iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** V **Month of publication:** May 2024

DOI: https://doi.org/10.22214/ijraset.2024.62421

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue V May 2024- Available at www.ijraset.com

Task Scheduling by Hybrid Metaheuristic Approach in Cloud Environment

Meena Jindal[1], Khushwant Kaur[2*]
Sri Guru Gobind Singh College Chandigarh, Panjab University
**Corresponding Author*

Abstract: ***Recent advancements in hardware, virtualization technology, distributed computing, and service delivery over the Internet have led to the formation of cloud computing. In a distributed computing platform, the scheduling of task workflow is a well-known NP-hard problem. This becomes more complex and challenging if the virtualized clusters execute a large number of tasks in a cloud computing platform. From low-level execution tasks in multiple processors to high-level execution tasks, many heuristics have been proposed. This chapter deals with the proposed min-min algorithm, max-min algorithm, PSO, and TS. The max-min and min-min algorithms estimate the execution time and overall completion time of all the tasks Then assign the tasks on better suitable resource. PSO is simple and effective in a wide range of applications with low computational cost and fast speed. TS approach is one of the popular local search techniques based on neighborhood search algorithm. TS can be directly applied to any kind of virtual optimization problem.***
Keywords: NP-hard problem, Genetic Algorithms (GA), Particle Swarm Optimization (PSO)

I. INTRODUCTION

Cloud computing is an innovative invention using advanced computational power with enhanced storage capacities through sharing over the Internet. It is a large group of computers that are interconnected with proper storing of data and running applications which are properly provisioned by a service provider with on-demand access over the internet. The advantage of cloud computing is of affordable price and the threat is its security. Through cryptographic techniques, security cannot be ensured (Yawale & Gadicha 2014).

The main aim of cloud computing is sharing of data; the users can save their information in a cloud where it can be shared but there is a possibility of breach of trust due to leakage of information. Cloud securing is an approach to protect data in the cloud computing environment. The following parameters influence cloud security: Networks, Databases, Virtualization, Resource scheduling, Operating system, Concurrency control Transaction management, Load balancing, and Memory management (Senthil et al. 2012).

The biggest issue in cloud computing is security, where an attacker imitates as a legitimate user and pilfer the information which leads to many problems as many users share data in the cloud. The various issues on safety faced in cloud computing are privacy, data, and security issues, and infected application (Kumar & Nithya 2014).

1) Data issues: Generally, sensitive data is stored in the cloud and data can be accessed at anytime from anywhere in the cloud. Two issues that are faced by data security are data integrity and data loss.
- Data integrity – There is a loss of data integrity in the cloud as many users can access and modify the data simultaneously.
- Data loss – Unauthorized users in one way or the other can view the data which are not supposed to lead to breaching of trust.

2) Privacy issues: Again, there is a breach of privacy as anyone can access data in the cloud and confidential data cannot be protected.

3) Infected application: It is the responsibility of the provider to monitor the server against malicious attack by infected application and this influences the cloud user.

4) Security Issues: Generally in two levels, security is breached in cloud computing – at the provider level and at the customer level.
- Security of provider level: It should be checked that the server is secure from threats which come from outside. If proper security is offered by the provider to the customer, then the cloud is good.
- Security on user level: It is the duty of the user to check that the data that is received by the provider is without any loss.

Emerging and next generation architecture for computing is Cloud Computing and this is a combination of computing resources which are accessible through the internet. Conventionally, data are stored by the client in data centers with firewall and other

©IJRASET: All Rights are Reserved

security techniques can protect data against intruders accessing the data.

The control of the users is more when the data is stored in a limited space at the boundaries of an organization and procedures that are properly defined could be used for accessing the data. But in Cloud Computing, as the data is stored anywhere across the globe, there is less control for the client over the stored data.

To lie emphasize on trust for the growth of cloud computing, the providers should protect the data from unauthorized users from access and disclosure. One of the techniques is that the data should be encrypted before storing it, but from the side of the user, it is a heavy burden in terms of key management and maintenance perspective etc., Another technique is that the provider should take the responsibility of encrypting and decrypting of data that is stored but again there is a risk of security in this as the provider has access to both storage and security service (Hulawale, 2013).

A trusted third party provider can be involved in client service where one provider can be in-charge of security services while the other could be a data storage provider. Here, the trusted third party provider should not store any data from his/her end and should confine himself/herself from providing only security services. Hashing algorithms like SHA-1, provide encryption/decryption using a symmetric algorithm like Advanced Encryption Standard (AES), will help in providing data integrity verification and restrict and define the strata of users who have accessibility to the shared data securely by defining access list (Bhagat & Sahu, 2013).

The software does not store any data in the trusted third party security system server and it just encrypts/decrypts/computes/verifies the hash of the data. The encrypted data is stored along with the original data hash in a separate cloud which is a security cloud and so it is difficult to access user data even though the storage system administrator has access. When a data is downloaded from the storage cloud, it is first decrypted and a new hash is calculated which in turn is compared with the hash of original data that is stored in security cloud. Ultimately, the application/software provides the user with the encrypted data in storage cloud and hash and encryption/decryption keys in security cloud service and a single service provider cannot get in touch with both. Another advantage of designating authority to a trustworthy 3rd party is that the client is relieved from any kind of key management or overhead maintenance of any key message related to data on its device as the client is allowed to use any browser-enabled devices to access such services.

To maintain reputation, the cloud provider might hide data corruptions. So as to avoid this issue, an effective Third Party Auditor (TPA) is introduced to audit user‟s outsourced data when necessary. TPA can act as an inspector; there are two types of audit ability – public and private. Higher scheme of efficiency is achieved through private audit ability but public audit ability allows anyone, not just the data owner or the client in order to challenge the cloud service for the correctness of data storage while keeping private information. To keep the data owner free from the burden of management, TP will audit the data for the client. It checks whether the data stored by the client is intact which can achieve economies of scale for cloud computing. Thus, the data owner is helped in making sure that his data is safe in the cloud and data management is easy and there are fewer burdens on data owner (Yawale & Gadicha, 2014).

II. OPTIMIZATION IN CLOUD COMPUTING

Under given circumstances, the process of finding the best solution is optimization. In general, optimization can maximize or minimize the value of the function and it can be local or global optimum. Based on their utilization, optimization can be linear, non-linear, dynamic etc and there are different techniques to solve problems (finding best solution). Task scheduling and resource management are the key technologies of the cloud environment. The manner in which QoS is maintained is considered through previously designed task scheduling algorithms but optimization of resources in gaining maximum profit is considered. As class optimization problems are NP hard, many optimization techniques are proposed as follows:

Genetic Algorithms (GA) are basically inspired from theory given by Darwin about evolution. The solution to the problem is given by genetic algorithm starting with a solution set (i.e. chromosomes) called the population. Solutions from a particular population are taken into account and used to generate a new population. When this is enthused, the new population will be improved from the old one. Solutions are selected to form new ones (offspring) which are selected according to their fitness (Bilgaiyan et al. 2015).

GA follows meta-heuristic approach. This heuristic is regularly used to produce useful solution to optimization and search problems. GAs belongs to the class of evolutionary algorithms, which generates solutions to optimize problems using technique inspired by natural evolution, such as inheritance, and mutation. GA is useful in solving a variety of optimization problems. It is basically applied to dynamic scheduling problems with multiple tasks. These Tasks must be non-identical, and numerous autonomous task. They must be distributed in shared multiprocessor memory system.Structural GA provides the option of solving the solution structure parameter problems with accuracy.

Particle Swarm Optimization (PSO) algorithm is a population- based random optimization method which is proposed by Russell Eberhart and James Kennedy in 1999. The development of this algorithm is inspired from the swarm behaviour of birds or fish. This system begins with a population which has random solutions and it updates the generation to find an optimal solution.

In contrast with GA, none of the evolutionary operators such as crossover and mutation are available in the PSO algorithm. Solutions in PSO algorithm are referred as particles which move in the problem search space and follow the current optimal particle (Pandey et al. 2010).

In this algorithm, each particle follows the particle which has a better fitness function among all the other particles. However, it does not forget its own experience. Hence, it follows the condition and state in which it has the best fitness function. Thus, in each iteration of the algorithm, each particle determines its next position based on two values: first, the best position that the particle has ever had indicated by pbest and also the best position that all the particles have ever had indicated by gbest. In other words, gbest refers to the best pbest in the entire population. Conceptually, pbest for each particle refers to the memory which a particle has experienced about its best position and gbest represents the public knowledge of the population when particles change their positions based on gbest, they try to keep up with the knowledge of the population. Conceptually, the best particle connects all the particles of the population with each other (Kamalinia & Ghaffari 2016).

The Cuckoo Search (CS) is an optimization algorithm developed by Xin-she Yang and Suash Deb in 2009. It is inspired by the obligate brood parasitism of some cuckoo species by laying their eggs in the nests of other host birds (of other species). Some host birds can engage direct conflict with the intruding cuckoos. For example, if a host bird discovers that the eggs are not their own, it will either throw these alien eggs away or simply abandon its nest and build a new nest elsewhere. Cuckoo characteristics could be described as a model for good behavior other than animals and this has extensive use in computing Intelligence Systems (Al-maamari & Omara 2015).

According to the CS algorithm, an initial set of nests, which represent the solutions are randomly generated. These solutions are then updated over multiple generations. The process of updating an individual solution is as follows; a random nest is chosen, and a new solution is generated by random walking from this previous solution. This new solution can then replace a different randomly chosen solution if it has a fitness value better than the original. After the possible replacement of a solution, all the nests are ranked by fitness and the worst fraction of the nests is replaced with random solutions. This combination of mechanism allows the solutions to search locally and globally at the same time for the optimal solution.

1) *Bacterial Foraging Algorithm:* In the process such as chemo taxis swarming, reproduction, elimination, and dispersal are present in bacterial foraging. The objective of searching food is depicted as the energy intake per time (E/T) is maximized (Jacob et al. 2014).

- *Chemo taxis:* The foraging behaviour of E. coli bacteria through swimming and tumbling with the help of flagella are the basis for this technique.
- *Swarming:* The group of bacteria moves towards nutrients but it can be either attracted or rejected. During stress, one bacteria will attract other bacteria.
- *Reproduction:* This is the next step to chemo taxis. The fitness value of the bacterium is calculated in this step and is sorted in an ascending order. Maintenance of the size of the swarm is constant as less healthy bacteria die and healthier ones split into two sexually.Elimination and dispersal: This process takes place to avoid local optima.

2) *Ant Colony Optimization (ACO):* For single objective combinatorial issues ACO is applied successfully. ACO can manage Multi objective Task scheduling problem including defining multi objective pheromone information, calculating probability matrix and strategies on pheromone update. Ants, when considered individually, are unsophisticated in nature. A restricted memory and individual behavior is exhibited which seems to have quite a large random component (Vijayalakshmi & Muthusamy 2014). But when they are collective, they perform a variety of tasks that are complicated with consistency and reliability. Though this is more of a self-organization, rather than learning, ants have to cope with a phenomenon which looks more like overtraining in reinforcement learning techniques.

   Different web services are implied through food in the context of cloud computing. The path with more pheromone trails is generally chosen by the ants as more used path will have more pheromone trails and will lead to a better food source. Also, the shorter path will have more pheromone trails than longer paths as the ants that chose the shorter path will reach food early and get back to the nest by the same path because of unelaborated pheromone (Ranjan Kumar et al. 2013). So, the shorter path will have more amount of pheromone and all ants choose the shorter path.

3) *Artificial Bee Colony (ABC):* Foraging behavior of honey bee colonies is the basis for swarm-based metaheuristic ABC algorithm. Three important elements are present in this model – employed, unemployed foragers, and food sources. Food source is abundant and close to their hive (Bolaji et al. 2013). The model describes two leading modes of behavior. These are essential for self-organization and collective intelligence. Positive feedback is obtained by recruitment of forager bees to rich sources and at the same time, poor sources are abandoned causing negative feedback.

The flashing characteristics of fireflies so as to develop firefly inspired algorithms. For simplicity in describing the new Firefly Algorithm (FA), it now use the following three idealized rules: (1) all fireflies are unisex so that one firefly will be attracted to other fireflies regardless of their sex; (2) Attractiveness is proportional to their brightness, thus for any two flashing fireflies, the less brighter one will move towards the brighter one. The attractiveness is proportional to the brightness and they both decrease as their distance increases. If there is no brighter one than a particular firefly, it will move randomly; (3) the brightness of a firefly is affected or determined by the landscape of the objective function. For a maximization problem, the brightness can simply be proportional to the value of the objective function. Other forms of brightness can be defined in a similar way to the fitness function in GAs (Florence & Shanthi 2014).

Simulated Annealing (SA) is an initial meta-heuristic algorithm inventing from an analogy of how an optimal atom arrangement is originated in the statistical mechanics. It is a general method to find a solution for some NP hard optimization problems in a heuristic process. The SA‟s effectiveness comes out from its extension of two primary heuristic techniques, iterative improvement scheme and divides and conquers approach. It processes the temperature as an explicit strategy to direct the search. The solution space is frequently discovered by taking random tries. The SA method randomly produces a large number of potential results, maintaining together good and bad solutions (Madni et al. 2016).

League Championship Algorithm (LCA): It is inspired by the contests of sport teams in a sports association (league). A league schedule is designed every week (iteration) for the teams (individuals) to play in pairs and the result is in the form of win or loss depending upon the playing strength (fitness value) of a team following a meticulous team formation/playing technique (solution). On the basis of prior week knowledge, the team makes changes in the formation (a new solution) for the next week competition and the championship continues till the specified number of seasons (terminating condition) (Kalra & Singh 2015).

BAT algorithm: Getting inspiration from echolocation behavior of bats, Yang introduced BAT algorithm, a novel optimization algorithm in 2010. Bats use echolocation to estimate the distance of their prey. They fly randomly with a velocity, position, frequency, loudness and pulse emission rate to seek for their prey. When they are hunting for their prey, they can adjust their frequency, loudness and pulse rate of emission based on the distance amid them and the prey. This behavior of bats has been used to formulate BAT algorithm.

III. METHODOLOGY

This section deals with the max-min algorithm, min-min algorithm, Particle Swarm Optimization (PSO) algorithm, and Tabu Search (TS).

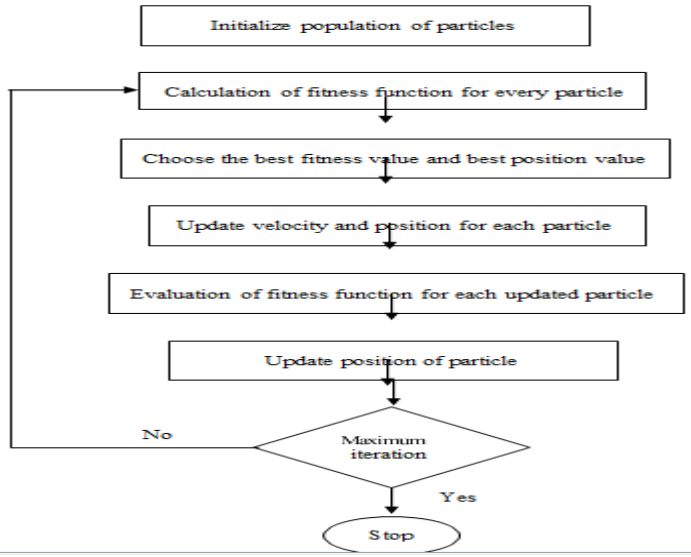


Figure 3.1 Flowchart for Particle Swarm Optimization (PSO) Algorithm

TABU Search

An iterative process that has meta-strategy for constructing extended surroundings is Tabu search. It lays a specific emphasis on eliminating the chances of being stuck in a local optimum. It has the ability to tackle any sequencing problem using the described structure. At present, it is the most preferred method for local searches. It gives a near optimal solution of many combinatorial optimization problems. Briefly, the basic idea of this method consists in starting from an initial basic permutation of n jobs and searching through its neighborhood, and a set of permutations related to the basic one, for a permutation with the lowest make span. The search is iterated taking into account the best permutation as the new basic permutation and there is a continuation of the process (Glover 1989, Glover 1990).

The moves generate the surroundings of the permutation. In the basic permutation, a move transforms some jobs‟ location. The procedure of Tabu list has been incorporated for avoiding cycling, and continuing the search in a restricted area and avoiding the local optimum. The elements of the tabu list define the move that cannot be applied currently which means that in the present neighbourhood that is analyzed, they deduce the permutations that are not allowed. Every time when a new basic permutation is found, the content of the list is refreshed. Eliminating the oldest element, the new one is added. When a predetermined number of iterations have been reached by no improvement in the present most superior makespan, or the time has expired or the given number of iterations has been performed by the algorithm, the search procedure ceases. It also stops when the surroundings are empty or the permutation with the makespan that is satisfactory has been discovered. It is an art to design a certain component of the search algorithm. The algorithm is influenced by the component construction, running time, convergence speed etc.

Instead of calculating the makespan explicitly for choosing the best solution, in algorithm TS(Grabowski & Mieczyslaw 2004), in order to reduce calculations for the search of the neighbourhood, lower bound is used on the makespan. Also, there is an application of dynamic length that also aids us to be stuck at the local optimum. Finally, by incorporating precise disturbances, the improved algorithm can avoid the local optimum. These disturbances comprise of many moves at the same time in one iteration and lead the search to a better area of the solution space where “good solutions” are found. An algorithm that solves large-size 'an own-shop case with high precision in minimal time is suggested. A basic TS algorithm as shown below (Pirim et al. 2008):

IV. RESULTS AND DISCUSSION

Simulations are carried out using 15 VM and a variable number of jobs with the computing power of 300-1500 jobs. The max-min, min-min, Particle Swarm Optimization (PSO) and Tabu Search (TS) are evaluated. The average schedule length and the ratio of successful execution are shown in Tables 3.1

Table 1 Average Schedule Length (in seconds)

NUMBER OF TASKS	MAX-MIN	MIN-MIN	PSO	TS
300	372	358	315	333
600	1240	1150	981	1035
900	1958	1863	1623	1710
1200	2810	2655	2253	2384
1500	3426	3293	2766	2909

From the Table 1, it is observed that the TS possess lower average schedule length when compared with max-min, min-min and PSO which varies from 11.06 to 18.02%, from 7.23 to 12.38% and from 5.03 to 5.65% for 300, 600, 900, 1200 and 1500 number of tasks respectively. The lower average schedule length is desired to schedule the task workflow when virtualized clusters execute a large number of tasks in a cloud computing platform. Here the PSO scheduling algorithm has lower average schedule length for various number of tasks compared to the other methods.

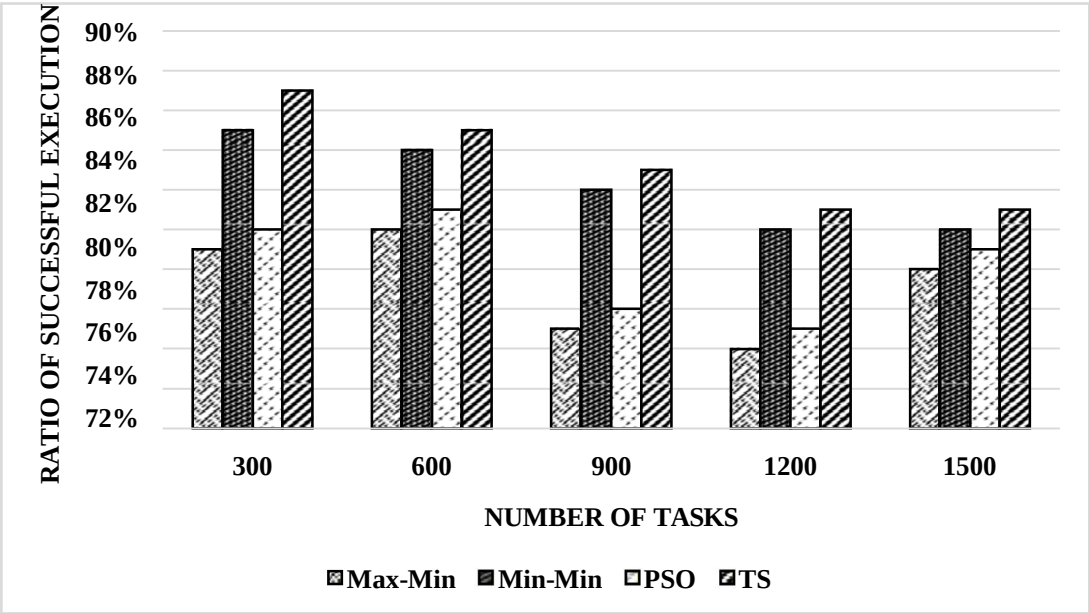


Figure 1 Ratio of Successful Execution

From the Figure 1 it is observed that the TS has higher ratio of successful execution than max-min, min-min and PSO which varies from 3.77 to 10.12%, from 1.18 to 2.32% and from 2.5 to 8.8% for 300, 600, 900, 1200 and 1500 number of tasks respectively. The higher ratio of successful execution is desired to schedule the task workflow when virtualized clusters execute a large number of tasks in a cloud computing platform. Here the TS scheduling algorithm has a higher ratio of successful execution for a various number of tasks compared to the other methods.

V. CONCLUSION

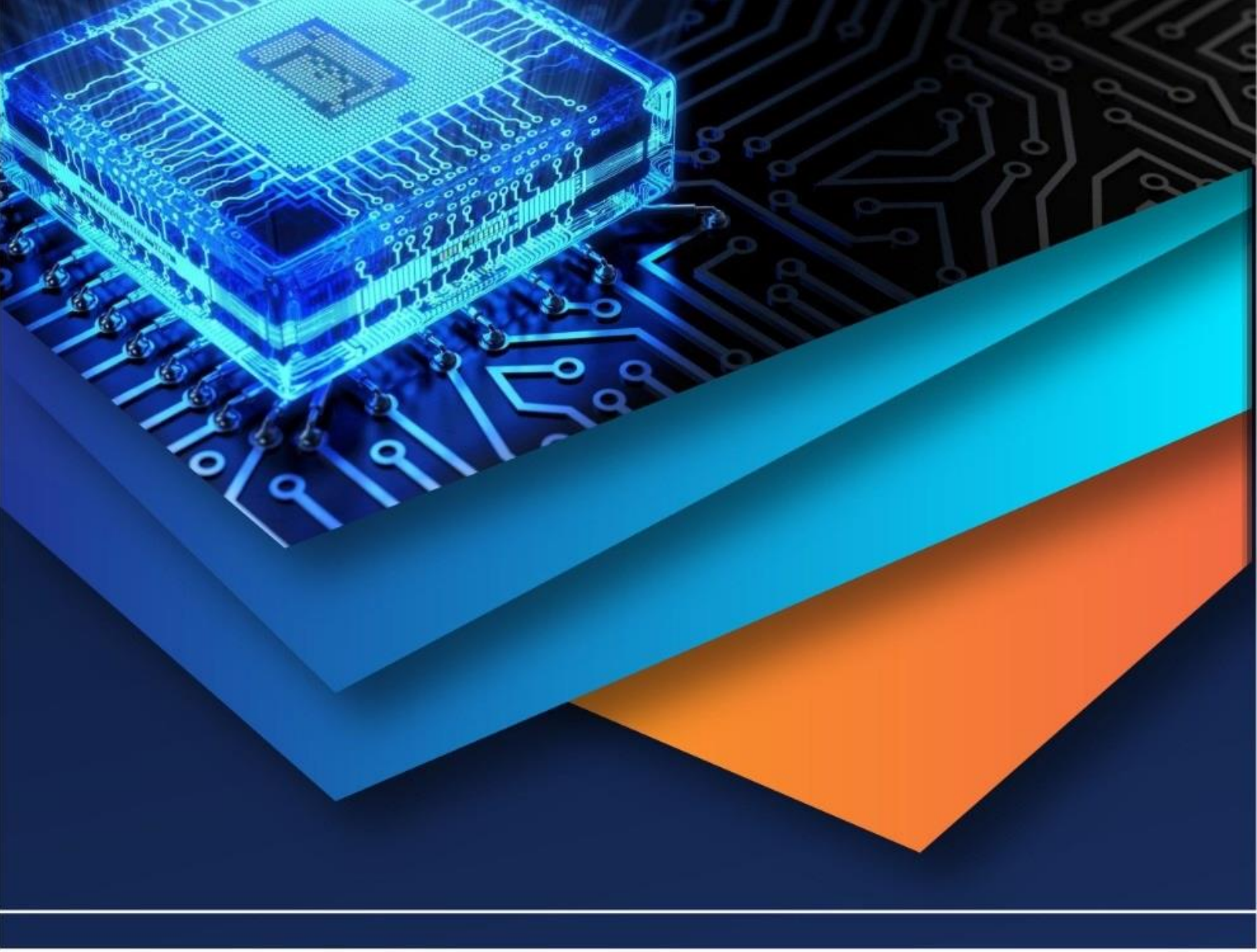
The simulations that were carried out using 15 VM and a variable number of jobs with the computing power of 300-1500 jobs shows that there is a gradual increase in the average schedule length for all the methods with the increase in the number of tasks and the ratio of successful execution of the tasks gradually decreases with increase in the number of tasks but as the task size increase beyond 1500 there is a significant increase in the ratio of successful execution. This behavior shows that the performance of the system is not stable. Hence in order to improve the scalability of the tasks and the performance of the scheduling system meta-heuristic approach can be considered.

REFERENCES

[1] Abdi, S, Motamedi, SA & Sharifian, S 2014, "Task scheduling using Modified PSO Algorithm in cloud computing environment". In International conference on machine learning, electrical and mechanical engineering, pp. 8-9.
[2] Aete, C 2012, "Areas of shared responsibility for public cloud security", hp Cloud Source Blog, vol. 7.
[3] Agarwal, D & Jain, S 2014, "Efficient optimal algorithm of task scheduling in cloud computing environment". arXiv preprint arXiv:1404.2076.
[4] Alkhanak, E, Lee, SP & Ling, TC 2015, "A Hyper-Heuristic Approach using a Prioritized Selection Strategy for Workflow Scheduling in Cloud Computing". In The 4th International Conference on Computer Science and Computational Mathematics (ICCSCM 2015), pp. 601-608.
[5] Alkhanak, EN, Lee, SP & Ling, TC 2015, "A Hyper-Heuristic Approach using a Prioritized Selection Strategy for Workflow Scheduling in Cloud Computing".The 4th International Conference on Computer Science and Computational Mathematics (ICCSCM 2015)
[6] Alliance, C 2011, "Security guidance for critical areas of focus in cloud computing v3. 0". Cloud Security Alliance.
[7] Al-maamari, A & Omara, FA 2015, "Task scheduling using hybrid algorithm in cloud computing environments". Journal of Computer Engineering (IOSR-JCE), vol. 17, no. 3, pp. 96-106.
[8] Al-maamari, A & Omara, FA 2015, "Task scheduling using PSO algorithm in cloud computing environments". International Journal of Grid and Distributed Computing, vol. 8, no. 5, pp. 245-256.
[9] Aswin, MR & Kavitha, M 2012, "Cloud intelligent track-Risk analysis and privacy data management in the cloud computing". In RecentTrends In Information Technology (ICRTIT), 2012 International Conference on IEEE pp. 222-227.
[10] Bateman, K 2010, Community cloud computing benefits and drawbacks.
[11] Bhadauria, R & Sanyal, S 2012, "Survey on security issues in cloud computing and associated mitigation techniques". ArXiv preprint arXiv:1204.0764.
[12] Bhadauria, R, Chaki, R, Chaki, N & Sanyal, S 2014, "Security Issues in Cloud Computing". Acta Technica Corviniensis-Bulletin of Engineering, vol. 7, no. 4, pp. 159.
[13] Bhagat, A & Sahu, RK 2013, "Cloud Data Security While using Third Party Auditor". International Journal of Computer Applications, vol. 70, no. 16.

[14] Bilgaiyan, S, Sagnika, S, Mishra, S & Das, M 2015, "Study of task scheduling in cloud computing environment using soft computing algorithms". International Journal of Modern Education and Computer Science, vol. 7, no. 3, pp. 32.
[15] Chauhan, T, Chaudhary, S, Kumar, V & Bhise, M 2011, "Service level agreement parameter matching in cloud computing". In Information and Communication Technologies (WICT), 2011 World Congress on IEEE, pp. 564-570.
[16] Chawla, Y & Bhonsle, M 2012, "A Study on Scheduling Methods in Cloud Computing". International Journal of Emerging Trends & Technology in Computer Science (IJETTCS), vol. 1, no. 3, pp. 12-17.
[17] Crago, S, Dunn, K, Eads, P, Hochstein, L, Kang, DI, Kang, M & Walters, JP 2011, "Heterogeneous cloud computing. In Cluster Computing (CLUSTER)", 2011 IEEE International Conference on IEEE, pp. 378-385.
[18] Delavar, AG & Aryan, Y 2014, "HSGA: a hybrid heuristic algorithm for workflow scheduling in cloud systems". Cluster computing, vol. 17, no. 1, pp. 129-137.
[19] Fernandes, DA, Soares, LF, Gomes, JV, Freire, MM & Inácio, PR 2014. ,"Security issues in cloud environments: a survey". International Journal of Information Security, vol. 13, no. 2, pp. 113-170.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)