International Journal of Hybrid Information Technology
Vol. 9, No. 5 (2016), pp. 371-380
http://dx.doi.org/10.14257/ijht.2016.9.5.31

A Distributed Management Method in Cloud Computing Environment

Xi Liu[1,a] and Jun Liu[2, b]

[1]*Colleg of Information Science and Engineering, Yunnan University, Kunming Yunnan China*
[2]*College of Mathematics and Information Science, Qujing Normal University, Qujing Yunnan China*
[a] *lxghost@126.com* , [b]*liujunxei@126.com*

Abstract

Recently, there are some problems in the centralized management, such as centralized management, heavy burden, excessive number of virtual machine migration, lack of mutual cooperation mechanism between nodes, can't adapt to the cluster of change. Aiming at the limitations of centralized management method, this paper presents a distributed management method in cloud computing environment. The method achieve self-management node by using P2P thinking. According to the predetermined threshold node themselves decided to create a virtual machine, virtual machine migration and whether sleep to save energy consumption and other operations. This topology can be well adapted to large-scale changes in the cluster. The experimental results show that the distributed management method not only can effectively improve the quality of service, and can well satisfy the clusters of massive growth. It has a good robustness.

Keywords*: cloud computing; virtual machine; distributed management; global management*

.

1. Introduction

Cloud computing[1-2] has become a hot research applications and is a new computing model which is the development of distributed computing, parallel computing and grid computing. Cloud Computing is not a totally new concept. It has its root in Grid Computing paradigm, and other technologies like cluster computing, utility computing and distributed computing in general. It can be defined as utilization of computer resources be it hardware and software from any device, anytime, any where and you pay as you need. Cloud Computing can be define in various terms. Here we present some of the key definitions of cloud: Cloud Computing is a broad array of web-based services aimed at allowing users to obtain a wide range of functional capabilities on a " pay-as-you-go " basis that previously required tremendous hardware and software investments and professional skills to acquire. The national institute of standards and technology to define the basic characteristics of cloud computing is on-demand self-service and rapid elasticity[3].To achieve these characteristics, the main use virtualization and its related technologies[4-5].

Cloud computing is mainly use virtualization technology [6-7].The current virtualization [8] is defined as: 1) The object of virtualization are all kinds of resources.2) The user does not need to know the details of the hidden in the logical resources virtualize process.3) In the user experience of virtualization, users can implement part or all of the functions in the real environment. Virtualization is a software implementation of a machine which will execute different program like a real machine. Two type of virtualization found in case of cloud computing. Para

ISSN: 1738-9968 IJHIT

Virtualization: It means the hardware allow multiple operating systems to run on a single machine. In pare virtualization all the services are not fully available rather the services are provided partially. Full Virtualization: It means a complete installation of one machine is done on another machine. Individual companies and research institutions have focused on the definition of cloud computing. In academia, the most interested and most researchers have studied the value of cloud computing feature is the use of powerful computing capabilities of cloud computing to solve complex problems. In this area, the cloud computing model is defined as a by massive computing network processing program automatically split into numerous smaller subroutines, and then handed over to multiple servers consisting of a large system by searching for calculation after analyzing the results back to the user. Through this technique, the network service provider may be within a few seconds, reaching tens of millions of information processing, to achieve and "supercomputer" is also a strong network of services, and at the same time can effectively control costs.

Core demand deployment of cloud computing to solve the on-demand deployment, must be resolved dynamically reconfigurable resources, monitoring, and automated deployment, and these also need to be virtualized, high-performance storage, processor, high-speed Internet and other technologies foundation [9-10]. Cloud computing truly on-demand computing, thus effectively improving the utilization of hardware and software resources. It is with the development of processor technology, distributed storage technology, automated management, virtualization technology and broadband Internet technology generated. From a technical level, the basic functionality of cloud computing depends on two key factors, one is the ability to store the data, and the other is a distributed computing power. Cloud computing has the following characteristics: low cost, including user terminals, IT infrastructure and honest people; a high-performance, including enhanced computing power, unlimited storage capacity and enhanced data security; high degree of compatibility between the user of the system and data formats; simple to update and maintain the software; low device-dependent; efficient group collaboration. So cloud computing is a way to quickly deploy with minimal management and service providers as well as the conditions and publish interactive, convenient and available to the user through the network, on-demand access to computing resources can be dynamically configured (eg, networks, servers, storage, applications and services) pooling service model.

Cloud Computing is a broad array of web-based services aimed at allowing users to obtain a wide range of functional capabilities on a “pay-as-you-go” basis that previously required tremendous hardware and software investments and professional skills to acquire[11]. There are three basic types of Service Models depending on the service each layer provides to its upper layers. The three models are: Infrastructure as a Service (IaaS): When user has got the maximum flexibility to deal with the Infrastructure, where he can even install his Operating System it is known as Infrastructure as a Service. In it the provider makes entire infrastructure available as a service. The end user can use the entire infrastructure on pay per usage basis. The ability to support an IaaS architecture is through a combination of some of the special characteristics of cloud computing. Platform as a Service (PaaS): A cloud platform offers an environment on which developers create and deploy applications and do not necessarily need to know how many processors or how much memory that applications will be using. In it the cloud provides software platform or middleware as service on which the applications run. The user is responsible for creation, updating and maintenance of the application. Software as a Service (SaaS): This is the topmost layer of the architecture, it has all the applications residing in it. Applications are provided as a Service and these Services can be accessed by end users through Web portals. Therefore, consumers are increasingly shifting from locally installed computer programs to on-line software services that offer the same functionally.

Features of Cloud Computing: resource pooling, rapid elasticity, scalability, on-demand self-service and efficiency. Resource pooling: The provider's computing resources are pooled to server 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 to be unlimited and can be appropriated in any quantity at any time. Scalability: Cloud computing providers' resources and services for users on demand. The resources are scalable over several data centers. In order to achieve a highly scalable system, balancing of the loads when the load increases at a large extent and a cloud user demands more resources online rapidly is very important. 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. Efficiency: An efficient cloud computing system should work for all the possible configurations of cloud where users are requesting the resources on any extent unknown to the cloud service providers providing the important features like rapid elasticity and high scalability with the much needed fault tolerance.

The cloud computing platform mainly adapts the way of centralized management. Centralized management style is unified management physical nodes and the resources of the virtual machine. Nodes implement centralized management of the virtual machine by virtual machine monitor. However, the current centralized management cannot adapt to a massive increase ,the management node burden, lack of interaction mechanisms between nodes, cannot trigger active migration, cannot decide whether active sleep in order to save energy consumption and other issues[8]. Therefore this paper proposes a distributed management method. This method implements something by the thought of P2P. Such as the node of self-management and meet the dynamic changes of the cluster.

2. Related Work

Currently used mainly centralized management, there are still some questions [12-15]:

(1) Centralized management approach uses a continuous optimization. Each optimized for all virtual machines will once again combined. This process has brought a lot of migration of virtual machines. The virtual machine migration brings network bandwidth utilization, the consumption of CPU resources and so on. Migration process will result in a cloud computing center service performance degradation. The system should try to avoid migration of virtual machines.

(2)Centralized management can be unified management and scheduling, simple structure, easy to operate, but it only applies to smaller clusters. When the system is extended to large-scale cloud computing clusters or the cluster scale dynamic change, centralized management difficult to adapt to the needs of system.

(3)Between physical nodes lack of coordination interaction mechanism. In centralized management, arranged by the management node, the node between don't need to interact with each other. It brings the task of the management node burden is heavy and the other node can only passive accept the arrangement of the management node. The node own independent can't execute virtual machine live migration, to achieve the dynamic allocation of resources.

(4)Centralized management has a single point of failure problem. If the management node is abnormal, it may cause the entire cloud computing platform can't work normally.

3. The Distributed Management Method

3.1. Define the Relevant Parameters

In order to better describe the performance of the virtual machine and the physical server, we use the relevant parameters to define the resource properties.

Virtual machine hardware resources mainly include CPU usage, memory usage, network bandwidth usage and hard disk resources. This paper reference Microsoft Virtual Machine Manager technical report of the 2008 for the physical server performance evaluation standard according to the following method to design virtual machine resources.

A=CPU utilization = CPU usage /（CPU total-CPU reserved amount）

B=memory utilization = memory usage /（memory total- memory reserved amount）

C=bandwidth utilization = bandwidth usage /（bandwidth total- bandwidth reserved amount）

D=disk utilization = disk usage /（disk total- disk reserved amount）

Define：a=1-A, b=1-B, c=1-C, d=1-D.

The vector U= [A, B, C, D] represents the resource usage of the current node. The vector u= [a, b, c, d] represents resources which nodes can be used. The vector o= [a1, b1, c1, d1] represents the new task needs to use the pre-determination.

We set High and Low threshold in the node. Define High= [A1, B1, C1, D1], Low= [A2, B2, C2, D2].

If U > High, the node at the High load no longer accept to create a new virtual machine. If there are idle resources around the node, then the trigger migration of virtual machines.

If U < low, then the node is in low load and accept to create a new virtual machine. If node not received to create a virtual machine request over a certain time, the node is in sleep mode to save energy after the completion of the task。

If Low<U<High, the node is in normal working condition. Node creates a new virtual machine based on the current specific resources.

3.2. Virtual Machine Processing

Virtual machine processing operations are divided into six, namely: create, delete, pause, resume, stop, and dynamic migration.

Create a virtual machine: Physical machine receive new information which is create new virtual machine. If u>o, physical machine create a virtual machine according to information and whether to create a successful returned back to the task allocation node.

Delete, pause, resume and stop the virtual machine. There is the interaction between the nodes and mutual coordination mechanism between nodes in the process of task execution. According to the feedback of the surrounding nodes, physical machine decides to stop, delete, suspend or resume virtual machine.

Live migration of virtual machine: If U>High, node implementations the dynamic migration process after node interact with the surrounding nodes. When the node receives feedback of surrounding nodes, it randomly selects a node to implement dynamic migration of virtual machines.

3.3 Model Topology

There is no central server and only the task of receiving node in the distributed management which is based on P2P peer network structure. The task of receiving nodes are used to receive the task distribution, the other nodes is self-managed state. Each node is independent manage their virtual machine resources. Node as a resource provider for tasks provides CPU, memory, network bandwidth and disk resources. Node as a consumer of resources implements the dynamic migration of virtual machine to another node. These actions are completed self-negotiation between nodes. Communication between nodes uses UDP. When nodes are in the same broadcast domain, they can exchange information through multicast. If the network condition is poorer, nodes use TCP exchange information.

When the task nodes receive new missions, they broadcast to each node by UDP packet. Nodes choose whether or not the respond according to their own resources. If u<o, the node sends a response by UDP packets. Task nodes which are receive the response packets randomly select nodes to distribute task.

When resource usage of the node is U>High, it sends the live migration of virtual machine packets to around nodes and trigger the node dynamic migration process. When resource usage of the around nodes is u<o and Low<U<High, they respond to this live migration of virtual machine packets. When the node receives response packets from around nodes, it chooses the one of the most idle node for the virtual machine migration

4. System Implementation and Experiment

4.1 System Implementation

We implement a distributed prototype system in this paper. System based on Windows 7 platform, using c # language, using visual studio 2010 development environment. The system uses the Timer class implements resource collection, multicast, task distribution, virtual machine operations and other operations, such as using Thread multithreading technology realize the thread concurrent execution module.

4.2. System Experiment

The system set the three different of nodes for experiment. The number of nodes arranged as shown in Table 1.

Table 1. Number of Nodes Arrangement

Experiment	Physical nodes	Virtual machine number
Experiment 1	40	100
Experiment 2	200	400
Experiment 3	500	1500

This paper sets the threshold is as follows, High = 80% * node greatest resources, and Low= 20% * node greats resources. Figure 1 shows the number of virtual machine migration under different cluster scale. Figure 2 shows a violation of the SLA probability under different cluster scale. SLA violation refers to the virtual machine can not get the

CPU resource. The lower the probability of violation of the SLA, the higher the quality of service systems.

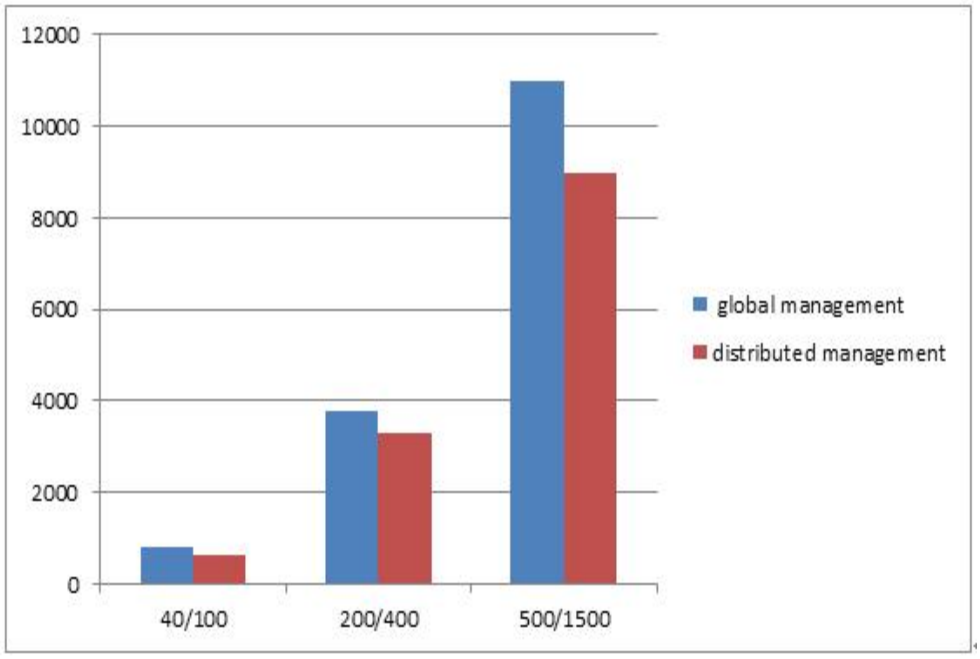

Figure 1. Number of Virtual Machine Migration under Different Cluster Scale

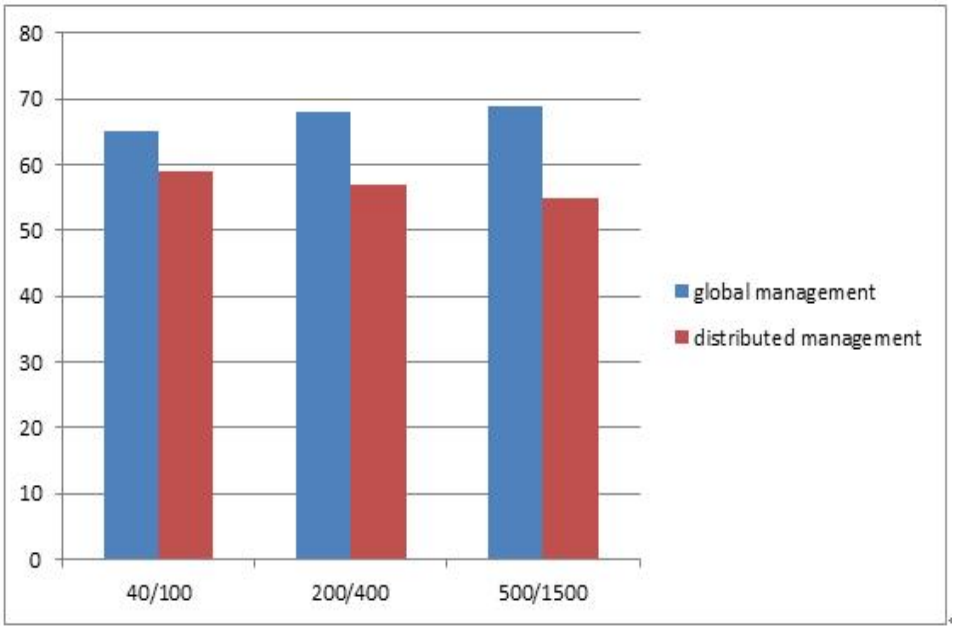

Figure 2. Comparison of SLA Probability under Different Cluster Scale

Table 2. Number of Nodes Arrangement

Time(s)	Users number lower	Users number Upper
100	10	15
200	20	25
300	25	35
400	40	45
500	40	45
600	45	50
700	50	55
800	55	60
900	55	60
1000	60	65
1100	60	65
1200-1900	70	100

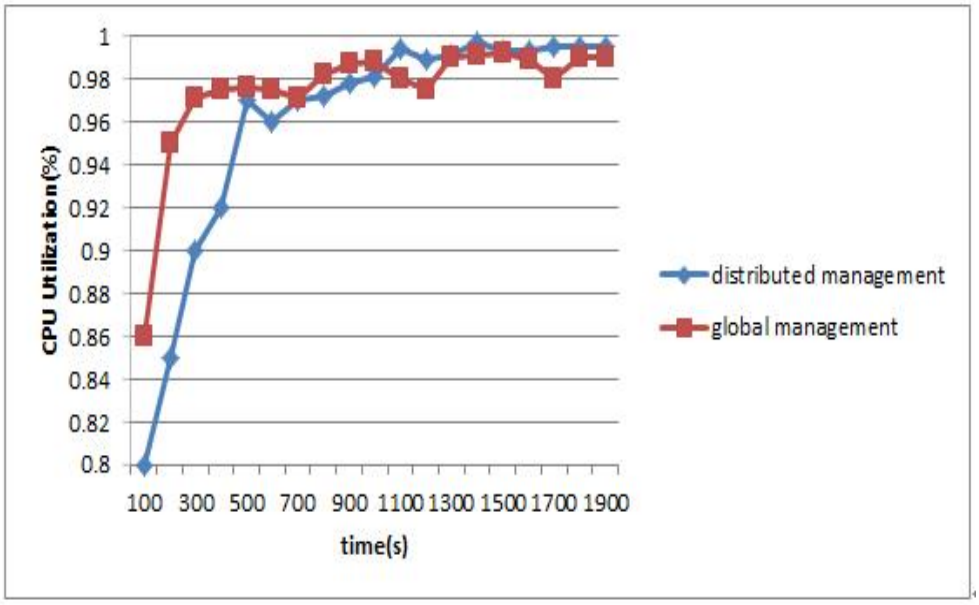

Figure 3. Cpu Utilization

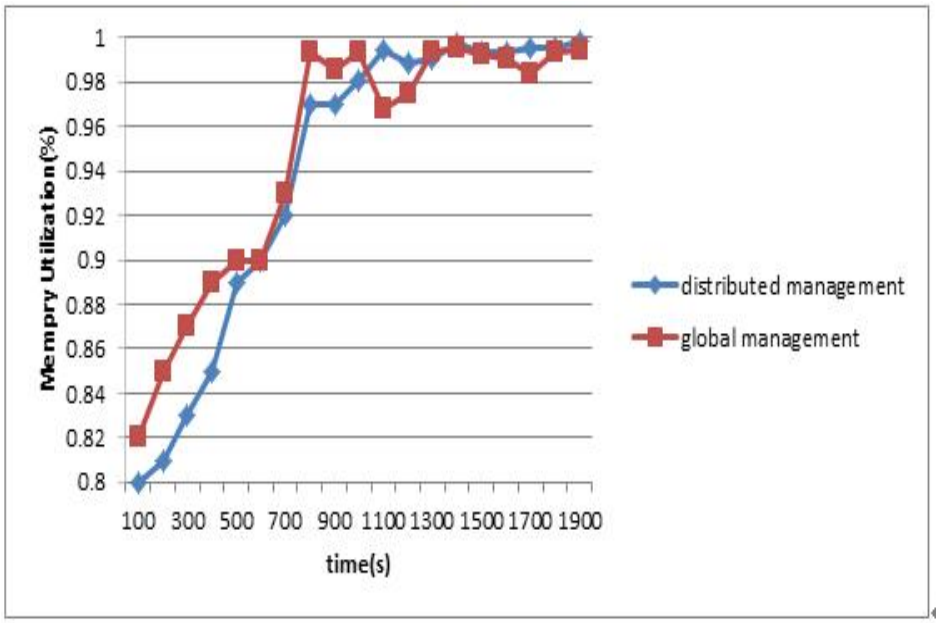

Figure 4. Memory Utilization

The experimental data show that compared with centralized management, distributed management of virtual machine live migration has the obvious drop, significantly lower than the probability of SLA violation centralized management. When the number of nodes is small, the gap is small. As the number of nodes increases, the gap is more obvious.

We simulate the real cloud computing environment, users at different times randomly added to the system. Figure3 shows the CPU utilization when users randomly added to the system, figure4 shows memory usage when the users randomly added to the system. The cloud system has 200 physical nodes and 400 virtual machines. We can see the figure 3 and 4, when the number of users is small, the centralized management has advantage. Because centralized management allocates and manages well to resources (CPU and memory) when the users of number in the system are small. When the number of users is large, distributed management expresses than centralized management on CPU utilization and memory utilization. The experiment results show that the distributed management method makes the system to improve the quality of service.

5. Conclusions

This paper presents a distributed management method. Experimental results show that in large- scale cloud computing systems, a distributed management can adapt to more large-scale cloud computing platform. While ensuring quality and reliable services, it can effectively reduce the migration of virtual machines. We only take into account the user added to the system, without taking into account the circumstances for leaving the system. The next step will be to study how to improve the interaction between nodes, and to ensure that if the task nodes are failure, the remaining nodes automatically select some nodes as task nodes.

Acknowledgements

The work was supported by Chinese Natural Science Foundation Grant No.11361048, Yunnan NSF Grant No.2010CD086.

References

[1] M. Ambrust, A. Fox and R. Griffith, "Above the Clouds: A Berkeley View of Cloud Computing [EB/OL]", http://www.eecs.berkeley.edu/pubs/techrpts/2009/EECS-2009-28, pdf.
[2] K. Chen and W. Zheng, "Journal Software", vol.5, no. 20, **(2009)**.
[3] J. Mell and T. Grance, "National Institute of Standards and Technology", vol. 3, no.25, **(2011)**.
[4] R. Buyya and C. S. Yeo, "Future Generation Computer System", vol. 2, no.25, **(2009)**.
[5] L. Inbger, "Math. Conput. Modeling". Vol.4, no. 12, **(1989)**.
[6] P. M. Pardal, H. E. Pomeijn and H. Tuyh, "Journal Comput and Appl. Math", vol. 2, no. 24, (**2009**)
[7] B. Noram, K. Andrzej and B. Kirk, "International Symposium on Integrated Network Management", vol. 2, no. 22, **(2007)**.
[8] Virtualization & Cloud Computing [EB/OL], [**2011**-10-20], http://www.vmware.com/solutions/cloud-computing/index.html.
[9] M. Armbrust, A. Fox and R. Griffith, EECS Department, University of California, Berkeley, Tech Rep UCB/EECS-2009-28, **(2009)**
[10] A. Elastic, "Compute Cloud", vol. 2, no. 10, **(2009)**.
[11] J.S. Chase, E.D. Irwin and E.L. Grit, Proceedings 12th IEEE International Symposium on High Performance Distributed Computing, **(2003)**, pp. 90-100.
[12] Y. D. Li, H. T. Liu and S. A. Bin, Computer Engineering and Design, vol. 6, no.32, **(2011)**.
[13] E. W. William, Editor, "Gerald Teasuro of jpb scheduiling algorithms in cloud computing environment", Beijing Jiaotong University Publishers, Beijing, **(2009)**.
[14] T. Gerald, D. Rajarshi and E. W. Willian, Proceedings of the Scend International Conference on Autonomic Computing, vol. 5, no. 12, **(2005)**.
[15] V. Akshat, A. Puneet and N. Anindya, "Power and Migration Cost Aware Application Placement in Virualized Systems"; Proceedings of the 9th ACM/IFIP/USENIX International Conference on Middleware, New York, USA: ACM Press, **(2008)**.

Authors

Xi Liu, Doctor, College of Information Science and Engineering, Yunnan University, Kunming Yunnan China

Jun Liu, Professor, College of Mathematics and Information Science, Qujing Normal University, Qujing Yunnan China