabavi, S. S., Ahmadi, V. (2016). An

overview of virtual machine placement schemes in cloud computing. Journal of Network and Computer Applications, 66, 106-127.

[13] Choudhary, A., Rana, S., Matahai, K. J. (2016). A critical analysis of energy efficient virtual machine placement techniques and its optimiza- tion in a cloud computing environment. Procedia Computer Science, 78, 132-138.

[14] Masdari, M., Gharehpasha, S., Ghobaei-Arani, M., Ghasemi, V. (2020). Bio-inspired virtual machine placement schemes in cloud computing environment: taxonomy, review, and future research directions. Cluster Computing, 23(4), 2533-2563.

[15] Karmakar, K., Banerjee, S., Das, R. K., Khatua, S. (2022). Utilization aware and network I/O intensive virtual machine placement policies for cloud data center. Journal of Network and Computer Applications, 205, 103442.

[16] Singh, A. K., Swain, S. R., Lee, C. N. (2023). A metaheuristic virtual machine placement framework toward power efficiency of sustainable cloud environment. Soft Computing, 27(7), 3817-3828.

[17] Laghrissi, A., Taleb, T. (2018). A survey on the placement of virtual resources and virtual network functions. IEEE Communications Surveys and Tutorials, 21(2), 1409-1434.

[18] Infantia Henry, N., Anbuananth, C., Kalarani, S. (2022). Hybrid meta- heuristic algorithm for optimal virtual machine placement and migration in cloud computing. Concurrency and Computation: Practice and Expe- rience, 34(28), e7353.

[19] Kuo, J. J., Yang, H. H., Tsai, M. J. (2014, April). Optimal approximation algorithm of virtual machine placement for data latency minimization in cloud systems. In IEEE INFOCOM 2014-IEEE Conference on Computer Communications (pp. 1303-1311). IEEE.

[20] Kim, K. H., Beloglazov, A., Buyya, R. (2011). Power-aware provisioning of virtual machines for real-time Cloud services. Concurrency and Computation: Practice and Experience, 23(13), 1491-1505.

[21] Alahmadi, A., Alnowiser, A., Zhu, M. M., Che, D., Ghodous, P. (2014, March). Enhanced first-fit decreasing algorithm for energy-aware job scheduling in cloud. In 2014 International Conference on Computational Science and Computational Intelligence (Vol. 2, pp. 69-74). IEEE.

[22] Kumaraswamy, S., Nair, M. K. (2019). Bin packing algorithms for virtual machine placement in cloud computing: a review. International Journal of Electrical and Computer Engineering, 9(1), 512.

[23] Nashaat, H., Ashry, N., Rizk, R. (2019). Smart elastic scheduling algorithm for virtual machine migration in cloud computing. The Journal of Supercomputing, 75, 3842-3865.

[24] Adelson-Velskii, Georgii Maksimovich, and Evgenii Mikhailovich Lan- dis. An algorithm for organization of information. In Doklady Akademii Nauk, vol. 146, no. 2, pp. 263-266. Russian Academy of Sciences, 1962.

[25] Q. Zhou, M. Xu, S. Singh Gill, C. Gao, W. Tian, C. Xu, R. Buyya, Energy Efficient Algorithms based on VM Consolidation for Cloud Computing: Comparisons and Evaluations, Proc. - 20th IEEE/ACM Int. Symp. Clust. Cloud Internet Comput. CCGRID 2020. (2020) 489–498.

[26] R. Zolfaghari, A.M. Rahmani, Virtual Machine Consolidation in Cloud Computing Systems: Challenges and Future Trends, Springer US, 2020.

[27] B. Prabha, K. Ramesh, P.N. Renjith, A Review on Dynamic Virtual Machine Consolidation Approaches for Energy-Efficient Cloud Data Centers, in: Springer, Singapore, 2021: pp. 761–780.

[28] Arshad, U., Aleem, M., Srivastava, G., Lin, J. C. W. (2022). Utilizing power consumption and SLA violations using dynamic VM consolida- tion in cloud data centers. Renewable and Sustainable Energy Reviews, 167, 112782.

[29] Magotra, B., Malhotra, D., Dogra, A. K. (2023). Adaptive computational solutions to energy efficiency in cloud computing environment using VM consolidation. Archives of Computational Methods in Engineering, 30(3), 1789-1818.

[30] G. Zhao, J. Liu, Y. Zhai, H. Xu and H. He, ”Alleviating the Impact of Abnormal Events Through Multi-Constrained VM Placement,” in IEEE Transactions on Parallel and Distributed Systems, vol. 34, no. 5, pp. 1508-1523, May 2023, doi: 10.1109/TPDS.2023.3248681.

[31] Guo, L., Lu, C., Wu, G. (2023). Approximation algorithms for a virtual machine allocation problem with finite types. Information Processing Letters, 180, 106339.