ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

Allocation of Datacenter Resources Based on Demands Using Virtualization Technology in Cloud

G.Rajesh
B.Tech Student,
Depatment of CSE,
TKR College of Engineering & Technology.

L.Bobbian Naik
B.Tech Student,
Depatment of CSE,
TKR College of Engineering & Technology.

K.Mounika
B.Tech Student,
Depatment of CSE,
TKR College of Engineering & Technology.

Dr. K.Venkatesh Sharma
Associate Professor,
Depatment of CSE,
TKR College of Engineering & Technology.

Abstract:

Cloud computing is computing in which large groups of remote servers are networked to allow centralized data storage and online access to computer services or resources. Clouds can be classified as public, private or hybrid.Cloud computing relies on restricting sharing of resources to achieve coherence and economies of scale, similar to a utility (like the electricity grid) over a network. At the foundation of cloud computing is the broader concept of converged infrastructure and shared services.Cloud computing facilitates Industrial and large business users to increase proportionallyand decrease accordingly theirdata center resource usage depending on requirements. Most of theflauntedbenefits in the cloud model originate from resource multiplexing through virtualization technology.

In this research paper, we studied and simulated a system that utilizes virtualization technology to distribute data center resources automatically depending on application requirements and maintain green computing by optimizing the number of servers in deployment. We analyzed the concept of "skewness" to determine the unevenness in the multi-dimensional resource utilization of a server. By minimizing skewness, we can merge different types of workloads nicely and develop the overall utilization of server resources. We developed a set of heuristics that avoid overload in the system efficientlyat the same time as saving powerutilized. Trace driven simulation and experiment results demonstrate that our algorithm achieves good performance.

Keywords:

Cloud computing, Dynamic Resource Allocation, Data center, Virtual machine, Ranking and Load balancing.

Introduction:

Cloud computing, or in simpler shorthand just "the cloud", also focuses on maximizing the effectiveness of the shared resources. Cloud resources are usually not only shared by multiple users but are also dynamically reallocated per demand.

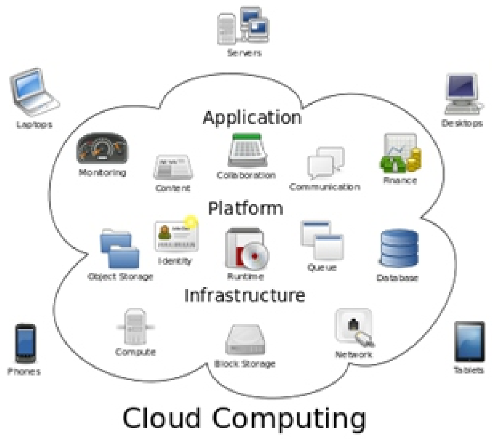
This can work for allocating resources to users. For example, a cloud computer facility that serves European users during European business hours with a specific application (e.g., email) may reallocate the same resources to serve North American users during North America's business hours with a different application (e.g., a web server).

This approach should maximize the use of computing power thus reducing environmental damage as well since less power, air conditioning, rackspace, etc. are required for a variety of functions. With cloud computing, multiple users can access a single server to retrieve and update their data without purchasing licenses for different applications.The National Institute of Standards and Technology's definition of cloud computing identifies "five essential characteristics":

On-demand self-service:

A consumer can unilaterally provision computing capabilities, such as server time and network storage, as needed automatically without requiring human interaction with each service provider.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal



Broad network access:

Capabilities are available over the network and accessed through standard mechanisms that promote use by heterogeneous thin or thick client platforms (e.g., mobile phones, tablets, laptops, and workstations).

Resource pooling:

The provider's computing resources are pooled to serve multiple consumers using a multi-tenant model, with different physical and virtual resources dynamically assigned and reassigned according to consumer demand.

Rapid elasticity:

Capabilities can be elastically provisioned and released, in some cases automatically, to scale rapidly outward and inward commensurate with demand. To the consumer, the capabilities available for provisioning often appear unlimited and can be appropriated in any quantity at any time.

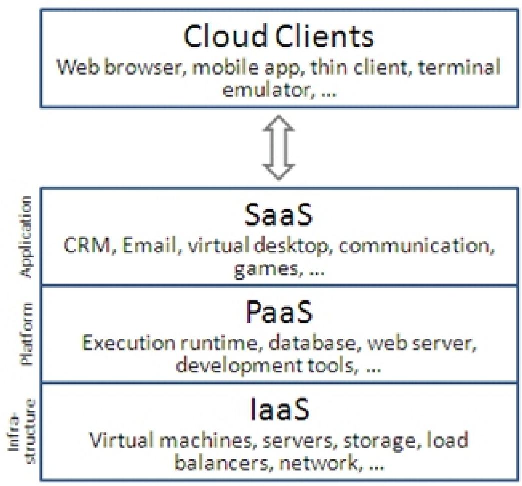
Measured service:

Cloud systems automatically control and optimize resource use by leveraging a metering capability at some level of abstraction appropriate to the type of service (e.g., storage, processing, bandwidth, and active user accounts). Resource usage can be monitored, controlled, and reported, providing transparency for both the provider and consumer of the utilized service.

Service models: Cloud computing providers offer their services according to several fundamental models:

Infrastructure as a service (IaaS):

In the most basic cloud-service model & according to the IETF (Internet Engineering Task Force), providers of IaaS offer computers – physical or (more often) virtual machines – and other resources. (A hypervisor, such as Xen, Oracle VirtualBox, KVM, VMware ESX/ESXi, or Hyper-V runs the virtual machines as guests. Pools of hypervisors within the cloud operational support-system can support large numbers of virtual machines and the ability to scale services up and down according to customers' varying requirements.) IaaS clouds often offer additional resources such as a virtual-machine disk image library, raw block storage, and file or object storage, firewalls, load balancers, IP addresses, virtual local area networks (VLANs), and software bundles. IaaS-cloud providers supply these resources on-demand from their large pools installed in data centers. For wide-area connectivity, customers can use either the Internet or carrier clouds (dedicated virtual private networks).



Platform as a service (PaaS):

In the PaaS models, cloud providers deliver a computing platform, typically including operating system, programming language execution environment, database, and web server. Application developers can develop and run their software solutions on a cloud platform without the cost and complexity of buying and managing the underlying hardware and software layers.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

With some PaaS offers like Microsoft Azure and Google App Engine, the underlying computer and storage resources scale automatically to match application demand so that the cloud user does not have to allocate resources manually. The latter has also been proposed by an architecture aiming to facilitate real-time in cloud environments. Platform as a service (PaaS) provides a computing platform and a key chimney. It joins with software as a service (SaaS) and infrastructure as a service (IaaS), model of cloud computing.

Software as a service (SaaS):

In the business model using software as a service (SaaS), users are provided access to application software and databases. Cloud providers manage the infrastructure and platforms that run the applications. SaaS is sometimes referred to as "on-demand software" and is usually priced on a pay-per-use basis. SaaS providers generally price applications using a subscription fee.

In the SaaS model, cloud providers install and operate application software in the cloud and cloud users access the software from cloud clients. Cloud users do not manage the cloud infrastructure and platform where the application runs. This eliminates the need to install and run the application on the cloud user's own computers, which simplifies maintenance and support. Cloud applications are different from other applications in their scalability—which can be achieved by cloning tasks onto multiple virtual machines at run-time to meet changing work demand. Load balancers distribute the work over the set of virtual machines. This process is transparent to the cloud user, who sees only a single access point. To accommodate a large number of cloud users, cloud applications can be multitenant, that is, any machine serves more than one cloud user organization.

Existing System:

Virtual machine monitors (VMMs) like Xen provide a mechanism for mapping virtual machines (VMs) to physical resources. This mapping is largely hidden from the cloud users. Users with the Amazon EC2 service [4], for example, do not know where their VM instances run. It is up to the cloud provider to make sure the underlying physical machines (PMs) have sufficient resources to meet their needs.

VM live migration technology makes it possible to change the mapping between VMs and PMs while applications are running.

Disadvantages of Existing System:

• A policy issue remains as how to decide the mapping adaptively so that the resource demands of VMs are met while the number of PMs used is minimized.

• This is challenging when the resource needs of VMs are heterogeneous due to the diverse set of applications they run and vary with time as the workloads grow and shrink. The two main disadvantages are overload avoidance and green computing.

Proposed System:

In this paper, we present the design and implementation of an automated resource management system that achieves a good balance between the two goal:

• Overload avoidance: the capacity of a PM should be sufficient to satisfy the resource needs of all VMs running on it. Otherwise, the PM is overloaded and can lead to degraded performance of its VMs.

• Green computing: the number of PMs used should be minimized as long as they can still satisfy the needs of all VMs. Idle PMs can be turned off to save energy.

Advantage of Proposed System:

• We develop a resource allocation system that can avoid overload in the system effectively while minimizing the number of servers used.

• We introduce the concept of "skewness" to measure the uneven utilization of a server. By minimizing skewness, we can improve the overall utilization of servers in the face of multi-dimensional resource constraints.

Module Description:

After careful analysis the system has been identified to have the following modules:

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

1. Cloud Computing Module.

2. Resource Management Module.

3. Virtualization Module.

4. Green Computing Module.

1.Cloud Computing Module:

Cloud computing refers to applications and services offered over the Internet. These services are offered from data centers all over the world, which collectively are referred to as the "cloud." Cloud computing is a movement away from applications needing to be installed on an individual's computer towards the applications being hosted online. Cloud resources are usually not only shared by multiple users but as well as dynamically re-allocated as per demand. This can work for allocating resources to users in different time zones.

2. Resource Management Module:

Dynamic resource management has become an active area of research in the Cloud Computing paradigm. Cost of resources varies significantly depending on configuration for using them. Hence efficient management of resources is of prime interest to both Cloud Providers and Cloud Users. The success of any cloud management software critically de-pends on the flexibility; scale and efficiency with which it can utilize the underlying hardware resources while pro-viding necessary performance isolation. Successful resource management solution for cloud environments, needs to provide a rich set of resource controls for better isolation, while doing initial placement and load balancing for efficient utilization of underlying resources.

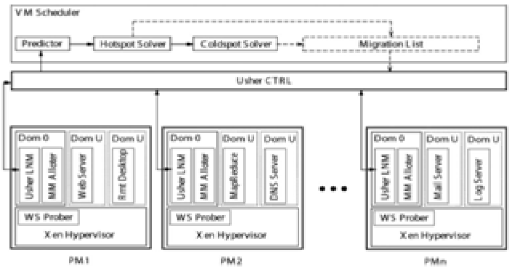
3. Virtualization Module:

Virtualization, in computing, is the creation of a virtual (rather than actual) Version of something, such as a hardware platform, operating system, and a storage device or network resources.VM live migration is a widely used technique for dynamic resource allocation in a virtualized environment. The process of running two or more logical computer system so on one set of physical hardware. Dynamic placement of virtual servers to minimize SLA violations.

4. GreenComputing Module:

Many efforts have been made to curtail energy consumption. Hardware based approaches include novel thermal design for lower cooling power, or adopting power-proportional and low-power hardware. Dynamic Voltage and Frequency Scaling (DVFS) to adjust CPU power according to its load in data centers. Our work belongs to the category of pure-software low-cost Solutions. It requires that the desktop is virtualized with shared storage. Green computing ensures end user satisfaction, regulatory compliance, telecommuting, virtualization of server resources.

System Architecture:



Conclusion:

Cloud computing can solve complex set of tasks in shorter time by proper resource utilization. To make the cloudto work efficiently, best resource allocation strategies have to be employed. Utilization of resources is one of the mostimportant tasks in cloud computing environment where the user's jobs are scheduled to different machines. In this paper a new VM load balancing algorithm was proposed and then implemented in Amazon EC2 Cloud computing environment using java language. Proposed algorithm find the available cpu cycle of each Virtual Machine (VM) and Send the ID of Virtual Machine to theCloud controller for allocating the new request. We conclude that Cloud controller utilizes the available resources on Virtual Machine then it effect the overall performance of the Cloud Environment and also decrease the average response time.

References:

[1] Zhen Xiao, Senior Member, IEEE, Weijia Song, and Qi Chen-"Dynamic Resource Allocation Using Virtual Machines for Cloud Computing Environment"- IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEMS, VOL. 24, NO. 6, JUNE 2013.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

[2] Jinhua Hu, Jianhua Gu, Guofei Sun, Tianhai Zhao "A Scheduling Strategy on Load Balancing of Virtual Machine Resources in Cloud Computing Environment" 3rd international Symposium on Parallel Architectures, Algorithms and Programming.

[3] Rashmi. K. S, Suma. V, Vaidehi. M "Enhanced Load Balancing Approach to Avoid Deadlocks in Cloud" on Advanced Computing and Communication Technologies for HPC Applications - ACCTHPCA, June 2012.

[4] Ke Yang, JianhuaGu,Tianhai Zhao,Guofei Sun "An Optimized Control Strategy for Load Balancing based on Live Migration of Virtual Machine" 2011 Sixth Annual China Grid Conference.

[5] Li Zhou, Yi-Cheng Wang, Ji-Lin Zhang, Jian Wan, Yong-Jian Ren "Optimize Block-Level Cloud Storage System With Load-Balance Strategy" 2012 IEEE 26th International Parallel and Distributed Processing Symposium Workshops & PhD Forum.

[6] Shinwon Lee, V. Meka, Mingu Jeon, Nagoo Sung, Jeongnam Youn "Dynamic load balancing for system on chip" IEEE Transactions on parallel and distributed systems 2012.

[7] Singh, M. Korupolu, and D. Mohapatra, "Server-Storage Virtualization: Integration and Load Balancing in Data Centers," Proc. ACM/IEEE Conf. Supercomputing, 2008.

[8] A. Singh, M. Korupolu, and D. Mohapatra, "Server-Storage Virtualization: Integration and Load Balancing in Data Centers" Proc. ACM/IEEE Conf. Supercomputing, 2008.

[9] Y. Agarwal, S.Savage and R. Gupta "sleep server: A software- only approach for reducing of energy consumption of PCS with in enterprise environment" Proc. USENIX Ann. Technical Conf., 2010.

About Author's:

G.Rajesh
B.Tech Student,
Depatment of CSE,
TKR College of Engineering & Technology.

L.Bobbian Naik
B.Tech Student,
Depatment of CSE,
TKR College of Engineering & Technology.

K.Mounika
B.Tech Student,
Depatment of CSE,
TKR College of Engineering & Technology.

Dr. K.Venkatesh Sharma
Associate Professor,
Depatment of CSE,
TKR College of Engineering & Technology.