JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p. 3644-3646
https://publishoa.com
ISSN: 1309-3452

"A study on deadline-aware priority scheduling (daps) model tasks scheduling with virtual machines in cloud computing"

Ph.D Scholar: - Rituraj

Subject: - Computer Science

Email. Id:- jainrituraj@yahoo.com

UNDER THE SUPERVISION OF

Dr. Manish Varshney

Dept. of Computer Science

MUIT University, Lucknow (U.P.)

Email. Id:- itsmanishvarshney@gmail.com

Abstract

Cloud computing refers to both the hardware and system software in the datacenters that provide the applications delivered as services via the Web. The services themselves have long been referred to as Software as a Service (SaaS). The hardware and software that make up the datacenters are what we will refer to as a cloud. When a cloud is made available to the general public on a pay-as-you-go basis, we refer to it as an open cloud. The service being offered is utility computing. You can create and improve the policies that will be used by all of the CloudSim components with the help of the toolkit. As a result of its ability to mimic the complexity that manifests itself in many circumstances, it will be viewed as a helpful study aid.

Keywords: Cloud Computing, Load, balance, task, scheduling, Deadline-Aware Priority Scheduling (DAPS) model, etc

1. INTRODUCTION

Cloud computing refers to both the hardware and system software in the datacenters that provide the applications delivered as services via the Web. The services themselves have long been referred to as Software as a Service (SaaS). The hardware and software that make up the datacenters are what we will refer to as a cloud. When a cloud is made available to the general public on a pay-as-you-go basis, we refer to it as an open cloud. The service being offered is utility computing. We use the term "Private Cloud" to refer to internal datacenters of a company or other organization that are not open to the general public. In this way, SaaS and Utility Computing together make up Cloud Computing; however Private Clouds are not included.

Task scheduling with a sense of load balance

The performance of the task scheduling process is significantly impacted by load balancing, another important restriction in cloud settings. It is important to make sure the resources are not in any manner overcrowded or under loaded, even throughout the job scheduling process. Additionally, because the bulk of the resources are anticipated to be dispersed simultaneously, the scheduling system must ensure that during the allocation process, all of the resources are balanced in accordance to their capacity. This helps to increase resource utilization because the problem isn't just overloaded, idle VMs; it's also under loaded, idle VMs. The revenue of the service provider is impacted when virtual machines (VMs) are idle. Therefore, an efficient load balance aware task scheduling algorithm must be included as a part of the task scheduling approach in order to make better use of the resources that are currently accessible. There are many ways to balance load, but because the cloud computing architecture is entirely real-time oriented, dynamic load

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p. 3644-3646
https://publishoa.com
ISSN: 1309-3452
balancing is typically preferred to static load balancing. There are two possibilities for dynamic load balancing: either it is done at the same time as task scheduling, or it is done after task scheduling.

LITERATURE REVIEW

Diwakar Ramanuj Tripathi, M.D. (2019) Different primary foci for the arrangement of applications with increasing data are ensured by cloud computing. A pay-as-you-go business show's reduced expense is one crucial assurance. Another guarantee is (generally speaking) limitless throughput by adding workers as the responsibility grows. This paper lists optional developments that have an impact on cloud computing for database applications and reflects on the findings of an in-depth analysis of previous corporate cloud advantages that have benefited from these strategies. In contrast to evaluation or OLAP jobs, which have recently received a lot of attention, the focus of this task is on trade dealing with (i.e., peruse and revitalise jobs). The outcomes are unexpected.

Gawali (2018) Cloud computing is a necessity in today's technologies. Critical elements of cloud computing include the allocation of resources and the scheduling of work. The modified analytic hierarchy process (MAHP), bandwidth aware divisible scheduling (BATS) + BAR optimization, longest anticipated processing time preemption (LEPT), and divide-and-conquer approaches are all included in this paper's heuristic approach to task scheduling and resource allocation. In this approach, each job is processed before being allocated to cloud resources, which is achieved by using an MAHP process. The bandwidth and load of the cloud resources are taken into consideration as constraints when assigning the resources using a combined BATS + BAR optimization technique. The proposed system, which is an expansion of the existing one, also preempts resource-intensive tasks via LEPT preemption. The divide-and-conquer strategy enhances the suggested system when turnaround time and response time are employed as performance indicators, as empirically shown through comparison with the current BATS and enhanced differential evolution algorithm (IDEA) frameworks.

Mitrevski and others (2017) New and intriguing in the IT sector is the term "cloud computing." The concept of combining computation and capacity in allocated data leads to the term "cloud computing." Its long-term goals are to provide the cloud client with a flexible, on-demand bundle, providing him much more possibility, adaptability, and dependability at the same time. It accomplishes all of the above by using a simple "utility computing paradigm." On-request estimating, fewer IT overhead, and the ability to quickly increase IT here and there are all promised benefits. This work focuses on transaction processing programmes that operate in multi-processor and cloud environments.

Study of the Objectives

- To investigate the role of load balance awareness in work scheduling.
- To experiment with DAPS model tasks using different VMs.

RESEARCH METHODOLOGY

It will be necessary to have a reliable tool for modelling job scheduling in both homogeneous and heterogeneous cloud computing settings. This tool must be adaptable enough to accommodate this environment and the growing quantity of user requests. The simulator-based toolkit will be housed in the Java programming language and run on an event-driven approach. You can create and improve the policies that will be used by all of the CloudSim components with the help of the toolkit. It will therefore be recognised as a valuable study.

Datasets

Depending on the applications being used, different datasets are used in the cloud computing environment. It will be used in this study to construct two different types of datasets: one that will be collected directly from the NASA website for research and another that will be produced using the CloudSim tools.

Recommendations

The initial scheduling model will be the Deadline-Aware Priority Scheduling (DAPS) model. With the help of this model, jobs will be scheduled and allocated to the right resources in order to cut down on the time needed to finish them by the deadline and the lowest completion time. The degree of user satisfaction with the system is one of the

factors that must be taken into account when we are designing an effective task scheduling model. The method will focus on scheduling tasks based on a deadline restriction, and as a result, the deadline will be taken into account as the amount of time needed to finish the activity.

CONCLUSION

When the deadline restriction is met, the Virtual Machine's (VM) state is deemed successful according to the DAPS model's procedure, which involves ranking tasks according to length priority in ascending order. After that, the jobs are assigned to the proper VM while minimizing the make span and completion time. In order to satisfy users, the DAPS model seeks to obtain the best performance by reducing metrics like average makespan, mean of total average reaction time, number of violations, and violation ratio. When a user is primarily concerned with task costs, such as the cost of processor, memory, bandwidth, and storage depending on the budgetary constraints, the DAPS model also seeks to maximize resource consumption and guarantee ratio.

REFERENCES

1. Framework for Resource Management in Cloud Computing, Gagandeep Kaur (2021).
2. Using machine learning to automate resource management in cloud computing environments, Shaw, Rachael (2021),
3. Harvinder Singh et al., "Cloud Resource Management: Comparative Analysis and Research Issues," International Journal of Scientific and Technology Research, Volume 9, 2020 Issue 06
4. "Log as a Secure Service Scheme (LASS) for Cloud", MohdAmeen Imran et al.
5. A survey on cloud computing security: Issues, challenges, and solutions, Singh, Saurabh, Young-SikJeong, and Jong Hyuk Park [5]. Journal of Network and Computer Applications 75 (2016): 200-222.