

# Chapter 6

## Dynamic Virtual Machine Placement in Cloud Computing

**Arnab Kumar Paul**  
Virginia Tech, USA

**Bibhudatta Sahoo**  
National Institute of Technology Rourkela, India

### ABSTRACT

*The aim of cloud computing is to enable users to access resources on demand. The number of users is continuously increasing. In order to fulfil their needs, we need more number of physical machines and data centers. The increase in the number of physical machines is directly proportional to the consumption of energy. This gives us one of the major challenges; minimization of energy consumption. One of the most effective ways to minimize the consumption of energy is the optimal virtual machine placement on physical machines. This chapter focuses on finding the solution to the problem of dynamic virtual machine placement for the optimized consumption of energy. An energy consumption model is built which takes into account the states of physical machines and live migration of virtual machines. On top of this, the cloud computing model is built. Unlike centralized approaches towards virtual machine placement which result in many unreachable solutions, a decentralized approach is used in this chapter which provides a list of virtual machine migrations for their optimal placement.*

### INTRODUCTION

Cloud computing is a technique for helping users to have access to computing resources on demand with very minimal effort in management. It has emerged as a popular computing model in order to support processing of large scale data. There are several cloud service providers, for example, Amazon, Microsoft, Google, IBM and Yahoo which have built platforms for other enterprises and cloud users in order to access the services provided in the cloud.

The areas in which cloud computing is used vary from nuclear physics to geothermal experiments. Thus we need data centers to help provide powerful computation resources for these vital areas. A Data Center (DC) consists of a large number of Physical Machines (PM) arranged on racks and packed densely

DOI: 10.4018/978-1-5225-1721-4.ch006

## ***Dynamic Virtual Machine Placement in Cloud Computing***

in order to increase space utilization. In cloud computing, one of the key concepts for data center management is virtualization. The most important advantage of virtualization is the capacity to run many instances of operating system in a single PM thus making maximum use of the hardware capacities of the PM which helps immensely in saving money for energy and hardware costs. These individual instances of operating system are called Virtual Machines (VM). Thus, in cloud computing, users access the computing resources of a DC with the help of VMs.

The scheduling of VMs in DCs and VM scheduling on PMs form an important part in cloud computing now more than ever because of the growing number of users. VM scheduling has a massive effect on the system's performance and throughput. Thus, the two types of placement of VMs include Virtual Machine Placement (VMP) over DCs and VMP over PMs. The effect that VMP provides is directly related to the Quality of Service (QoS) provided by the services in the cloud. The major objective of VMP is to make maximum utilization of a data center, thus making use of a lesser number of DCs. This helps in achieving more availability and flexibility of DCs, while reducing the costs incurred during operating and maintaining hardware.

## **Anatomy of Cloud Computing**

In cloud computing, virtualization is the most important factor in giving dynamic and scalable architectures. Apart from providing the important aspects of resource sharing and scalability, virtualization also contributes to the capability of virtual machine migration between PMs in order to balance the load (Jones, 2010).

Figure 1 depicts the key elements in a single node in the environment of cloud computing. The component of virtualization in a cloud node is given by the Hypervisor, which is also called as the Virtual Machine Monitor (VMM). This layer is responsible for providing the interface of executing many instances of operating systems in one PM. Thus, hypervisors create virtual machine objects which provide encapsulation for operating system, applications and configuration. Another important aspect in a cloud node known as device emulation is given either in the hypervisor or as a VM. VM management takes place both in local PMs as well as in global DCs.

Thus in order to form a DC, the cloud node shown in Figure 1 is multiplied over a network provided with the management orchestration over the complete infrastructure:

- **Hypervisors:** This is the initial level of a PM (a single cloud node). It provides the virtual operating platform, thus managing the execution of guest operating systems, known as VMs. The VMs share the virtualized hardware resources of the cloud node. One of the best examples of hypervisors for production environments is the Linux Kernel Virtual Machine (KVM).
- **Device Emulation:** Hypervisors are responsible for providing the platform for sharing virtualized physical resources. But in order to achieve complete virtualization, the whole cloud node must be virtualized. This is the job of a device emulator. QEMU is the example of a complete package combining emulator and hypervisor.
- **Virtual Networking:** As the number of cloud users grow, the number of VMs increases. In order to achieve intensified networking system, virtual machines are required to communicate on a virtualized network rather than on a physical level. This reduces the load on the physical infrastructure of the system. Virtual switches are also introduced for effective communication among the VMs.

30 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:  
[www.igi-global.com/chapter/dynamic-virtual-machine-placement-in-cloud-computing/171351](http://www.igi-global.com/chapter/dynamic-virtual-machine-placement-in-cloud-computing/171351)

## Related Content

---

### Blockchain Technology in Cloud Security

Manivannan Karunakaran, Kiran Bellam, J. Benadict Rajaand D. Shanthi (2024). *Emerging Technologies and Security in Cloud Computing* (pp. 53-75).

[www.irma-international.org/chapter/blockchain-technology-in-cloud-security/339396](http://www.irma-international.org/chapter/blockchain-technology-in-cloud-security/339396)

### Policy-Driven Middleware for Multi-Tenant SaaS Services Configuration

Khadija Aouzal, Hatim Hafiddiand Mohamed Dahchour (2019). *International Journal of Cloud Applications and Computing* (pp. 86-106).

[www.irma-international.org/article/policy-driven-middleware-for-multi-tenant-saas-services-configuration/236128](http://www.irma-international.org/article/policy-driven-middleware-for-multi-tenant-saas-services-configuration/236128)

### Green Fashion Evolution and Its Impact on Indian Youth in This Modern Era

Kumari Nayan Tara Singhand Baisakhi Dalapati (2025). *Establishing AI-Specific Cloud Computing Infrastructure* (pp. 557-576).

[www.irma-international.org/chapter/green-fashion-evolution-and-its-impact-on-indian-youth-in-this-modern-era/374454](http://www.irma-international.org/chapter/green-fashion-evolution-and-its-impact-on-indian-youth-in-this-modern-era/374454)

### Ethical AI and Human-in-the-Loop Security in Industrial Operations

S. Ida Evangeline (2026). *Cybersecurity for Industrial IoT in Hybrid Cloud Environments* (pp. 483-506).

[www.irma-international.org/chapter/ethical-ai-and-human-in-the-loop-security-in-industrial-operations/412070](http://www.irma-international.org/chapter/ethical-ai-and-human-in-the-loop-security-in-industrial-operations/412070)

### Impacts of Digital Technologies Across Generations

Ketki P. Kshirsagarand Anup Ingle (2025). *Bridging Academia and Industry Through Cloud Integration in Education* (pp. 1-36).

[www.irma-international.org/chapter/impacts-of-digital-technologies-across-generations/369571](http://www.irma-international.org/chapter/impacts-of-digital-technologies-across-generations/369571)