A LOAD BALANCING ALGORITHM USING PARTICLE SWARM OPTIMIZATION FOR THE APPLICATIONS OF CLOUD COMPUTING

Rajavenkatesswaran K C[1], Abieshkumar R[2], Arun kumar P[3], Nijanthan M[4], Praveenkumar A[5]

[1]Assistant Professor, Department of Information Technology,
Nandha College of Technology, Perundurai 638 052, Tamilnadu, India
[2,3,4,5]UG Students - Final Year, Department of Information Technology,
Nandha College of Technology, Perundurai 638 052, Tamilnadu, India.

ABSTRACT

Virtual Machines (VMs) in Cloud systems are scheduled to hosts based on their instant resource consumption (e.g., hosts with the greatest accessible RAM), rather than their overall and long-term utilization. Furthermore, the scheduling and placement operations are often computationally intensive and have an impact on the performance of deployed VMs. In this paper, we provide a Cloud VM scheduling method that considers existing VM resource consumption over time by assessing previous VM utilization levels in order to schedule VMs while maximising performance using the PSO technique. Because Cloud management activities such as VM placement have an impact on previously deployed systems, the goal is to minimise such performance deterioration. Furthermore, because overcrowded VMs tend to grab resources from neighbouring VMs, the task enhances the VMs' true CPU consumption. The results reveal that our method refines traditional Instant-based physical machine selection as it learns and adapts to system behaviour over time. The idea of VM scheduling based on resource monitoring data taken from previous resource utilizations (VMs). The PSO classifier reduces the physical machine count by four.

Keywords - Load balancing, Virtual machine, Task scheduling, Particle swarm optimization, IaaS.

1. INTRODUCTION

Cloud computing is the computational paradigm of the future. It is quickly solidifying its position as the future of distributed on-demand computing. Cloud Computing is growing as a critical backbone for various online enterprises by utilising the notion of virtualization. On the other hand, Internet-enabled business (e-Business) is quickly becoming one of the most successful company models in the modern era. To meet the needs of internet-enabled businesses, computing is being turned into a paradigm of commoditized services offered in a way akin to conventional utilities such as water. Customers may access services according on their needs, regardless of where they are housed or how they are provided. Many computing paradigms have pledged to provide utility computing. One such dependable computing paradigm is cloud computing. A front end and a back end comprise the architecture of cloud computing. These two ends are linked through the Internet or an intranet. Client devices such as thin clients, fat clients, and mobile devices are included in the front end. Clients require an interface and apps in order to access the cloud computing infrastructure. The many servers and data storage systems comprise the back end. There is also a "Central Server" server. The cloud system is managed by a centralised server. It also monitors general traffic and responds to client requests in real time.

2. LITERATURE REVIEW

2.1 Identity-Based Remote Data Integrity Checking with Perfect Data Privacy Preserving for Cloud Storage

This work was proposed by Yong Yuet.al. Remote data integrity checking (RDIC) allows a data storage server, such as a cloud server, to demonstrate to a verifier that it is honestly storing a data owner's data. A number of RDIC protocols have been presented in the literature to date, however the most of the constructs suffer from the issue of sophisticated key management, that is, they rely on the expensive public key infrastructure (PKI), which may impede RDIC implementation in practise. In this research, we present a novel identity-based (ID-based) RDIC protocol that uses key-homomorphic cryptographic primitives to decrease system complexity and costs associated with creating and administering the public key authentication framework in PKI-based RDIC schemes. We formalise ID-based RDIC and its security model, which includes protection against a rogue cloud server and zero knowledge privacy from a third-party verifier. Throughout the RDIC procedure, the proposed ID-based RDIC protocol does not provide any information about the stored data to the verifier. In the generic group model, the novel design has been shown secure against the malicious server and provides zero knowledge privacy against a verifier. Comprehensive security research and implementation results show that the suggested protocol is both provably safe and applicable in real-world situations.

2.2 Service Level Agreement in Cloud Computing: A Survey

This paper was proposed by UsmanWaziret.al. Cloud computing makes scattered resources available to people all over the world. Cloud computing has a scalable design that delivers on-demand services to enterprises across several disciplines. Yet, there are several obstacles with the cloud services. In the cloud services, several solutions have been presented to address various types of issues. This study examines the many models presented for SLA in cloud computing in order to tackle the issues that exist in SLA. Problems with performance, customer satisfaction, security, profit, and SLA violations. SLA architecture in cloud computing is discussed. Next, we review existing SLA models presented in various cloud service formats such as SaaS, PaaS, and IaaS. With the use of tables, we explore the benefits and limits of current models in the following section. We summarise and offer a conclusion in the last part.

2.3 A Review of Game-Theoretic Approaches for Secure Virtual Machine Resource Allocation in Cloud

This paper was proposed by PritiNarwalet.al. Cloud computing is a rapidly evolving and dynamic platform that employs virtualization technologies. Virtualization isolates the hardware system resources in software in the Cloud computing environment so that each application may operate in an isolated environment called the virtual machine, and the hypervisor allocates virtual machines to various users who are hosted on the same server. Although it has several advantages like as resource sharing, cost-efficiency, high-performance computing, and lower hardware costs, it also has a variety of security risks. Threats can exist directly on Virtual Machines (VMs) or indirectly on Hyper-visors via virtual machines hosted on them. This paper provides an overview of all potential security risks as well as responses using Game Theoretic techniques. Because of the autonomous and strategic rational decision-making nature of cloud users, where each player competes for the best feasible answer in a safe manner, Game Theory may be employed as a protective mechanism.

2.4 Representing Attribute Based Access Control Policies in Owl

This work was proposed by Nitin Kumar Sharma et al. Attribute Based Access Control (ABAC) models are intended to solve the problems of traditional access control models (DAC, MAC, and RBAC) while integrating their benefits. ABAC provides access control based on generic properties of entities. Many organisational security rules make access decisions dependent on qualities. OWL may be used to officially describe and process security policies represented by ABAC models. With OWL, we created models, domains, data, and security regulations, and then used a reasoner to determine what is permissible. We offer a method for representing the ABAC model using Web Ontology Language in this work (OWL). The EYE reasoner infers the logical link and deduces the access permit for each requested activity to enforce policies. We demonstrated how the Attribute Based Access Control model may be expressed using Web Ontology Language in this study (OWL).

2.5 Auditing a Cloud Provider’s Compliance with Data Backup Requirements: A Game Theoretical Analysis

This work was proposed by ZiadIsmailet al. Cloud computing advancements have created substantial security difficulties in ensuring the confidentiality, integrity, and availability of outsourced data. Typically, a Service Level Agreement (SLA) is signed between the cloud provider and the customer. It is critical to check the cloud provider's compliance with data backup standards in the SLA for redundancy considerations. There are several security methods in place to ensure the integrity and availability of outsourced data. This work can be completed by the client or assigned to an independent body known as the verifier. Nevertheless, evaluating data availability incurs additional fees, which may dissuade customers from doing data verification on a regular basis. Game theory may be used to capture the interaction between the verifier and the cloud provider in order to determine the best data verification approach. This problem is formulated as a two-player non-cooperative game in this paper. We explore the scenario where each piece of data is copied a number of times depending on a set of factors such as its size and sensitivity. We investigate the cloud provider's and verifier's tactics at the Nash Equilibrium and deduce the predicted behaviour of both participants. Lastly, we numerically validate our model using a case study and describe how we assess the model's parameters.

3. EXISTING SYSTEM

Lately, service composition has received considerable attention as a promising paradigm for improving cloud computing data accessibility, integrity, and interoperability. In this paper, we propose an efficient agent-based ant colony optimization (ACO) technique for solving the cloud service composition (CSC) issue. The CSC challenge seeks to meet the complex and demanding needs of enterprises/users in a cloud environment. The complexity of such a scenario is the proliferation of suppliers offering identical services with similar functionality but varied quality of service (QoS) qualities. Because the problem's complexity is NP-hard, which is high, several swarm-based techniques have been proposed to tackle it. To achieve this, a multi-agent based on ACO is suggested and evaluated with four distinct algorithms on 25 different actual datasets. The computational findings on 25 real-world datasets validate the efficacy of the ACO multi-agent distribution mechanism. Furthermore, comparisons with the findings of the four algorithms in the literature show that the multi-agent ACO strategy is comparable with cutting-edge algorithms.

4. PROPOSED SYSTEM

The goal is to offer the idea of VM scheduling based on resource monitoring data taken from previous resource utilizations and to assess previous VM usage levels using two classification techniques such as PSO in order to schedule VMs while maximising performance. The suggested VM scheduling technique improves the VM selection phase by collecting real-time monitoring data and analysing physical and virtual resources. Our goal is to improve VM scheduling by including factors relating to real VM use levels, so that VMs may be deployed while reducing the penalization of overall performance levels. The optimization approaches incorporate analytics on previously deployed VMs in order to (a) maximise utilization levels and (b) minimise performance losses. Users have underused VMs and do not have the same resource usage pattern throughout the day. Lastly, cloud management activities such as VM placement have an impact on previously deployed systems (for example, performance loss in a database cluster) since heavily loaded VMs tend to steal CPU time from neighbouring VMs. These are basic examples that show the need for more precise VM scheduling that might increase performance. VMId, CPU, RAM, and BW are the input datasets. The cloud, which generates CPU, Memory, and bandwidth. The VM machine of the allocation has been allocated to the host of that particular host id. The VM is migrated so that the assigned host may be scheduled.

4.1 VM Scheduling

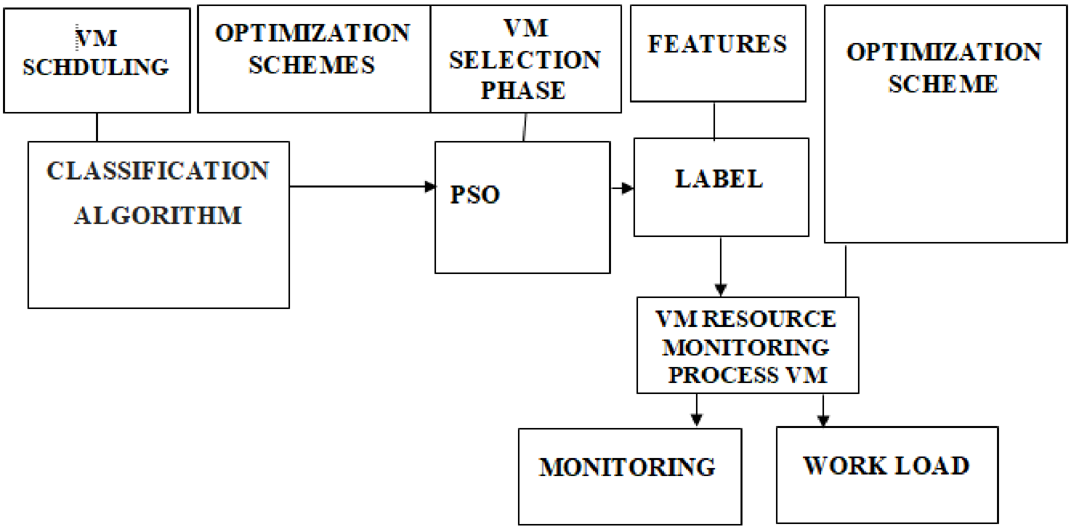
The proposed technique improves the VM selection phase by collecting real-time dataset monitoring data and analysing physical and virtual resources. Our goal is to improve VM scheduling. To add criteria relating to real VM use levels, so that VMs may be put while reducing the penalization of overall performance levels. The optimization approaches use analytics to previously deployed VMs in order to (a) maximise utilization levels and (b) minimise performance losses. A monitoring engine that collects data about VM resource utilization monitoring online. The engine may gather system data at regular intervals and store it in an online cloud service where it can be processed. Data is gathered at regular intervals (e.g., every 1 second) and saved in a temporary local file.

4.2 Classification Algorithm

When considering supervised machine learning algorithms for classification, the input dataset should be labelled.

4.3 Particle Swarm Optimization

Particle swarm optimization (PSO) is a basic bio-inspired technique for searching for an optimal solution in the solution space. It differs from other optimization techniques in that it requires only the objective function and is not affected by the gradient or any differential form of the goal. It also features a small number of hyper parameters. Using an example, you will discover the logic for PSO and its algorithm in this tutorial.



After finishing this tutorial, you will understand:

What is a particle swarm and how does it behave in the PSO algorithm?

What types of optimization issues can PSO solve?

How to Use Particle Swarm Optimization to Solve a Problem

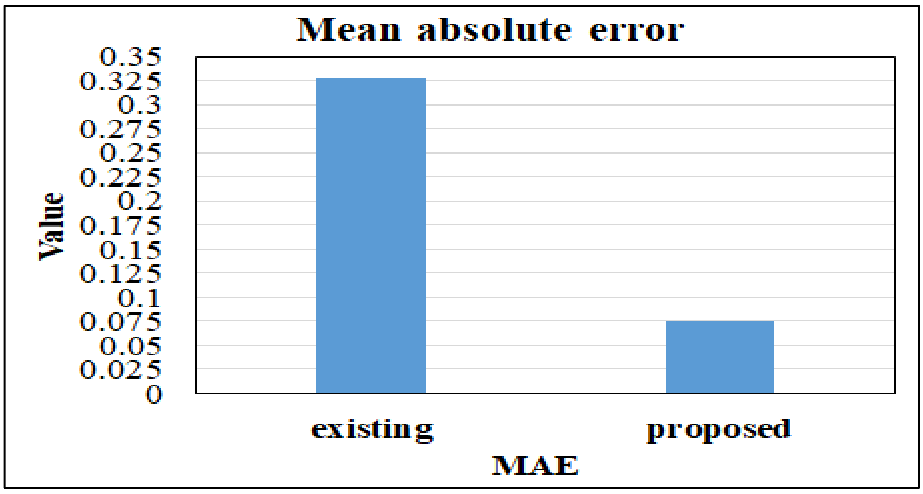
PSO is most useful for determining the maximum and minimum of a function specified on a multidimensional vector outputs a real value from a vector parameter (such as a plane coordinate) and can take on nearly any value in space (for example, is the altitude and we can locate one for any point on the plane), then we can use PSO. The PSO algorithm will return the parameter that produced the lowest result.

4.4 Optimization Scheme

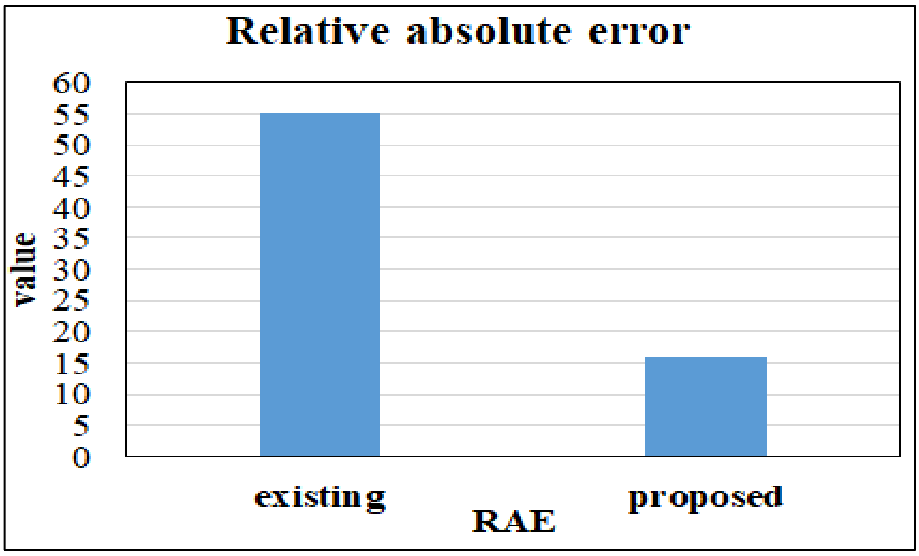
The goal of these optimization approaches is to define the weight of the VM based on the VMs' resource utilization. This will give information about the state of previously deployed VMs, such as whether or not a workload is executing. We present two optimization strategies to do this. The PSO is used to classify the VM state in terms of its current resource utilization. The virtual machine resource utilization information is initially gathered and monitored, and the acquired data is subsequently categorized using machine learning algorithms such as PSO.

5. RESULTS AND DISCUSSIONS

The emphasis is on cloudsim, an open-source programme for creating private and public clouds. Cloudsim’s default setup involves installing virtual machines by picking the host with the greatest available ram until the number of virtual machines exceeds the limit.



Virtual Machines (VMs) are assigned to hosts based on their immediate resource consumption (e.g., hosts with the greatest available RAM), without regard for their overall and long-term utilization. Furthermore, the scheduling and placement operations are often computationally intensive and have an impact on the performance of deployed VMs. As a result, the standard VM placement method does not take into account previous VM resource use levels.



This is addressed by implementing a VM scheduling method. The notion of VM scheduling based on resource monitoring data derived from previous resource utilizations (including VMs and VMs), and the resource data are categorised using the optimization techniques PSO, thus scheduling is performed. The programme analyses previous resource consumption levels and classifies them based on overall resource usage. Finally, a list of possible hosts is generated, and the resources are sorted accordingly. In detail, VMs are re-ranked using this method based on the optimization technique chosen and their VM consumption. For example, we utilise resource information from a 24-hour monitoring as the data set and a seven-day resource consumption monitoring as the training set. The analytics are (a) based on usage levels throughout time, which are classified as low, medium, and heavy, and (b) based on on-going data (e.g. memory percent that increases over time). The method applies a weighing mechanism to the selected VMs based on several characteristics (e.g. CPU, RAM percentage).

6. CONCLUSION

Various virtual machine placement techniques were employed for scheduling by selecting physical computers based on system statistics (i.e. CPU, memory, and bandwidth utilization) in a cloud system. The current VM placement does not consider real-time VM resource use levels. In this section, we present a novel VM placement technique based on previous VM usage experiences. The VM consumption is then observed, and the data is trained using machine learning models (PSO) to anticipate VM resource utilization and arrange VMs accordingly. It was proposed an algorithm that permits VM placement based on PM and VM consumption levels, as well as a computational learning approach based on the notion of assessing previous VM resource utilization based on historical data to optimise the PM selection phase. A virtual machine placement technique based on real-time virtual resource monitoring was introduced, with machine learning models used to train and learn from prior virtual machine resource utilization. As a result, a monitoring engine providing resource utilization statistics is assumed. Using the PSO classifier instead of the Support Vector Machine (SVM) classifier reduces the physical machine count by four. The work was completed using ten virtual machines.

7. REFERENCES

[1] Y. Yong, M. H. Au, and G. Ateniese, Identity-based remote data integrity checking with perfect data privacy preserving for cloud storage, IEEE Transactions on Information Forensics and Security, vol. 12, no. 4, pp.767–778, 2017.

[2] U. Wazir, F. G. Khan, S. Shah, Service level agreement in cloud computing: A survey, International Journal of Computer Science and Information Security, vol. 14, no.6, p. 324, 2016.

[3] P. Narwal, D. Kumar, and M. Sharma, A review of game-theoretic approaches for secure virtual machine resource allocation in cloud, in Proceedings of the Second International Conference on Information and Communication Technology for Competitive Strategies, 2016.

[4] N. K. Sharma and A. Joshi, Representing attribute based access control policies in owl, in 2016 IEEE Tenth International Conference on Semantic Computing (ICSC), 2016, pp. 333–336.

[5] Z. Ismail, C. Kiennert, J. Leneutre, and L. Chen, Auditing a cloud provider’s compliance with data backup requirements: A game theoretical analysis, IEEE Transactions on Information Forensics and Security, vol.11, no. 8, pp. 1685–1699, 2016.

[6] E. Furuncu and I. Sogukpinar, Scalable risk assessment method for cloud computing using game theory (CCRAM), Computer Standards & Interfaces, vol. 38, pp.44–50, 2015.

[7] J. Li, J.W. Li, and X. F. Chen, Identity-based encryption with outsourced revocation in cloud computing, IEEE Transactions on Computers, vol. 64, no. 2, pp. 425–437, 2015.

[8] X. Yi, F. Y. Rao, and E. Bertino, Privacy-preserving association rule mining in cloud computing, in Proceedings of the 10th ACM Symposium on Information, Computer and Communications Security, 2015, pp.439–450.

[9] J. Lou and Y. Vorobeychik, Equilibrium analysis of multi-defender security games, in Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence (IJCAI), 2015, pp. 596–602.

[10] M. Nabeel and E. Bertino, Privacy preserving delegated access control in public clouds, IEEE Transactions on Knowledge and Data Engineering, vol. 26, no. 9, pp.2268–2280, 2014.