# A Survey on Load Balancing Techniques Based on Virtualization in Cloud

**Anupama Tiwari[1], Pankaj Richhariya[1], Dr. Satyaranjan Patra[1]**

BITS, Bhopal[1]

**Abstract:** Cloud computing refers to online available collection of resources like processing elements and storage. In cloud a single server is shared among various users. Virtualization is the concept which provides this service. In this paper a survey on load balancing techniques is performed for cloud.

**Keywords:** Load Balancing Techniques, Virtualization, PaaS, SaaS, IaaS.

## I. INTRODUCTION

The cloud computing is computing technique within which giant teams of remote servers square measure networked to permit centralized information storage and on-line access to laptop services or resources. It provides flexibility and completely different computing platform for organizations. The Interest is associate example of cloud computing, it's a free web site that needs registration to access, wherever users will transfer, save, type and manage pictures called pins and different media content like videos and pictures, through collections called inboard. It viewed by around seventeen million folks per month and encompasses a nice storage capability. It's hosted on Amazon's cloud platform [2]. Initially, because of absence of cloud computing, maintaining security of the knowledge had been terribly tough task. However, introduction of the cloud computing has created life straightforward. The cloud consists of bound components like purchasers, servers and therefore the main center wherever all servers square measure managed [3].

Cloud computing could be a model for providing service as Platform, Software, Hardware as a service over web facultative present, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) which will be quickly provisioned and free with nominal management effort or service supplier interaction [1]. Cloud computing is that the latest effort in delivering computing resources as a service. User will rent computing resources on-line as a product these computing resources are utilized by user on-line [1], this helps in reducing prices {and offer and supply}rest can take hold of it and provide on demand service that's delivered to customers over the net from large-scale information centers or "clouds". While cloud computing is gaining growing quality within the IT trade. Cloud Computing contains assortment of distributed servers called masters World Health Organization offer demanded services and resources to the purchasers called central controller or cloud manager in a very network with the measurability and liableness of servers. On-demand service is provided by distributed servers.

Services are also of computer code resources (e.g. computer code as a Service, SaaS) or physical resources (e.g. Platform as a Service, PaaS) or hardware/infrastructure (e.g. Hardware as a Service, HaaS or Infrastructure as a Service, IaaS). Amazon EC2 (Amazon Elastic work out Cloud) is associate example of cloud computing services. Cloud services square measure provisioned to use by service suppliers, for instance, Amazon, Google on the net. Usually, the resources accessible to the user of the cloud square measure virtualized that's (PaaS, IaaS, and SaaS) services square measure virtual service. User gets needed service with none dependencies or constraints reciprocally firms can take some charge for mistreatment their services that square measure nominal as compared to the particular price of that specific service, thanks to this cloud is changing into standard. Cloud computing technology uses web and central remote servers to keep up information and applications [2].

Cloud computing permits any user from anyplace to use the updated version of services and application. We'd like to not purchase computer code with the license as a result of change and maintaining computer code square measure the server's responsibility, solely we'd like to own web affiliation; therewith we will use applications while not putting in it on our system. Gmail, yahoo mail or Hotmail, etc. square measure the common and wide used cloud example, for mistreatment any reasonably cloud service, you need to have a web affiliation. Scientific application desires an outsized quantity of calculation and storage that you desire an outsized computation storage and power. Initially, all the scientific applications square measure deployed on Grid [3].

However, Grid computing is dear and not accessible everywhere the word. Thus the scientific applications square measure moving toward the cloud, Cloud provides another to grid and supercomputers for a person in a very lower price. Cloud is associate rising space and excellent for this type of application. For the advance of deploying associate application of cloud there square measure such a

big amount of methods developed, for instance load reconciliation, planning algorithmic program for VM allocation in cloud.

## II. TYPES OF CLOUD

There are four types of cloud computing exists which are described as follows.

**Public cloud –** publically cloud, any subscriber will access data from the cloud house with an online affiliation. Public clouds are in hand and operated by third parties; they deliver superior economies of scale to customers, because the infrastructure prices are unfold among a combination of users, giving every individual shopper a beautiful inexpensive, "Pay-as-you-go" model. All customers share a similar infrastructure pool with restricted configuration, security protections, and convenience variances. These are managed and supported by the cloud supplier. one among the benefits of a Public cloud is that they'll be larger than AN enterprises cloud, so providing the flexibility to scale seamlessly, on demand.

**Private cloud -** a non-public cloud is for specific cluster or organization and limit access to simply that cluster. Non-public clouds are designed solely for one enterprise. They aim to deal with issues on information security and supply larger management that is usually lacking in a very public cloud.

**Community cloud -** In community cloud, 2 or a lot of organizations that have similar cloud needs, share their intelligence.

**Hybrid cloud –** Hybrid cloud contains of the mix of minimum 2 clouds, will be a mix of public, community or non-public cloud. Hybrid Clouds mix each public and personal cloud models. With a Hybrid Cloud, service suppliers will utilize third party Cloud suppliers in a very full or partial manner so increasing the pliability of computing. The Hybrid cloud setting is capable of providing on-demand, outwardly provisioned scale. The flexibility to reinforce a non-public cloud with the resources of a public cloud will be wont to manage any surprising surges in work.

## III. VIRTUALIZATION IN CLOUD

In cloud computing the resource allocation possesses associate awfully important role within the performance of the whole system and conjointly the extent of client satisfaction provided by the system. But whereas providing the utmost client satisfaction the service supplier needs to make certain the profits that incur to them conjointly. The resource allocation ought to be economical on each view i.e. on the tip user and therefore the service supplier perspective. Thus on get such a system the new technologies insist that the system ought to be with minimum SLA (Service Level Agreements) violation

**SLA**: The service level agreement [18] is a part of the terms that is offered by the service provider to give assurance to the end user regarding the level of service that it can provide to the end user. In short, for a customer high QoS suggests few SLA violations

Virtualization is a popular solution that acts as a backbone for provisioning requirements of a cloud-based solution**.**

**Virtualization**: Virtualization is the use of hardware and software resources to create the perception that one or more entities exist, although the entities, in actuality are not physically present. Using virtualization, we can make one server appear to be many, a desktop computer appear to be running multiple operating system simultaneously, many network connection appear to exist, or a vast amount of disk space or a vast number of drives to be available. The ability to create virtual machines (VMs) [19] dynamically on demand is a popular solution for managing resources on physical machines. Virtualization provides a "virtualized" view of resources used to instantiate virtual machines (VMs). A VM monitor (VMM) or hypervisor manages and multiplexes access to the physical resources, maintaining isolation between VMs at all times. As the physical resources are virtualized, several VMs, each of which is self-contained with its own operating system, can execute on a physical machine (PM). The hypervisor [3], which arbitrates access to physical resources, can manipulate the extent of access to a resource (memory allocated or CPU allocated to a VM, etc.). In the context of virtualization, almost every component of IT systems can be virtualized. This includes Servers, Desktops, Applications, Wide Area Networks, Local Area Networks, Storage And Appliances A cloud provider's resource management actions toward simultaneously minimizing resource usage and maximizing SLA adherence can be classified as follows:

**Load Balancing**: There are various resource management policies for balance load in datacenter. The goal of load balancing is to avoid a situation where there is a large discrepancy in resource utilization levels of the PMs (refer to Figure 3.2). A desired scenario could be to have equal residual resource capacity across PMs (to help increase local resource allocations during increase demands). Virtual machine migrations can be employed to achieve this balance. Load balancing [20] is of two types:
**Static Load balancing-** In this approach of load balancing, we consider static information of system to choose the least loaded node. It performs better in terms of complexity issue but compromises with the result as decision is made on statically gathered data.
**Dynamic Load balancing-** In this strategy, current system state plays major role while making decisions. Despite the fact that dynamic load balancing has higher run rime complexity then static one, dynamic has better performance report as it considers current load of system for choosing next datacenter to serve the request.
This will surely provide an optimal choice from available ones for that state of system.

**Power Saving**: One of the main aspects of Resource management techniques to minimize power consumption at datacenter. To achieve energy efficiency in cloud computing the following methods are useful.

**Server Consolidation**: The goal of consolidation is to avoid low-resource-usage of hosts. VMs on lightly loaded hosts can be "packed "onto fewer machines to meet resource requirements. The freed-up PMs can either be switched off (to save power) or represent higher-resource availability bins for new VMs.

## IV. LITERATURE REVIEW

In [12] an admission control and scheduling mechanism proposes which not only maximizes the resource utilization and profit, but also ensures that the QoS requirements is proposed. Mixed Workload Aware Policy (MWAP) is implemented to consider the workload of different types of application such transactional and non-interactive batch jobs. The proposed mechanism provides substantial improvement over static server consolidation and reduces SLA violations. • In [13] VM consolidation problem which is a NP Hard problem is solved by applying meta-heuristic algorithm ACO. The objective is to lower down the energy consumption of the overall algorithm. And the algorithm also reduces VM migrations. In [6] a novel allocation and selection policy for the dynamic virtual machine (VM) consolidation in virtualized data centers to reduce energy consumption and SLA violation. Firstly, it detect overloading hosts in virtual environments and then apply a method to select VMs from those overloading hosts for migration. VM Provisioning Method to Improve the Profit and SLA Violation of Cloud Service Providers. In [2] authors proposed an Threshold based algorithm for VM provisioning among multiple service providers that reduces SLA Violation. It uses two threshold values and two type of VMs (on-demand and reserved),These threshold values will be decided by the cloud federation depending on the environmental conditions like current workload, idle capacity of each cloud provider, etc. In [11] a power friendly algorithm is proposed. This paper compared live and non live VM migration in terms of power consumption.

In [16] authors developed an objective method to facilitate the comparison of different virtual machine placement algorithms in the cloud. In [17]stable matching framework to decouple policies from mechanisms when mapping virtual machines to physical servers are presented and a general resource management architecture called Anchor is proposed. In [18] the resource allocation problem to be a convex optimization problem and proposed a self-organizing cloud architecture is discussed. Speitkamp and Bichler [20] studied the static consolidation problem with a mathematical programming approach.In [19] they modelled the consolidation as a modified bin-packing problem. These works focus on the initial VM deployment or static consolidation problem based on resource utilization and do not consider VM migration overhead.

## REFERENCES

[1] Komal Singh Patel and A. K. Sarje, "VM Provisioning Method to Improve the Profit and SLA Violation of Cloud Service Providers," IEEE International Conference, Cloud Computing in Emerging Markets (CCEM) 11-12 Oct. 2012.

[2] K. S. Patel and A.K. Sarje, "VM Provisioning Policies to Improve the Profit of Cloud Infrastructure Service Providers," ICCCNT-12, July.2012.

[3] Gundeep Singh Bindra, Prashant Kumar Singh,Seema Khanna,Krishen Kant Kandwal, "Cloud Security : Analysis and Risk Management of VM Images," Proceeding of IEEE International Conference on Information and Automation Shenyang, China, June 2012.

[4] Eeraj Jan Qaisar, "Introduction to Cloud Computing for Developers," In IEEE ©2012.

[5] A. Beloglazov, R. Buyya, "Optimal Online Deterministic Algorithms and Adaptive Heuristics for Energy and Performance Efficient Dynamic Consolidation of Virtual Machines in Cloud Data Centers," Concurrency and Computation: Practice and Experience (CCPE), Wiley Press, New York, USA, Sep. 2012, pp. 1397–1420, doi: 10.1002/cpe.1867.

[6] Zhibo Cao and ShoubinDong , "Dynamic VM consolidation for energy-aware and SLA violation reduction in cloud computing," 13th International Conference on Parallel and Distributed Computing, Applications and Technologies 2012.

[7] YonggenGu, Wei Zhang, YonggenGu, Jie Tao, "A Study of SLA Violation Compensation Mechanism in Complex Cloud Computing Environment," In IEEE © 2012.

[8] C. Belady, "In the data center, power and cooling costs more than the equipment it supports," 2007. URL http://www.electronicscooling.com/articles/2007/feb/a3/.

[9] http://en.wikipedia.org.

[10] http://www.sciencedirect.com/science/article/pii/S1877705811054117.

[11] David Aikema, AndreyMirtchovski, Cameron Kiddle, and Rob Simmonds "Green Cloud VM Migration: Power Use Analysis" in IEEE 2012.

[12] Saurabh Kumar Garg, Adel NadjaranToosi, Srinivasa K. Gopalaiyengar, RajkumarBuyya, "SLA-based virtual machine management for heterogeneous workloads in a cloud datacenter," Journal of Network and Computer Applications 1 August 2014 .

[13] Rafid Sagban, Ku Ruhana Ku Mahamud, Muhamad Shahbani Abu Bakar "Reactive Memory Model for Ant Colony Optimization and Its Application to TSP" in 2014 IEEE International Conference on Control System, Computing and Engineering, 28 - 30 November 2014, Penang, Malaysia.

[14] M. Veluscek, T. Kalganova, P. Broomhead "Improving Ant Colony Optimization Performance through Prediction of Best Termination Condition" in IEEE 2015.

[15] Fahimeh Farahnakian, Adnan Ashraf, Tapio Pahikkala, PasiLiljeberg, JuhaPlosila, Ivan Porres, and HannuTenhunen "Using Ant Colony System to ConsolidateVMs for Green Cloud Computing" in IEEE TRANSACTIONS ON SERVICES COMPUTING, VOL. 8, NO. 2, MARCH/APRIL 2015.

[16] K. Mills, J. Filliben, and C. Dabrowski, "Comparing vm-placement algorithms for on-demand clouds," in Proc. IEEE 3rd Int. Conf. Cloud Comput. Tech. Sci., 2011, pp. 91–98.

[17] H. Xu and B. Li, "Anchor: A versatile and efficient framework for resource management in the cloud," IEEE Trans. Parallel Distrib. Syst., vol. 24, no. 6, pp. 1066–1076, Jun. 2013.

[18] S. Di and C.-L. Wang, "Dynamic optimization of multi-attribute resource allocation in self-organizing clouds," IEEE Trans. Parallel Distrib. Syst., vol. 24, no. 3, pp. 464–478, Mar. 2013.

[19] S. Srikantaiah, A. Kansal, and F. Zhao, "Energy aware consolidation for cloud computing," in Proc. Conf. Power Aware Comp. Syst., 2008, pp. 10–10.

[20] B. Speitkamp and M. Bichler, "A mathematical programming approach for server consolidation problems in virtualized data centers," IEEE Trans. Serv. Comput., vol. 3, no. 4, pp. 266–278, Oct. 2010.