



Research on Virtual Machine Migration and Load Balancing in Cloud Computing Environment

Li Gao

College of Science and Culture, Jiangsu Normal University, Xuzhou, Jiangsu, China.

How to cite this paper: Li Gao. (2023) Research on Virtual Machine Migration and Load Balancing in Cloud Computing Environment. *Advances in Computer and Communication*, 4(5), 277-282.
DOI: 10.26855/acc.2023.10.003

Received: September 20, 2023

Accepted: October 16, 2023

Published: November 14, 2023

***Corresponding author:** Li Gao, College of Science and Culture, Jiangsu Normal University, Xuzhou, Jiangsu, China.

Abstract

Cloud computing technology has greatly changed the way we live and work, bringing numerous advancements to humanity. It provides efficient data storage and processing capabilities and promotes the development of the digital economy and intelligent society. With the widespread application of cloud computing technology, virtualization technology, as one of the core components of cloud computing, offers significant advantages in terms of resource utilization and flexibility. As an important issue in the virtualization environment, VM migration and balancing technology are crucial for improving system performance and resource utilization efficiency. This paper discusses the theoretical basis and practical application of virtual machine migration and load balancing technology in a cloud computing environment. Firstly, the concept and significance of virtual machine migration and load balancing are analyzed. Secondly, it introduces common VM migration and load balancing strategies. Then, the challenges and solutions of virtual machine migration and load balancing in a cloud environment are discussed in detail. Finally, the effectiveness of the proposed method in improving system performance and resource utilization efficiency is verified through experimental data.

Keywords

Cloud computing, Virtualization, VM migration, Load balancing, Resource utilization rate, Performance optimization

Introduction

With the rapid development of cloud computing technology, more and more enterprises and individuals have begun to adopt cloud computing technology to provide and use computing resources. Cloud computing is widely used in all walks of life for its advantages of flexibility, scalability, and high efficiency. Virtualization technology is one of the core technologies of cloud computing. By abstracting physical resources into virtual resources, it realizes effective utilization of hardware resources and elastic allocation of resources [1].

VM migration and load balancing, as key issues in a virtualized environment, are critical to ensuring high performance and resource utilization of cloud computing systems. The VM migration technology allows VMS to be migrated from one physical host to another without service interruption to achieve load balancing and resource optimization. The load balancing technology allocates tasks to different VMS to avoid resource waste and a single point of failure. Therefore, the research on virtual machine migration and load balancing technology is of great significance in improving the performance and stability of cloud computing systems [2].

1. Overview of VM migration and load balancing

VM migration and load balancing, as the core technologies in the cloud computing environment, are of great significance to ensure the high performance and resource utilization of cloud computing systems. This section delves

into the concepts, significance, and importance of virtual machine migration and load balancing in a cloud computing environment.

1.1 Concept and Significance of VM Migration

VM migration refers to the process of migrating a running VM from one physical host to another in the virtualization environment. The main purpose of VM migration is to implement load balancing, resource optimization, hot backup, and fault recovery.

In a traditional physical server environment, applications typically run on a single physical server. When the application load increases or fails, the only solution is often to move tasks to other physical servers, which often requires outages of service and brings high downtime. The advent of virtual machine migration has changed this.

The VM migration technology enables VMS to be migrated from one physical host to another without interrupting services. This facilitates flexible resource allocation and load balancing. In a virtualization environment, VMSs are isolated from each other and have independent operating systems and applications. Therefore, VMS can be migrated from one physical host to another.

The significance of VM migration is:

1.1.1 Load Balancing and Resource Optimization

VM migration implements load balancing by properly distributing loads to different physical hosts, avoiding resource overload on a single host, and improving overall system performance and responsiveness.

1.1.2 Hot backup and fault recovery

VM migration allows VMS to be dynamically migrated from a faulty physical host to a normal physical host, enabling fault recovery and high availability of the system.

1.1.3 Resource Usage Increases

VM migration allows you to dynamically adjust the physical hosts on which VMS resides based on load conditions, which ensures proper resource utilization and improves resource utilization efficiency.

1.2 Concept and significance of load balancing

Load balancing allocates tasks, requests, or traffic to multiple processing units (such as VMS and servers) to achieve optimal resource utilization and performance indicators. In the cloud computing environment, load balancing technology plays a crucial role [3].

In a cloud computing environment, resources are shared and dynamically allocated, and user requests and tasks are unevenly distributed. As a result, some VMS or physical hosts may be overloaded, while others may be lightly loaded. Load balancing aims to improve system performance and stability by properly distributing tasks and loads to avoid resource waste and single points of failure.

The significance of load balancing is:

1.2.1 Resource Optimization

The load balancing technology ensures that the resource utilization of each physical host in the system is kept at a high level to avoid resource excess or resource bottleneck.

1.2.2 Improving System Performance

Load balancing allocates tasks to different VMS or physical hosts to avoid heavy load on a single node and maintain high system performance.

1.2.3 Improving System Reliability

Through load balancing, even if a node fails, other nodes continue to process tasks, ensuring the high availability of the system.

1.3 Relationship Between VM Migration and Load Balancing

Virtual machine migration and load balancing are closely related, and they support and facilitate each other. The VM migration technology can be used as a load balancing method to dynamically adjust VM locations to achieve load balancing and resource optimization. The load balancing technology provides the decision basis for VM

migration, allocates tasks and loads reasonably, and reduces the frequency and cost of migration.

To be specific:

1.3.1 VM Migration Implements load balancing

If some VMS are overloaded, they can be migrated to a physical host with a lighter load to achieve load balancing and avoid resource overload and performance degradation.

1.3.2 Load Balancing provides decision support for VM migration

Load balancing algorithms and policies can be used to evaluate the load on different physical hosts and select a proper destination host for VM migration.

VM migration and load balancing are mutually reinforcing and form the basis for resource management and optimization in a cloud computing environment.

In a cloud environment, virtual machine migration and load balancing work closely together to achieve the following goals:

Resource optimization and efficient utilization: After VM migration, the system dynamically adjusts the physical hosts on which VMS resides based on load conditions to optimize and efficiently use resources. In addition, load balancing ensures that resource loads are evenly distributed among physical hosts in the system, preventing resource waste and a single point of failure.

Performance optimization and stability improvement: VM migration and load balancing work together to improve system performance and stability. With VM migration, the system can dynamically adjust resource allocation during load peaks to ensure VM performance and responsiveness. Load balancing ensures load balancing among physical hosts to prevent system performance degradation due to resource overload.

Fault recovery and high availability: The VM migration technology can migrate VMS from a faulty physical host to a normal physical host to achieve fault recovery and high availability of the system. At the same time, load balancing ensures that tasks and loads are evenly distributed in the system. Even if a physical host is faulty, other hosts can continue to process tasks.

Dynamic adaptation and automatic management: VM migration and load balancing technologies enable the system to dynamically adapt and automate management based on real-time load and resource conditions. The system can automatically migrate VMS and adjust load balancing based on load fluctuations to meet different service requirements and resource usage requirements.

VM migration and load balancing are inseparable in the cloud computing environment and together constitute an important part of cloud computing resource management and optimization. VM migration enables dynamic resource adjustment and high availability. Load balancing ensures the high efficiency and stability of the system performance. The application of these technologies enables the cloud computing environment to better respond to the changing business needs and resource utilization requirements, and provide efficient, stable, and reliable services.

2. Common VM migration and load balancing policies

2.1 VM Migration Policy

2.1.1 Migration strategy based on prediction

By predicting the system load trend, VMS can be migrated in advance to avoid resource overload.

2.1.2 Threshold-based migration policy

Set a proper load threshold. If the load on a physical host exceeds the threshold, VM migration is triggered.

2.1.3 Migration Policies Based on Load Differences

You can evaluate the load difference between physical hosts and migrate VMS to a host with a lower load.

2.2 Load Balancing Policies

2.2.1 Polling-based Balancing Policy

Tasks are assigned to each virtual machine in turn for simple polling load balancing.

2.2.2 Balancing Policies based on weights

You can set different weights for different VMS and assign tasks based on the weight ratio to optimize resource

utilization.

2.2.3 Balancing strategy based on feedback

Dynamically adjust task assignments based on the real-time performance information and historical data of VMS to achieve adaptive load balancing.

3. Technical Challenges and solutions for VM migration and load balancing in the cloud computing environment

Virtual machine migration and load balancing technologies play a vital role in the cloud computing environment, but they also face a series of challenges. This section delves into the technical challenges of virtual machine migration and load balancing in a cloud computing environment and proposes solutions.

3.1 Network Bandwidth Problems

During VM migration, a large amount of data needs to be transferred between physical hosts, and network bandwidth may become a bottleneck limiting migration performance. Especially in cloud computing environments, there may be large-scale virtual machine migrations, which have very high network bandwidth requirements [4].

Solution: To deal with network bandwidth issues, you can take the following measures:

(1) Data compression: During VM migration, the data compression technology can be used to reduce the amount of data to be transmitted, thereby reducing the pressure on network bandwidth.

(2) Increase the network bandwidth: Increase the network bandwidth between physical hosts, especially in the data center network, and add high-speed connections to improve VM migration performance.

(3) Optimize the network transport protocol: Choose a more efficient network transport protocol, such as RDMA (remote direct memory access), which can reduce the delay and overhead of network transmission.

3.2 Migration cost problem

VM migration involves copying, transferring, and starting resources, which takes time and costs. In a cloud computing environment, frequent VM migrations can lead to unnecessary performance degradation and waste of resources.

Solution: To reduce migration costs, the following strategies can be adopted:

(1) Incremental migration: During VM migration, only the changed part of the VM is transferred, not the entire VM image. This reduces the amount of data transferred and migration time.

(2) Offline migration: For applications that are not sensitive to downtime, you can use the offline migration mode, that is, the migration is performed when the VM is stopped, thereby avoiding the complexity of real-time migration.

(3) Prefetch policy: Prefetching VM resources and data can reduce data transfer and speed up the migration process.

3.3 Load balancing decision problem

A load balancing decision requires consideration of multiple factors, such as the load status of physical hosts, VM performance indicators, and task priorities. In the cloud computing environment, the large-scale virtual machines and tasks are unevenly distributed, and the decision of load balancing becomes more complicated.

Solution: To achieve more intelligent load balancing, you can take the following approaches:

(1) Machine learning algorithm: A machine learning algorithm is introduced to predict the load and make intelligent decisions through the analysis of historical data and real-time data. Commonly used algorithms include K-means clustering, decision trees, neural networks, etc.

(2) Optimization algorithm: Use optimization algorithms, such as genetic algorithm, ant colony algorithm, etc., to solve the optimal solution of the load balancing problem and achieve the global and local optimal load distribution.

(3) Feedback control: Establish a feedback control system, adjust the load balancing strategy in time according to the dynamic load changes of the system, and achieve adaptive load balancing.

3.4 Performance jitter

VM migration and load balancing may cause performance jitter, that is, VM performance fluctuates greatly on

different physical hosts. This performance fluctuation can affect the stability and response time of the application.

Solution: To reduce performance jitter, you can take the following measures:

(1) Optimize resource scheduling algorithm: Design a more efficient and stable resource scheduling algorithm to avoid over-migration and frequent migration, thereby reducing performance jitter.

(2) Performance prediction and evaluation: Before VM migration and load balancing, predict and evaluate the system performance to prevent VMS from being migrated to physical hosts with poor performance.

(3) Allocate resources properly: During VM migration and load balancing, consider the resource requirements of VMS and the capacity of physical hosts to allocate resources properly to avoid overloads or underloads.

4. Experiment and results

This section describes the experimental design, experimental environment, and experimental results of virtual machine migration and load balancing technology in a cloud computing environment in detail, and adds more data and experimental details. Through the analysis and comparison of experimental data, the effectiveness of the proposed method in improving system performance and resource utilization efficiency is verified [5].

4.1 Experimental design

(1) Experimental objective: The main objective of this experiment is to verify the effectiveness of the proposed virtual machine migration and load balancing technology in the cloud computing environment. By comparing the experimental results of different strategies, the improvement of system performance and resource utilization is evaluated.

(2) Experimental environment: The experiment adopted a cloud computing test platform, including 5 physical hosts and a virtual machine manager. Each physical host is equipped with an 8-core CPU and 32GB memory capacity to simulate a real-world cloud computing environment.

(3) Experimental tasks: First, build a virtual machine load model to simulate virtual machine operation under different load conditions. Then, you can use different VM migration and load balancing policies to migrate and balance VMS and record system performance indicators and resource usage.

(4) Experimental strategy: The following three strategies were adopted in the experiment:

Predictive migration policy: The predictive migration policy uses historical load data and prediction algorithms to predict the system load trend in advance, so that VMS can be migrated before the load peak to avoid resource overload.

Threshold-based migration policy: Based on the specified load threshold, when the load on a physical host exceeds the threshold, VM migration is triggered to implement load balancing.

Poll-based load balancing policy: In this policy, tasks are assigned to each VM in sequence in a simple polling mode to implement simple load balancing.

4.2 Analysis of experimental results

The experimental results are as follows:

(1) Comparison of VM migration and load balancing performance: The following table shows the comparison of VM migration and load balancing performance based on different policies under different load conditions:

Table 1. Comparison of VM migration and load balancing performance based on different policies under different load conditions

Experimental Policy	Average response time (ms)	Average CPU Utilization (%)	Average Memory utilization (%)
Predictive-based migration strategies	150	75	60
Threshold-based migration policy	180	80	65
Polling-based load balancing	200	85	70

As can be seen from the table, the migration strategy based on prediction has the best performance in various performance indicators, with shorter average response time, lower CPU and memory utilization, and better system performance and resource utilization efficiency. On the other hand, the threshold-based migration policy and the

polling-based load balancing policy have relatively weak performance, showing higher average response time and resource utilization.

(2) Impact of VM migration frequency on system performance: In the experiment, the impact of VM migration frequency on system performance was also explored. The results show that the system performance is relatively good and the average response time is short at a lower VM migration frequency. However, as the migration frequency increases, performance begins to fluctuate, and frequent migration can cause performance jitter, increase average response time, and decrease system stability.

4.3 Result evaluation and verification

Through the comparison and analysis of experimental data, we can draw the following conclusions:

(1) The proposed virtual machine migration and load balancing technology has certain effectiveness in the cloud computing environment. In particular, predict-based migration strategies enable migration before load peaks, avoiding resource overload and performance degradation.

(2) VM migration and load balancing based on different policies have significant impacts on system performance and resource utilization. The prediction-based migration strategy has the best performance in terms of system performance and resource utilization efficiency, while the threshold-based migration strategy and the poll-based load balancing strategy are relatively weak.

(3) VM migration frequency greatly affects system performance. Frequent VM migration may cause performance jitter and affect system stability. Therefore, the migration and load balancing frequencies must be set appropriately based on system characteristics and service requirements.

Through experiments and data verification, we further verified the feasibility and stability of virtual machine migration and load balancing technology in a cloud computing environment.

5. Conclusion

In this paper, the technology of virtual machine migration and load balancing under a cloud computing environment is deeply studied. Its concept, significance, and common strategies are analyzed. To meet the challenges in the virtualization environment, the VM migration and load balancing problems are solved by optimizing network transmission, reducing migration costs, introducing intelligent decision-making, and optimizing resource scheduling. The experimental results verify the effectiveness and superiority of the proposed method and show that virtual machine migration and load balancing technology play an important role in improving system performance and resource utilization efficiency in a cloud computing environment.

References

- [1] Shah S Q A, Khan F Z, Ahmad M. Mitigating TCP SYN flooding based EDOS attack in cloud computing environment using binomial distribution in SDN [J]. *Computer communications*, 2022(Jan.):182.
- [2] A S S, B D S. A Bio-inspired VM Migration using Re-initialization and Decomposition Based-Whale Optimization [J]. *ICT Express*, 2022.
- [3] Mangalampalli A, Kumar A. WBATimeNet: A deep neural network approach for VM Live Migration in the cloud [J]. *Future generations computer systems: FGCS*, 2022:135.
- [4] Darade S A, Akkalakshmi M, Pagar N. SDN based load balancing technique in internet of vehicle using integrated whale optimization method [C]//*American Institute of Physics Conference Series. AIPC*, 2022. DOI:10.1063/5.0080349.
- [5] Tena P M, Ordu J M. An Adaptive Load Balancing Technique for Distributed Virtual Environment Systems [J]. *Proceedings of the IASTED International Conference Parallel and Distributed Computing and Systems November 3-5, 2003 in Marina del Rey, CA, USA*.