

# Virtual Machine Placement in Cloud Computing

Arunkumar Kumar, Ciddhesh Sathasivam\* and Prakash Periyasamy

Department of Computer Science Engineering, Amrita Vishwa Vidyapeetham, Amrita Nagar,  
Coimbatore - 641112, Tamil Nadu, India;  
arunkumark@outlook.com, ciddhesh1995@gmail.com, p\_prakash1@cb.amrita.edu

## Abstract

**Objectives:** Hardware virtualization is an evolving technology, because it has the potential to facilitate the consolidation of multiple workloads on a single physical host directly through cloud. We introduce the First-Last algorithm to minimize energy consumption and migration. **Methods:** The problem is treated as an instance of bin packing algorithm which focuses on saving energy. However the overloading of active hosts and the resultant migration are the drawbacks of this approach. This algorithm was tested on a set of heterogeneous virtual machines which have demands in both single and multi-dimension and came either statically or dynamically to the server. **Findings:** The virtual machine placement problem can be treated as a bin packing problem, wherein the physical machines/hosts are the bins and the virtual machines are the objects. It can then be solved through various approaches such as first fit, best fit, next fit et cetera. Using the first fit algorithm, the virtual machine is placed into the first active host which can accommodate it. If none of the active hosts satisfy the VM's demand, then a new host is made active. In the best fit method, all the active hosts are checked first and then the host which would suffer minimum resource wastage upon VM placement is chosen. Next fit algorithm places the VM in the last /most recently activated host. While first fit is best suited for optimal performance, best fit keeps the least number of hosts active, thereby saving energy but at the cost of performance since migration overhead increases. Next fit approach would fail unless the virtual machines have been arranged in the increasing order of their demands. The first-last algorithm that we introduce balances on energy consumption as well as performance by reducing migration. It also reduces the time complexity to a certain extent. **Applications:** The algorithm has a comparatively lower time complexity than the existing virtual machine placement algorithms. It also reduces the energy consumption as well as migration to a certain extent.

**Keywords:** CloudSim, First Last Algorithm, Hardware Virtualization, Migration, Resource Scalability

## 1. Introduction

Cloud computing is an emerging technology which is steadily evolving into a major domain of computer science. Unlike conventional local dedicated servers, cloud computing provides computing resources to multiple organizations connected to remote servers on a network. Based on the mode of deployment cloud computing can be classified into four, namely public, private, hybrid and community<sup>1</sup>. Resource scalability is the strength of public cloud. An organization whose primary goal is to provide abundant resources to its client opts for public cloud service. However, an organization

whose chief concern is security chooses private clouds although they are expensive. Hybrid cloud service has the advantages of both public and private. Various services provided by cloud computing are Software as a Service (SaaS), Platform as a Service (PaaS) and Infrastructure as a Service (IaaS)<sup>2</sup>.

In an Infrastructure as a Service model, cloud service provider hosts hardware such as memory, storage, network bandwidth, CPU time etc. These services are made available to the client through hardware virtualization. The number of clients that the provider can host is generally more than the available resources. This is possible through virtualization wherein the resources are

\* Author for correspondence

used to their maximum capacity. The client on the other hand does not have to make huge capital investment in hardware and instead pays on per-use basis. They also have the advantage of resource scalability<sup>3</sup>.

As and when hardware requests are made by the clients, corresponding virtual machines are created in the cloud and mapped on to the most appropriate host. Finding the most appropriate host is termed as the virtual machine placement problem. Often the client’s demands tend to expand over a period of time. In this situation, if the host onto which the virtual machine is placed is unable to service the rise in demand, then the virtual machine migrates to the next most appropriate host causing an overhead. Cloud service provider tries to minimize the number of hosts that are active in order to reduce energy consumption. Several algorithms such as first fit, best fit, first fit decreasing etc exist to address the mapping problem. However there is a tradeoff between minimizing energy consumption and minimizing migration when these algorithms are used. In this paper we propose an algorithm which minimizes both energy consumption

and migration overhead. We approach the problem by splitting it into three phases.

- Static, single-dimension VMs
- Dynamic, multi-dimension VMs
- VMs that require migration

In this paper we explain the virtual machine placement problem in terms of a bin packing problem. In the following sections, we introduce a novel algorithm called the First-Last algorithm to solve the same. This new algorithm can solve the mapping problem for both single dimension virtual machines requesting for hardware offline as well as multi dimension virtual machines arriving online.

## 2. Methodologies

### 2.1 System Architecture Diagram

Multiple users request for hardware to the cloud service provider[Figure 1]. These requests can be for any one

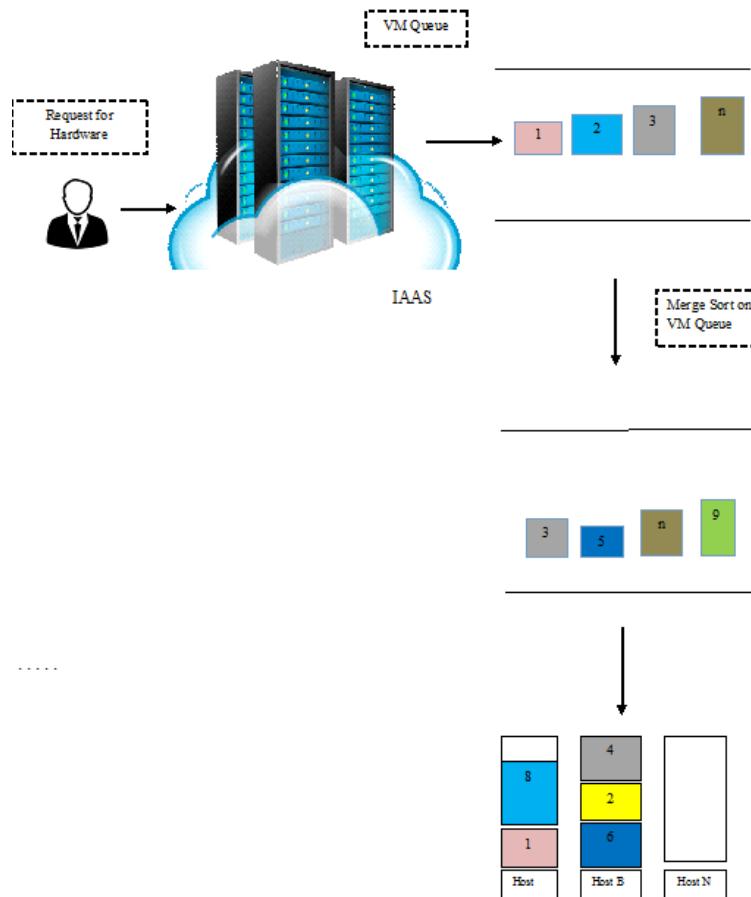


Figure 1. System Architecture Diagram.

or more of the following resources: network bandwidth, CPU time, storage, memory etc. If the request is only for one resource then it is called Single dimension else it is known as multi dimension. Multi dimension will be reduced to the case of single dimension through rank calculation for each request. This is explained in the later sections. When these user requests reach the server, corresponding virtual machines are formed and placed in a queue. The requests can be static (offline) or dynamic (online). Dynamic requests are handled as batches and thereby reduced to the static virtual machine placement. This is also explained in the following sections. Using merge sort these virtual machines are rearranged in the decreasing order of their demand/rank(in case of multi-dimension). Then using the First-Last Algorithm they are placed onto the most appropriate hosts.

### 3. FIRST-LAST ALGORITHM

#### 3.1 Case I: Offline Requests in One Dimension

In offline mode of operation, the number of virtual machines requesting for hardware is known in advance. Also these requests are only for a single resource. When a request is made to the server, the algorithm checks whether the request can be fulfilled by the physical machine. If it can be fulfilled, then a corresponding virtual machine is created and placed in the waiting queue, otherwise the request is denied. Once all the requests have been converted to virtual machines, then the queue is sorted using merge sort in descending order. Two pointers, namely FIRST and LAST are initialized to the head and the tail of the queue. The first host is activated and the first VM is placed into it. The FIRST pointer is then incremented. Then the last VM is placed into the first host[Figure 2].

(If it fits, else into a new host) LAST pointer is decremented. Whenever the VM pointed by the LAST pointer activates a new host, then the last active host is closed and pushed to the end. (This is because if the currently smallest VM could not fit in, it logically follows that the remaining unplaced VMs would not either.) This entire process is repeated till all the VMs have been placed. Using this approach, hosts get packed quickly since each host gets the largest and the smallest VM thereby optimizing the use of its resources as well. Also, as and when the host is full/ cannot accommodate future VMs, it

is pushed to the end. By doing this we reduce the number of comparisons required to choose a host. This algorithm chooses the same host as best fit but is better because it does not check all the active hosts. It simply places the VM in the first active host that can accommodate it. At all times, X% of the host's resources is kept free. This gives room for further expansion of VM leading to reduction in migration.

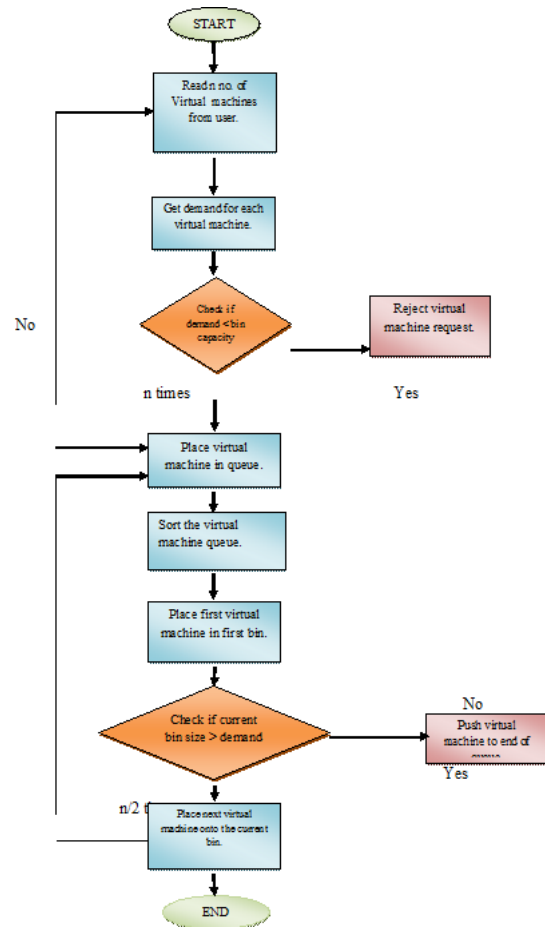


Figure 2. Flowchart of the First-Last algorithm.

#### 3.2 Case II: Online Requests in Multi Dimension

- Start
- Read N number of VMs from the user.
- Repeat step 4 N times.
- Get demand for each VM.
- Check if demand of each VM is within the bin capacity
  - Place it as a node in the queue.
  - Otherwise reject.

- Goto step 3.
- Sort using merge sort in descending order.
- Repeat step 9  $N/2$  times.
- Place the first VM from the queue onto the first bin.
- Reduce bin size
- If bin is not full, it remains open for the next VM.
- Otherwise push VM to the end of the queue and open a new bin.
- Place the last VM from the queue onto the next open/new bin.
- Reduce bin size
- If bin is not full, it remains open for the next VM.
- Otherwise push VM to the end of the queue and open a new bin.
- Goto step 8.
- End

#### 4. Example

For example, let the host size be 15 and the no. of VM requests be 10. The VM requests are as follows: 13, 9, 9, 7, 5, 15, 14, 8, 2, 1 and is then sorted in decreasing order as: 15, 14, 13, 9, 9, 8, 7, 5, 2, 1.

The VM of size 15 is placed in HOST1 whose capacity now decreases to zero, the FIRST pointer is then incremented, and the last VM of size 1 is now placed onto the new host (HOST2) where capacity of HOST2 is reduced to 14 units. The LAST pointer is decremented. The process continues till all the VMs are placed onto a host. The total number of hosts used in this scenario is 6 and no. of comparisons performed are 14 which is comparatively less than First Fit and First Fit Decreasing approach [Figure 3].

Cloud computing is the new paradigm of distributed computing which is still in its evolution phase. With the help of hardware virtualization technologies, it provides resource scalability and pay per use metered services to its clients on remote servers. Several challenges are present in this field and mapping the virtual machine to the most appropriate host is one of them. A lot of work has been done in order to solve the placement problem optimally. In the following section some of the most relevant prior works have been reviewed.

<sup>4</sup>Suggested an offline virtual machine placement method through emulated virtual machine migration and in online mode it is solved by real virtual machine migration through MBVMP algorithm. Furthermore,

they propose a hybrid scheme where a batch is employed to accept incoming virtual machine requests for the online scenario. Since the overall aim is to reduce total job completion time, if the time taken by virtual machine to migrate is too large then they adopt either direct placement or migration based placement. Their algorithm helps us to identify various ways to reduce total job completion time during migration.

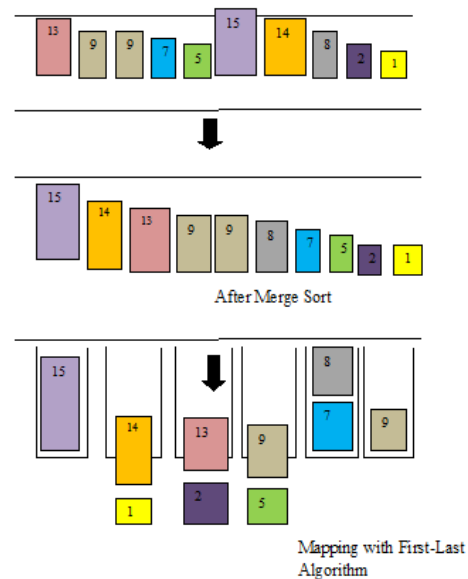


Figure 3. Example using First-Last algorithm.

In this paper<sup>5</sup>, the authors discuss about the energy cost and performance of a single virtual machine migration and dynamic virtual machine consolidation. The dynamic virtual machine consolidation is solved by using optimal online deterministic algorithm. The single virtual machine migration problem is solved by using either optimal offline algorithm or optimal online deterministic algorithm where parameter chosen is CPU performance. SLA violation occurs when virtual machine exceeds their allowed CPU request demand causing a single virtual machine to be migrated leading to decrease in CPU demand. This algorithm allows us to reduce cost considerably.

<sup>6</sup>Suggested various migration techniques towards dynamic resource management. The paper highlights the various advantages of using cloud-based services such as: flexibility in resource, resource on demand, pay-as-use model. The components identified are: when to migrate a virtual machine, where to migrate, which virtual machine to migrate etc. Migration allows load balancing,

consolidation of resources. A few challenges faced by following these techniques are: wide area migration, storage migration and also may lead to network reconfiguration. They conclude that virtualization is the backbone in providing cloud-based solution.

<sup>7</sup>Designed and implemented an automated resource management system which concentrates on overload avoidance and green computing. They introduce a concept of skewness which measures the uneven usage of servers which on decrease will improve utilization. Load prediction algorithm helps to find a pattern in resource demand by virtual machines. The result of with and without load prediction is compared, which suggest that migration is smaller with load prediction than virtual machine.

<sup>8</sup>Presented a virtual machine replication placement technique in order to reduce the overall energy consumption of the data center. The authors approach the problem by placing multiple copies of virtual machines on different servers and distributing the incoming client requests on to these virtual machines by using dynamic programming and local search methods. Thus through consolidation, servers are used to their maximum capacity. However, resource scalability is compromised since each virtual machine is given a fixed memory bandwidth for utilization.

<sup>9</sup>Focused on reducing the energy consumption cost and utilizing dynamic consolidation of Vms .They designed a new cloud technology called “Aneka” which helps to reduce the energy consumption of the datacenter. The proposed a new energy-aware resource allocation algorithm utilize the dynamic consolidation of VMs. The experimental results of the proposed algorithm shows that energy consumption is reduced significantly compared to static allocation method. They used a resource manager where Aneka and other cloud technology can interact with each other.

## 5. Conclusion

This work has shown that the bin packing problem can be effectively solved by First-Last algorithm. The First-

Last algorithms time complexity is much more efficient than the existing algorithms like Best Fit, First Fit, Best Fit Decreasing, Worst Fit etc. We address various issues in two stages: static placement and dynamic placement which is converted to a static placement scenario. The proposed evaluation method permitted virtual machine placement onto the most appropriate host, thereby decreasing energy consumption and migration to an extent. Even though migration is unavoidable, we make sure that the host has space to expand in future. This can be further extended to address issues like handling increasing demand in future, back-up storage of data or resources etc.

## 6. References

1. Shyamala K, Sunitha Rani T. An analysis on efficient resource allocation mechanisms in cloud computing. *Indian Journal of Science and Technology*. 2015 May; 9(9).
2. Veerapan G, Suresh Gnana Dhas C. Secure Resource Allocation and Control of Virtualized Datacenter using SLO. *Indian Journal of Science and Technology*. 2015 Jun; 8(12). Doi: 12.17485/ijst/2015/v8i12/61939.
3. Lavanya Suja T, Savithri V. Backtracking algorithm for virtual cluster migration in cloud computing. *Indian Journal of Science and Technology*. 2015 Jul; 8(15). Doi: 10.17485/ijst/2015/v8i15/73232.
4. Thiruvendakam T, Kamalakkannan P. Energy efficient multi dimensional host load aware algorithm for virtual machine placement and optimization in cloud environment. *Indian Journal of Science and Technology*. 2015 Aug; 8(17). Doi:10.17485/ijst/2015/v8i17/59140.
5. Types of Cloud Computing :Private, Public and Hybrid Clouds, appcore.com, Citrix Service provider, Web, Date Accessed-12/3/2015.
6. What is Cloud, ibm.com, IBM CLOUD, Web, Web, Date Accessed-08/4/2015.
7. Infrastructure as a Service (IaaS), searchcloudcomputing.techtarget.com, TechTarget, Web, Web, Date Accessed-02/3/2015.
8. Li K, Zheng H, and Wu J. Migration-based Virtual Machine Placement in Cloud Systems. 2nd International Conference on Cloud Networking (CloudNet), San Francisco, CA. 2013. p. 83 –90.
9. Beloglazov A. Energy-Efficient Management of Virtual Machines in Data Centers for Cloud Computing. *Computing and Information Systems*. 2013; pp. 44–74.