ADVANCED HETEROGENEOUS EARLIEST FINISH TIME METHODOLOGY IN A CLOUD ENVIRONMENT FOR TASK SCHEDULING

Sharad Awasthi, M.Tech Scholar, Department of Computer Science & Engineering, Kanpur Institute of Technology, Kanpur, India.
Nikita Gupta, Assistant Professor, Department of Computer Science & Engineering, Kanpur Institute of Technology, Kanpur, India.

Abstract— **Cloud computing has fundamentally revolutionized the information technology industry by making it possible to provide users and customers with services that are delivered through the internet. These services can be broken down into two distinct classifications: software services and hardware services. As a direct consequence of this, both the costs associated with the development of physical resources and the expenses associated with delivering the essential software licenses are decreased. The problem of task scheduling is one of the most major and noticeable challenges that the cloud computing system must overcome. This challenge must be tackled using a range of approaches, with the first step being an improvement in the scheduling of jobs within the Improved –IHEFT.**
Cloud computing is presently the preeminent technology in the field of high-performance distributed computing. It offers resource polling and on-demand services via the internet. Cloud computing has taken over as the main technology in this space. As a direct consequence of this, task scheduling has developed into an important study topic in the field of cloud computing as a direct consequence of the fact that the service requirements of users change on a regular basis. It is impossible to distribute the work in an effective manner because the Heterogeneous Earliest Finish Time (HEFT) makes it impossible. We make certain adjustments to the Improved HEFT algorithm in order to reduce the amount of power that the programme needs to run and to ensure that the burden is distributed evenly across all of the processors.

Index Terms— Improved -HEFT, Cloud Computing, Task Scheduling,Min-Min concept

I. INTRODUCTION

Computing on the cloud has quickly become one of the most fascinating specialized sectors to appear in the modern age, and it is just getting started. It has been established that it has an effect on the storage of data, information technology, the design of programming, and the organizational structures of businesses. The following is how the National Institute of Standards and Technology (NIST) defines cloud computing: "the cloud computing is a paradigm to offer access for resources pooling, convenience, on-demand, and ubiquitous delivery that can be conveniently delivered with multiple forms of service provider interaction" [1].
Cloud computing is a new paradigm in the world of computing, and it is commonly regarded as the most rapidly developing new innovation. As a result of its rapid development and the increasing number of new customers and suppliers that it attracts as a result of this, cloud computing is rapidly becoming a more popular option. The rapid development of cloud computing is being accelerated even further by the increasing number of computing breakthroughs, which are being developed at reasonable and reasonable costs in terms of infrastructure and capacity capabilities. This development is taking place at a time when cloud computing is also being developed at reasonable and reasonable costs.
The term "cloud computing" refers to the practice of storing and transmitting data over the Internet rather than using the more conventional way of reading the data directly from the hard drive of the computer. The concept of cloud computing can be traced back to the days when flowcharts and presentations were utilized as modes of communication with the computers that acted as the backbone of the Internet. Local storage and computing are the locations at which data is stored or projects are maintained operating from a hard drive. This enables clients to have simple and speedy access to data and information that is stored locally. It is not necessary for there to be cloud computing services or applications in operation in order for there to be dedicated hardware server equipment placed in a home. Either the information should be able to be accessed through the use of the Internet, or it may be necessary for the information to be synchronized with data that is able to be accessed through the use of the Internet. The term "cloud computing" refers to the process of outsourcing computing operations in order to take advantage of the scalability of surplus assets, the availability of services on demand, and the reduction in associated expenses. One of the most significant benefits of cloud computing is the reduction in the need for substantial upfront investments in new or upgraded information technology infrastructure.
The following is a definition of cloud computing taken from [2] and [3]: "The cloud" is a type of parallel and distributed system that is made up of a collection of interconnected and virtualized personal computers. These PCs are gradually provisioned and

introduced as at least one centralized computing asset based on administration level understandings that are established through an agreement between the service provider and their customers. [09] The use of computing done on the cloud is often regarded as being of great benefit to smaller businesses. Additionally, it enables them to make investments in emerging technology, which wasn't conceivable in the past. Furthermore, it gives them power by helping them compete against other independent businesses or even larger organizations. Because of this arrangement, the company will spend less money to have someone come in and settle or introduce an application, which will result in the organization having more money left over. Consequently, making use of apps that are hosted in the cloud is more cost-effective than purchasing a wide selection of separate software programmes. Having access to a single cloud that hosts many applications satisfies the needs of each and every employee at the company.

Because of the application programming interface (API), which is used to determine which cloud-based programmes are suitable for the objectives of the business, the integration of these programmes into the company's operations will go off without a hitch. Because cloud computing is regularly updated, the company does not need to pay money to stay up with the times in order to maintain its competitive advantage. The use of cloud computing offers businesses the opportunity to reduce the amount of money they spend on operational expenses [4-6].

The utilization of globally empty resources is required in order to raise the utilization rate and earnings from resources by improving the economic efficiency of these resources; the cloud model is the most appropriate for accomplishing this objective because it allows for the greatest amount of flexibility and scalability. The main objective of the notion of cloud computing is to make its underlying data and resources accessible to the greatest number of users as feasible. The following is a definition of a platform that is used for offering services and applications to its customers: Software as a service (SaaS), platform as a service (PaaS), and infrastructure as a service (IaaS) are the three distinct categories of cloud computing services that are now on the market [7, 8].

Customers are charged on a pay-per-use basis for access to a wide range of services, including shared computer resources, servers, data storage, applications, and networks. The following are some examples of such services: Software as a Service (SaaS) refers to a model of online service delivery in which consumers obtain software licenses in proportion to the amount of money they pay for the service. Any computer with an internet connection and a web browser installed can use the aforementioned services. The user is able to develop his/her own services using cloud-based services that are currently available through the PaaS platform, and then deploy those services to their own machine. Within the framework of IaaS, consumers have access to the organizational infrastructure they require through the use of the internet. In order for the customer to make use of the infrastructure, it is not necessary for them to have prior knowledge of its internal architecture. Customers rent the necessary infrastructure on an as-needed basis rather than acquiring it entirely to meet the needs of their businesses. When their companies no longer have a demand for the infrastructure, the customer keeps the money that was paid for the services and puts it to other use. In recent years, there has been an increase in the number of people using cloud computing services, which has resulted in an increase in the number of jobs that need to be managed propositionally. As a result, task scheduling has become an absolute necessity.

The remaining parts of the paper are structured as described in the following paragraphs. In Section II, you'll find the introduction to the literature review that was written. In Section III, the technique for Task Scheduling is broken down into its component parts in greater depth. The analysis of the results is offered in the fourth section. The last part of the paper is Section VI, which wraps everything up.

II. Literature Review

The development of the technology known as cloud computing is predicated on the utilization of distributed systems, which is one of the fundamental components of computing. In order for cloud computing to function properly, it is required to make use of dispersed resources that may either be assigned or de-allocated on the fly [8]. The concept of computing done via the internet in the cloud is a relatively new one. It is rapidly growing in size as a result of the significant growth in internet-based services over the past few decades. This technology is distinguished from others in that it not only offers services to customers but also makes it possible for customers to share cloud resources with one another. The user's activity is taken into account to determine the appropriate amount of payment for the resources that have been depleted.

Conventional scheduling strategies were presented by Pop, Florin, and others [34]. These techniques took into account, among other things, the processing of asynchronous tasks with several queues for aperiodic and periodic activities. In addition, these techniques were offered. When determining the number of resources necessary to plan a collection of aperiodic tasks, both the expenses associated with the tasks' execution and the data transfer associated with those tasks were taken into consideration. The time limit was the most important factor that had to be considered when developing the optimization metric.

In the research paper that Peng, Zhiping, and their colleagues published in 2016 [36], they offered a model for a system that consisted of three parts: a portal, a job scheduler, and a resource pool. They developed a novel method for work scheduling that is based on reinforcement learning by evaluating the execution process of user jobs in order to minimize the makespan, average waiting time, and virtual machine resources within the constraints of a deadline constraint and virtual machine resources. This method minimizes the makespan, average waiting time, and virtual machine resources.

The Adaptive Two-Stage Deadline constrained Scheduling (ATSDS) approach, which was developed by ReihanehKhorsand et al., was presented in 2017 [37]. When compared to other ways, the data demonstrated considerable improvements in process completion time, bandwidth, instances of missed deadlines, and the costs associated with the employment of virtual machines. Proportional Deadline Constrained (PDC) and Deadline Constrained Critical Path (DCCP) are two time-constrained algorithms that were introduced by Vahid Arabnejad et al. 2017 [38]. Both of these algorithms were developed by Vahid Arabnejad et al (DCCP). In order to make the process of identifying the priority of jobs and the filling in of resources more accurate, the algorithms have to be updated and improved. The results of the algorithms PDC and DCC show that they have a

higher rate of success while also achieving larger cost efficiencies when compared to other algorithms. This is according to the findings of the algorithms.
In the research paper authored by Toosi, A. N., et al. 2018 [39], a resource provisioning strategy was suggested to assist data-intensive apps in meeting their deadline requirements in hybrid cloud environments. The findings of the experiments revealed that the proposed algorithm was capable of allocating resources in a more effective manner when compared to other algorithms already in use.

III. TASK SCHEDULING

The process of cloud computing begins with the user submitting a request or task to a cloud computing service provider. This request or task may include a variety of information regarding the user's requirements, including a constraint, a priority, or other information. On the other hand, at a service provider that possesses a specialised task scheduling system, the scheduler receives requests from users to schedule their tasks in accordance with the terms of the service level agreement contract between the users and the cloud service provider. This is done in order to maintain the quality of service while simultaneously earning a profit from the services that are utilised by customers.
To carry out these tasks in accordance with certain constraints, the scheduler chooses the heterogeneous cloud resources from which it draws the heterogeneous cloud resources from which it draws the most appropriate resources to use, and the scheduler in this case is represented by the proposed model and serves as an intermediary between the users and the service providers (see Figure 1).

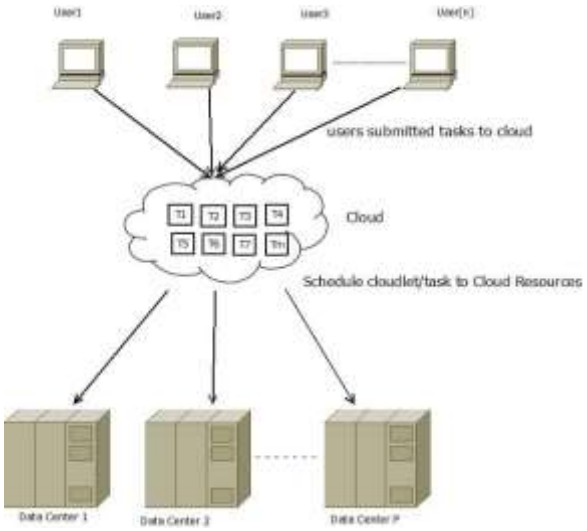

Figure 1: Scheduling in Cloud Computing

When it comes to task scheduling, the individual who will be carrying out the activities or tasks is the one who decides the appropriate order in which they should be completed. It is the process of assigning the appropriate resources to a certain work that is uploaded to the cloud in order to be completed. NP-hard problems are issues that cannot be solved in a finite amount of time and have an extremely large number of possible solutions. This problem fits into this category due to the large number of possible solutions. It is a methodology for the management of resources that are hosted in the cloud. The difficulty of selecting which resources should be given to which tasks in order to maximize resource consumption while lowering execution time can be overcome by scheduling jobs according to their priority. The scheduling algorithm needs to be effective in order for performance to be increased. It also needs to take into account a variety of elements, including load balancing across the entire system, handling interruptions, fault tolerance, and minimizing the overall amount of time needed for execution.
Once users have uploaded their tasks to the cloud in order for them to be performed, the tasks will need to be assigned to a processor in order for them to actually be finished. The question that needs to be answered now is how the duties are distributed throughout the processors in such a way that the owner of the cloud earns the maximum profit in the shortest period of time possible while also minimizing the amount of time it takes to carry out the tasks. As a result, task scheduling is able to address the challenge of allocating jobs to the processor that is most suited to perform those duties while also taking into account the other relevant parameters. The scheduling of tasks is one of the most effective techniques to improve economic efficiency and resource utilization. Several different approaches to the scheduling of work have been discussed and put to the test in a variety of settings.

We have two sorts of job scheduling algorithms to choose from while working in a cloud-based environment. The level of complexity distinguishes between distributed scheduling, in which tasks are assigned to different resources that are not geographically located in the same place, and centralized scheduling, in which all resources are located in the same place but the complexity level is lower than in distributed scheduling. Distributed scheduling involves tasks being assigned to different resources that are not geographically located in the same place. There are three different kinds of distributed scheduling

approaches that can be employed. These are heuristic methods, hybrid methods, and meta-heuristic methods. There are four categories of dynamic heuristic approaches: cost-based methods, energy-based methods, efficiency-based methods, and quality-of-service (QoS)-based methods. Static and hybrid methods are subdivided further into cost-based methods, energy-based methods, efficiency-based methods, and so on. The natural world serves as an important source of inspiration for the meta-heuristic research methodology. A few examples of static job scheduling methods are the min-min algorithm, the Symbiotic Organism Search (SOS) algorithm, the FUGE algorithm, the HEFT algorithm, and the CPOP algorithm.

IV. PROPOSED ALGORITHM

When we have access to accurate forecasts of the amount of time this method will take to complete a task, we will be able to improve the performance of the Greedy technique by incorporating it into the algorithm. The second method that we use is to priorities jobs according to their level of significance. In order to do this, we first assess how much time the virtual machine (VM) will have available both when the work is being done and when a new task is scheduled. In order to accomplish what needs to be done for this First Next Step, the estimation time calculation with IHEFT will be used. We made a comparison between the activity that was picked and other tasks that had the same finish time but a greater priority, and we decided to go with the task that had the higher priority. As part of our procedure, we adhere to the Min-Min principle and perform periodic verification and validation of the VM state to ensure that it is accessible for other tasks and is not being used for anything else.

Algorithm-

Enhanced -IHEFT New Algorithm

1. Create a DAG for all the submitted tasks Ti in Cloud.
2. Set the Computation Cost of tasks Ti and communication edges between the processor/resources R_j.
3. The Task ordering according to finish time that we calculate finish time and calculate average time
4. Sort the Task List with finish time
5. Repeat until task list
6. check the list for minimum the task finish time remove from list until size is zero
7. We Calculate the Virtual Machine Wait time for assigning the task
8. We calculate the Virtual Machine Makespam and compare with finish time from task list
9. Arrange tasks in a list in decreasing manner on the basis of their order of task OTi value.
10. for task in the list
11. map task to the processor which have the minimum execution time
12. end for
13. End

V. RESULTS AND ANALYSIS

I-HEFT SPEEDUP

	IHEFT SPEEDUP
10	7.35
20	10.59
50	15.8

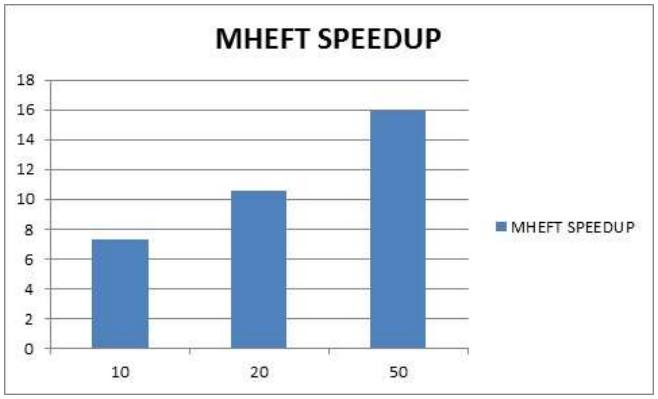

Figure 2: IHEFT Speedup

NEW SPEEDUP

	NEW IHEFT SpeedUp

10	7.12
20	10.53
50	15.64

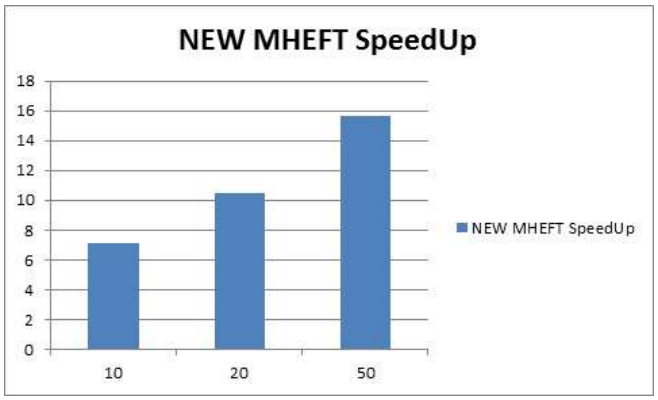

Figure 3: NEW IHEFT Speedup

A system's **speedup** can be thought of as a number that measures how quickly it can solve a problem when compared to another system. To put it another way, it is the increase in speed at which a task can be carried out when it is carried out on two architectures that are quite similar but have different resources.

The Modified HEFT Speedup is depicted in Figure 2. Figure 3 illustrates the newly improved speed of the Efficient Modified HEFT, and Figure 4 depicts the power consumption of the Efficient IHEFT.

IHEFT Power Consumption

	IHEFT Power Consumption
10	65313.36
20	38013.19
50	21022.72

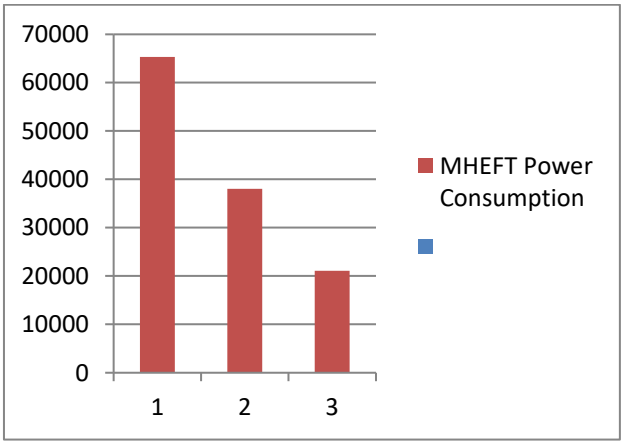

Figure 4: IHEFT Power Consumption

The method must be completed in a series of stages before it can be considered complete. These stages include compressing pages that share the same information in memory space and putting the space that was saved in a condition that uses less energy. The use of electricity by servers is strictly monitored and managed, and efforts have been made to cut down on the amount of electricity used by CPUs and memory. The amount of energy that servers need to run is cut down, and the efficiency with which memory is used is improved.

VI. CONCLUSION

In this study, we provide an improved version of the HEFT algorithm, which organizes the jobs into groups determined by their ratings and then assigns each group to the processor that is most suited to handle them. Following that, the workloads should be distributed throughout the various processors in the system in order to lower the overall power consumption of the system. The performance of the suggested technique was examined using the cloudsim simulator, and the estimated findings reveal that the suggested algorithm decreases the amount of time it takes for jobs to be finished and boosts the effectiveness of the system.

References

[1] Arabnejad, H., and Barbosa, J. G. (2014). List Scheduling Algorithm for Heterogeneous Systems by an Optimistic Cost Table. IEEE Transactions on Parallel and Distributed Systems, 25(3), pages 1-14.

[2] Bailin, P., Yanping, W., Hanxi, L., and Jie Q. (2014). Task Scheduling and Resource Allocation of Cloud Computing Based on QoS. Advanced Materials Research, Vols. 915-916, pages 1382-1385.

[3] S. Raghavan, P. Sarwesh, C. Marimuthu , K. Chandrasekaran, "Bat algorithm for scheduling workflow applications in cloud," International Conference on Electronic Design, Computer Networks & Automated Verification (EDCAV), Shillong, 2015, pp 139-144.

[4] Byun, E. K., Kee, Y. S., Kim, J. S., and Maeng, S. (2011). Cost Optimized Provisioning of Elastic Resources for Application Workflows. Future Generation Computer Systems, 27(8), pages 1011–1026.

[5] G. Wang, H. C. Yu, "Task Scheduling Algorithm Based on Improved Min-Min Algorithm in Cloud Computing Environment", Applied Mechanics and Materials, Vols. 303-306, 2014, pp 2429-2432.

[6] G. Wang, H. C. Yu, "Task Scheduling Algorithm Based on Improved Min-Min Algorithm in Cloud Computing Environment", Applied Mechanics and Materials, Vols. 303-306, 2014, pp 2429-2432.

[7] Abdulhamid, S.M., Latiff, M.S.A., Idris, "Tasks scheduling technique using league champion ship algorithm for makes pan minimization in IaaS cloud.", ARPN J. Eng. Appl. Sci., Volume 9, 2015, pp 2528–2533.

[8] Bansal, N., Awasthi, A. &Bansal, "Task Scheduling Algorithms with Multiple Factor in Cloud Computing Environment", Information Systems Design and Intelligent Applications, Springer.

[9] Babu, K.R., Samuel, P., "Enhanced bee colony algorithm for efficient load balancing and scheduling in cloud", Innovations in Bio-Inspired Computing and Applications , 2016, pp 67-78.

[10] Pop, Florin, et al. "Deadline scheduling for aperiodic tasks in inter-Cloud environments: a new approach to resource management." The Journal of Supercomputing 71.5 (2015): 1754-1765.

[11] Pandey, S., Wu, L., Guru, S. M., and Buyya, R. (2010). A Particle Swarm Optimization-based Heuristic for Scheduling Workflow Applications in Cloud Computing Environments. In 24th IEEE International Conference on Advanced Information Networking and Applications (AINA). IEEE.

[12] Parsa, S., and Entezari-Maleki, R. (2009). RASA: A New Task Scheduling Algorithm in Grid Environment. World Applied Sciences Journal, 7 (Special Issue of Computer & IT), pages 152-160.

[13] Rodriguez, M. A., and Buyya, R. (2014). Deadline Based Resource Provisioning and Scheduling Algorithm for Scientific Workflows on Clouds. IEEE Transactions on Cloud Computing, 2(2), pages 222-235

[14] Chun-Yan Liu, Cheng-Ming Zou, and Pei Wu. "A Task Scheduling Algorithm Based on Genetic Algorithm and Ant Colony Optimization in Cloud Computing"., Distributed Computing and Applications to Business, Engineering and Science (DCABES '14). IEEE Computer Society, Washington, DC, USA, 2014, pp 68-72.

[15] Topcuoglu, H., Hariri, S., Wu, W. Min-You. (2002). Performance-Effective and Low- complexity Task Scheduling for Heterogeneous Computing. IEEE Transactions on Parallel and Distributed Systems, 13(3), pages 260-274.

[16] H. Yang, "Improved Ant Colony Algorithm Based on PSO and its Application on Cloud Computing Resource Scheduling", Advanced Materials Research, Vols. 989-994, pp. 2192-2195, 2014.