Tuijin Jishu/Journal of Propulsion Technology
ISSN: 1001-4055
Vol. 44 No. 5 (2023)

A Complete Study on Cloud Virtual Machine for Data Analyses

[1]Poojitha G S , [2]Dr.puttegowda D

[1]Research Scholar, Department of Computer Science and Engineering, ATME College of Engineering,
Visvesvaraya Technological university, Belagavi, India

[2] Professor, Department of Computer Science and Engineering, ATME College of Engineering,
Visvesvaraya Technological university, Belagavi, India

Abstract:

In today's data processing era, cloud computing has become the essential technology for storing, processing and distributing large amounts of data. Scheduling is an important part of cloud computing and plays an important role in integrating distributed computing resources and managing distributed workloads. This article provides an overview of the key concepts and challenges in planning data distribution projects in the cloud.

Cloud computing uses virtual devices to reduce server power costs and hardware costs. Additionally, large data sets allow users to access multiple stores. To get the most out of cloud computing, developers need to develop a process that optimizes the architecture and deployment of usage models. The role of virtual machines (VMs) has become an important topic as cloud computing tools are made available through virtualization technology. To improve the overall performance of cloud computing, virtual machines need to be carefully managed so that they can run efficiently and allocate physical resources correctly.

There is a lot of SLA based big data processing and MapReduce scheduling in the cloud environment, but they do not dynamically configure the cloud environment. Instead, they create a virtual cluster by first provisioning (static) resources in a private cloud. We believe that cloud services should be provisioned dynamically based on the application's workload and data size. This introduces new challenges such as: a) how much and what type of air to provide; b) private producers or public air providers selected for demand with budget and time constraints; c) provide information on which fields and resources should be selected to minimize data transmission and processing costs.

1. Introduction

In today's data driven world, organizations rely on cloud computing to manage and process large amounts of data efficiently and cost-effectively. An important aspect of cloud-based deployment management is task scheduling. Task scheduling involves allocating computing resources and workloads in the cloud environment to ensure efficient processing of heavy workloads and efficient use of resources.

Cloud data distribution generally involves storing and processing data on a cluster of servers and data centers. These categories form the backbone of many applications, from big data analytics and machine learning to content distribution and real-time operations. Planning in such an environment is critical to achieving high performance, scalability and reliability. The main purpose of job scheduling in cloud distributed data is to effectively balance computing resources, reduce job completion time, and complete data on time. The job is not without its challenges.

Cloud environments are inherently characterized as allowing resources to be assigned and removed on-demand. Additionally, data can be distributed across different regions, making data entry time critical in making planning decisions. Planning tasks in the cloud data distribution environment is important for improving resource utilization in the cloud environment and increasing data efficiency and distribution. The main goal is to effectively distribute computing resources and tasks to meet performance, scalability and reliability. Below is a summary of the main points:

2. Related Work:

Scheduling is an important issue in Web server deployment, which has a strong impact on network connectivity and quality of service. Previous research has explored this area from various angles to achieve a network connectivity for efficient deployment. However, the rapid development of heterogeneous cloud computing has brought many new challenges to the network's ability to manage, develop and maintain. Solving network capacity in heterogeneous cloud systems has become an important issue in ensuring network security.

Wang Yongming and colleagues[1] worked on hop-recursive local search. Getting a jump start on local research steps will allow you to delve even lower into these valleys. Some workshop planning problem evaluation tests demonstrate the value and efficiency of cross-local research. Swachil J. Patel [2] proposed and implemented the development of the importance of task scheduling algorithm based on multi-modal and multi-attribute decision making in cloud computing and was successful. Using the reverse process to enhance the importance of mission planning in the cloud - Potu Narayana[3], this study examined the classification of mission planning and the traditional methods used to plan airborne operations and their benefits and limitations. This detailed analysis will help researchers find better solutions to the problems cloud computing currently faces in big data processing. Many researchers and investigators have proposed and implemented a method for scheduling tasks based on improved Monte Carlo tree search [4].

Experimental results show that this algorithm is effective in scheduling edge jobs. This algorithm can reduce the probability of premature growth, improve reverse search performance, and obtain more non-dominated solutions than other algorithms. Some research on virtual machine selection strategy based on Gray Wolf optimizer in cloud environment [6] is a great way to introduce the GWO algorithm based on the results. This algorithm will work best in virtual machine selection. Parinaz Haraian et al. [7] proposed an adaptive fuzzy control framework (AFRM) to study the adaptive fuzzy control system in cloud computing; here the final resources of each virtual machine are collected from environmental sensors and referenced to fuzzes. controller, controller. AFRM then analyzes the data received to determine how to allocate resources at each stage of the transition process. Chen Mingzhe [8] conducted a research on wireless virtual reality (VR) under the research title Information Communication-Information Management in Wireless Virtual Reality (VR): A Change in State Echo. determine this location. Data correlation to better manage uplink and downlink traffic. This relationship between capacity information can be taken into account in resource allocation to reduce uplink and downlink traffic. DYVINE: Deploying a Fitness-Based Dynamic Virtual Network in Cloud Computing, by Chinmaya Kumar Dehury [9].

The VNE algorithm allows the virtual network to be dynamic; This means that the structure and requirements may change during execution. Data analysis and line analysis of air operations - Abdullah Alenizi [10]. In this work, the authors apply the results of this approach to multiple queued processes and use the well-known pollaczek-khintchine model to estimate the average waiting time and completion time of job submissions. We then use different scheduling methods, Shortest Job First (SJF) and Longest Job First (LJF), and compare them with First Come First Served (FCFS).

Build performance across memory-intensive, compute-intensive, and I/O-intensive workloads across different virtual machine configurations. Due to the upper limit of memory bandwidth and I/O bandwidth and the presence of multiple synchronization overhead, two virtual machines with different parameters will have the same performance. For a given workload, two virtual machines with different configurations can provide the same performance. This research believes that all research done on the virtual machine is the same for this study. This study also proposes a performance estimation method to avoid excessive costs in comparing virtual machines. Performance prediction across a set of virtual machines.

3. Conclusion:

There are obvious changes in the licensing of cloud and virtual machines. To realize the full potential of these systems, features such as live migration are important for both the performance and usability of the system. But immigrant security is still in its infancy. Therefore, this paper uses the X.805 standard to examine the security issues of instant migration. In this context, eight security threats were analyzed and discussed. This article also discusses some suggestions to ensure a quick transition. In the discussion, it was emphasized that there are different claims regarding different security threats; But a good approach to all these problems is not

yet ready. Therefore, our future work is to create a good security system to deal with the security threats mentioned in this article. Overall, we believe that residential migration is a useful tool to improve living and working conditions. But the transition itself creates new security issues that need to be resolved before widespread use.

Virtual machine consistency: if a workload has the same or similar performance characteristics in different virtual ma- chines size, we think that the performance of the two virtual machines is consistent relative to this workload.

Virtual machines with more resources, its performance may not improve too much. For example, for memory-intensive workloads, due to the limitation of memory bandwidth, in-creasing the number of VCPU of virtual machine after reaching the maximum memory bandwidth will not improve the performance of workload too much. Therefore, it is possible that the performance of virtual machines with different configurations is consistent relative to some workloads. So, the performance of virtual machine is related to the workload running in the virtual machine, which makes it meaningful to compare the performance consistency of virtual machine. When the performance of multiple virtual machines is the same, we can choose the virtual machine with the lowest cost

References:

[1] Yung Ming WANG, Hang Li YIN Iterative Local Search with Leap-frog Steps for Job Shop Scheduling Problems, 2013 IEEE.

[2] Swachil J. Patel, Upendra R. Bhoi Improved Priority based Job Scheduling Algorithm in Cloud Computing using Iterative Method, 2014 Fourth International Conference on Advances in Computing and Communications

[3] Potu Narayana, Prof. Premchand Parvataneni A Research on Various Scheduling Strategies in Fog Computing Environment, 2020 International Conference on Emerging Trends in Information Technology and Engineering.

[4] Haobin Wang, Chen Li, A Job Scheduling Algorithm for Edge Computing Based on Modified Monte Carlo Tree Search, 2021 IEEE.

[5] Jiao Xuan, Li Chengyang, Research on Multi-objective Fuzzy Flexible JobShop Scheduling Based on Cloud Computinger, 2020 IEEE.

[6] Priyanshi Singh, Murtaza Abbas Rizvi, Virtual machine selection strategy based on Grey Wolf Optimizer in Cloud Environment: A study, 2018 8th International Conference on Communication Systems and Network Technologies.

[7] Parinaz Haratian, An Adaptive and Fuzzy Resource Management Approach in Cloud Computing, 2020 IEEE.

[8] Mingzhe Chen, Walid Saad, Data Correlation-Aware Resource Management in Wireless Virtual Reality (VR): An Echo State Transfer Learning Approach, 2019 IEEE.

[9] Chinmaya Kumar Dehury, Prasan Kumar Sahoo, Senior Member, IEEE, DYVINE: Fitness based Dynamic Virtual Network Embedding in Cloud Computing,2019 IEEE.

[10] Abdullah Alenizi, Reda Ammar, Queue Analysis for Probabilistic Cloud Workflows, 2020 IEEE International Symposium on Signal Processing and Information Technology.

[11] M. Aiash, G. Mapp, A. Lasebae, R. Phan. Providing Security in 4G Systems: Unveiling the Challenges. Sixth Advanced International Conference on Telecommunications. 2010.

[12] M R. Anala, J. Shetty, G. Shobha. A Framework for Secure Live Migration of Virtual Machines. International Conference on Advances in Computing, Communications and Informatics. 2013.

[13] D. Perez-Botero. A Brief Tutorial on Live Virtual Machine Migration From a Security Perspective.

[14] S. Biedermann, M. Zittel and S. Katzenbeisser. Improving Security of Virtual Machines during Live Migrations. Eleventh Annual Conference on Privacy, Security and Trust (PST). 2013.

[15] J. Oberheide, E. Cooke, F. Jahanian.Empirical Exploitation of live migration of virtual machines. Proc of Black Hat DC, March 24, 2008.

[16] A. Back, U. Mller, and A. Stiglic. Traffic Analysis Attacks and TradeOffs in Anonymity Providing Systems. Information Hiding, volume 2137 of Lecture Notes in Computer Science, page 245-257. Springer, 2001.

[17] T. Ristenpart, E. Tromer, H. Shacham, and S. Savage.Hey, You, get off my cloud: exploring information leakage in third-party compute clouds. In Proceedings of 16th ACM conference on computer and communication security, 2009, pp 199-212

[18] Cisco, Inc, Cisco IOS Software Configuration Guide, Release 12.2SX (VLAN Configuration Guidelines and Restrictions). http://www.cisco.com/en/US/docs/switches/lan/catalyst6500/ios/ 12.2SX/configuration/guide/vlans.pdf. [Last Accessed 04-12-13].

[19] Juniper Networks, Inc. Alternatives for Securing Virtual Networks: A Different Network Requires a Different ApproachExtending Security to the Virtual World. white paper 1000220-012-EN Dec 2011, .

[20] C. Xianqin, G. Xiaopeng, W. Han, W. Sumei, L. Xiang. ApplicationTransparent Live Migration for virtual machine on network security enhanced hypervisor . Research paper. China Communications. Page 32 42, 2011.

[21] Trusted Computing Group. http://www.trustedcomputinggroup.org /resources/tpm main specification.

[22] S. Berger, R. Caceres, K. A. Goldman, R. Perez, R. Sailer, and L. Doorn, Virtualizing the trusted platform module. in In USENIX Security, pp. 305320, 2006.

[23] Xen Project. http://www.xenproject.org/.[Last Accessed 05-12-13].

[24] T. Dierks, E. Rescorla. The Transport Layer Security (TLS) Protocol Version 1.2. Request for Comments: 5246 . 2008.

[25] W. Wang, X. Wu, B. Lin, K. Miao, X. Dang, Secured VM Live Migration in Personal Cloud. In Proceedings of ICCET, China, 2010.

[26] P. Dewan, D. Durham, H. Khosravi, M. Long, and G. Nagabhushan. hypervisor-based system for protecting software runtime memory and persistent storage. In Proceedings of the 2008 Spring Simulation Multiconference (Ottawa, Canada, April 14 - 17, 2008).