Enhancing Scheduling Multi - Resource Jobs in the cloud by Using Hungarian Model

Jagruti Soni
Computer Science
JSPM's Bhivarabai Sawant Institute of Technology & Research, Pune, India

Professor P. B. Jawalkar
Computer Science
JSPM's Bhivarabai Sawant Institute of Technology & Research, Pune, India

***Abstract*— As we know cloud is the most trusted and huge structure as well as it is the most complex structure too. This makes the cloud to play an important role in all of its three functional joints like Iaas,Pass and Saas. So a perfect cloud is designed to perform all the functions of three categories stated before. The functions or tasks in cloud are generally referred as the jobs, and scheduling these tasks is the most cumbersome task in the cloud based on the available resources. Each and every job in the cloud is related to one or more resources, So assigning these multi resources jobs need top precision to avoid the long job queues within the stipulated time is very important to maintain the cloud's health. Some systems are existed to do this by using the linear job priority queues and other are using the randomized algorithms for scheduling the multi resource jobs in the cloud. These methodologies always have inconsistency in scheduling as they are not able to consider all the possibilities based on the available resources. To enhance the working strategy of cloud in Scheduling Multi-Resource Jobs, proposed model uses Hungarian model and Decision tree techniques. Here all the possibilities of the scheduling are considered to yield the best scheduling decision which eventually enhances the performance of the cloud.**

Keywords - Job Scheduling, Resource Allocation, Decision tree, Hungarian Task allocation.

• I INTRODUCTION

Due to the internet boom, there has been increased utilization and manufacture of various devices capable of using the internet through various different channels. This has led to the emergence of various industries that are based on the versatility of the internet platform. One of these technologies is the cloud platform. The cloud platform is highly dependent on the internet infrastructure, which is not the same as requiring the internet to be of any significance, this is a very common misconception shared with the majority of the people.

Therefore, the cloud platform is dependent on the internet infrastructure which does not mean it would not function without an internet connection. This misconception may have something do to with the coining of the term cloud. as it was not officially declared as a cloud anywhere. The name was adopted due to completely different reasons. Usually, in the early stages of the conception of the internet, most of the researchers and authors depicted the internet as an ominous cloud in various drawings and other diagrams.

This representation of the internet as a cloud therefore, just got stuck and hence we have one of the largest technologies based off the internet to be named as the cloud. the cloud platform has seen a meteoric rise, with the majority of the users and large organizations clamoring to implement their functions and offloading their files on to the cloud to make them accessible ubiquitously anywhere in the world.

The exponential increase in the number of cloud providers and the willing customers ready to adopt the new technology has invariably reduced the prices of this particular technology. This has, in turn, led to a lot more adoption by the masses as the lower prices are quite affordable. The main take away from this technology is that it would cost a lot more to maintain the same amount of storage and also making it ubiquitously available everywhere, therefore, the lower prices of the cloud platform makes a lot more sense.

The cloud platform is in charge of managing the data and making it available to the user anywhere in the world, this is a very difficult task, where the cloud service provider segments the data into small segments and distributes them to various servers across the globe. These servers manage the data, make backups in case of a failure and also present the data to the customer as and when requested. This is a highly complex technique that maintains the data throughout the various servers and is charged with accurate search and retrieval when requested.

The cloud platform is not only utilized for storage purposes, even though the storage accounts for the majority of the usage of the cloud. But building on the cloud platform, the cloud providers and researchers have developed innovative services that can be availed from the cloud platform. the services are broadly categorized into three different categories, namely, SaaS or Software as a Service, PaaS or Platform as a Service and IaaS which stands for Infrastructure as a Service.

These services enable a high level of computing through the introduction of the cloud. The SaaS platform provides various applications and software which can be used by the customer while being installed on the cloud, this eliminates the process of installing the software and various hardware requirements, etc. this completely bypasses a lot of hurdles and limitations faced by a customer. The PaaS allows the development and deployment of various Web-based applications, this eliminates the need to install and maintain various support structures for the web apps. The IaaS acts as a large resource pool, capable of providing various elements such as processing power, storage, etc for a development environment.

Managing and handling various tasks such as these on the cloud platform is one of the most important and essential concepts. This is due to the fact that most of the organizations and individuals depend on the services provided by the cloud ever increasingly. If the tasks are not managed properly, it could lead to a lot of losses as the dependency on the cloud has reached extreme lengths.

Task scheduling is one of the most essential issues surrounding the cloud platform. Task scheduling is responsible for the smooth functioning of the multitudes of companies. As the main objective of the task scheduler is to manage and schedule tasks that can better segregate and utilize the resources on the cloud. The task scheduler has the great responsibility of allotting various resources to various tasks taking place simultaneously at the cloud. When a task is encountered, the task scheduler is responsible for allotting the processing power and storage for a set amount of time to that particular task.

Task scheduling has to take into account all the various tasks being executed at that instance, the tasks that are waiting and also anticipate the various tasks that could come in the future. Optimum task scheduling attains much better performance, efficiency and quality of service for the customer. This would, in turn, reduce the costs that are incurred for providing the service to the customers. to achieve the optimum task scheduling, there have to be various parameters that need to be monitored, such as completion time and task completion cost.

Therefore, there is an urgent need to develop efficient task scheduling to reach an optimum state that reduces energy consumption and can better save the resources at the cloud center. Task scheduling has been defined as an NP-hard problem; therefore, it is quite complicated and difficult to achieve the optimum levels of efficiency. But it is imperative that the task scheduling algorithms need to increase the efficiency by a large margin.

The Hungarian algorithm is one of the oldest techniques developed in statistical mathematics for the purpose of solving various combinatorial assignment problems in polynomial time. The algorithm has been in use in the early 1950s after being introduced as a Hungarian Algorithm, which was named as such due to the algorithm being based on the research done by two Hungarian mathematicians.

The Hungarian algorithm is a novel technique that can solve the assignment problem in polynomial time, this is highly useful as it can solve highly complex weighted assignment problems. What that means is that the Hungarian algorithm is capable of finding the way to complete a task with a minimum amount of cost incurred. Therefore, the Hungarian Algorithm is capable of finding out which task could be done by which person having multiple quotes in the least cost.

The Hungarian Algorithm is a very simple algorithm that is based on the concept of Matrices. It is highly accurate and can achieve an extremely low computational complexity. Therefore, it is a great addition to the Task allocation procedure in the cloud platform, an almost excellent choice to achieve high levels of optimization and efficiency for cloud's resources.

This research paper dedicates section 2 for analysis of past work as literature survey, section 3 deeply elaborates the proposed technique and whereas section 4 evaluates the performance of the system and finally section 5 concludes the paper with traces of future enhancement.

• LITERATURE REVIEW

F. Li states that task scheduling is one of the most complex tasks in the field of cloud computing. But since the cloud is a pervasive and multi-functional paradigm, it is highly essential that the tasks are managed efficiently. Due to this,there have been copious amounts of research being done in this sector but most of it is concentrated around the

scheduling algorithms and not on the task decomposition. Therefore, the authors in this paper propose a technique for the modeling scheduling process that takes into consideration the various task decomposition process in a cloud. The system has not been implemented in real time therefore that is one drawback of this paper. [1]

H. Wei [2] explains that there is still a lot of research left to utilize the cloud platform to its potential. The main objective in the cloud platform is the resiliency of the cloud to reduce the number of errors and single point failures. The authors propose an innovative hybrid cloud management system that provides a robust and resilient cloud platform. The proposed technique utilizes a custom scheduling process that balances the load and minimizes the amount of user request makespan. The one drawback in this technique is that it has not been tested on a large enough cloud provider yet.

M. Sarvabhatlaexpresses that there has been an increase in networking and mobile technologies due to modernization. This has led to the development of various Service Level Agreements (S.L.A) and Quality of Service (QoS) tasks exponentially. Due to this, there has been a huge spike in the demand for energy efficient task scheduling techniques. therefore, the authors have presented an innovative technique that can scrutinize the energy efficient task scheduling [3]. The main drawback in this methodology is that the system has only been tested in simulations, therefore, lacks real-world statistics.

A.Panneerselvamintroduces the immense growth in the field of computers and electronics and their utilization for processing Big Data. Cloud is one of the most popular sources of Big Data processing as they are distributed frameworks[4]. As these two technologies are invariably connected due to the inner workings of the cloud, the authors developed a hyper-heuristic MapReduce workflow scheduling algorithm, which can efficiently organize workflows by utilizing the MapReduce functions. This technique has been tested extensively and generated improved results at the cost of increased computational complexity.

B. Ravandi elaborates on the Software Defined Storage (SDS) as it is the element that separates the data and control from each other giving way for the data management to be done through automation [5]. As most of the cloud block storage does not have an SDS framework, this cannot be achieved. Therefore, to solve this issue, the authors have developed a framework that can automatically optimize the scheduling and the workload on block storage in the cloud. the authors have planned to implement dynamic migration and feedback into the system for future directions.

N. Sooezi states that there has been a paradigm shift for the cloud computing field as it has been increasingly developed to support a lot of different workflows. This leads to a problem due to the fact that scheduling a workload on multiple clouds is highly difficult and is classified as an NP-Hard problem [6]. Therefore, the authors have presented anovel communication-based approach for scheduling the workload for a multi-cloud environment. The main drawback in this technique is the time complexity as it increases significantly based on the amount of data.

X. Zhao explains that the FY-4 system is a very complex mechanism which involves complex data structures to make the communications happen. Therefore, the authors have planned to increase the efficiency of the system by implementing a resource scheduling technique in a cloud-based platform for the FY-4 system [7]. The authors have deployed a hierarchical scheduling design to maximize the efficiency of the system. The only drawback of the system is that the system has only been in the design phase and the authors have not performed any tests on the system to ascertain its performance accurately.

Y. Cui expresses that the problem of task scheduling is one of the most important problems that plague the cloud environment now and need to be solved. As the task scheduling is not optimized, it increases the cost of execution for all the tasks, therefore, to reduce the cost of operation, the authors propose a technique to the task scheduling optimization problem that is based on the Genetic Algorithm [8]. The presented technique has been extensively tested in a simulation environment and has resulted in exceptional results. The only drawback is the increased computational load on the cloud.

X. Zhu elaborates that there is an increased need for the conservation of energy on all fronts and especially in the field of cloud computing as it would be highly beneficial even in the short term. Energy conservation would help make the system robust and also reduce operating costs. Therefore, to ameliorate these effects, the authors have presented an innovative technique for the scheduling of real-time tasks in virtualized clouds through an oriented Energy-aware scheduling algorithm. The main drawback in this paper is the lack of vertical scaling of virtual machines with respect to their CPUs. [9]

A.Rezaeianintroduces the concept of having the personal and sensitive data to be saved on the private cloud, whereas the rest of the data to be uploaded on to the public cloud to maintain the privacy [10]. To decrease this habit and provide a solution to this, the authors implemented a (BCHCS) budget constrained hybrid cloud scheduler. This

scheduler maintains the privacy of the data while scheduling the workload in a cloud environment. The only drawback in this technique is the increased time consumption on large datasets which can be attributed to the longer makespan.

M. Rekik [11] states that the Business Processes are quite difficult to schedule on a cloud environment as most of them are part of a larger complex task which needs to be performed simultaneously. This is a problem for the cloud platforms as they cannot handle concurrent operation and schedule them easily. Therefore, the authors in this paper have proposed a Context-Based Scheduling technique for the scheduling of the adaptive Business Processes executing in the cloud. the main drawback for this technique is that the system does not account for the delay and energy while scheduling the Business Processes.

R. Patel explains that the cloud has been one of the most significant advances of the 21st century combined with the services that it offers; it has been one of the fastest growing technologies. But as the traffic and the number of users increase, the cloud is tasked with a very difficult job for the management and scheduling of the tasks in a cloud environment. Therefore, the authors have proposed a technique for the scheduling of jobs based on the Hungarian Model [12]. The proposed technique has been extensively tested to prove its effectiveness.

N. Dangeelaborates that there has been an exponential increase in the number of mobile devices and smartphones in recent years. This has led to an increase in the number of users of the mobile cloud which has grown enormously due to this. But the mobile devices are unlike the personal computer counterparts, as they have a limited battery and memory which is not the case in PCs, therefore the authors have presented a technique for the scheduling of tasks in a collaborative environment on the mobile cloud platform. the specialized algorithm is capable of efficiently scheduling the various tasks in a mobile cloud platform. [13]

S. Akintoye introduces the concept of scheduling and allocation of resources in a cloud environment as one of the most important concepts for the optimization of the cloud. If the cloud is utilizing all of its resources efficiently, this would mean that the optimization of the cloud is very high [14]. To get the perfect optimization, the authors have presented a novel solution for the resource allocation problem by utilizing machine learning through a Hungarian Algorithm based Binding Policy. This technique has not been implemented in a fog-based platform which will be the future direction for the researchers.

[15] Z. Mann states that mapping virtual machines to the physical machines are a very difficult and essential problem in a cloud environment. It is one of the most essential optimization problems that can have longstanding impacts on energy consumption, performance and the running costs of the cloud. Therefore, the authors have presented a multicore-aware virtual machine placement technique for the cloud data centers, which can solve this problem and help achieve high levels of optimization in the cloud data center. The main disadvantage of this system is that it does not take into account the different resources while scheduling such as I/O and memory.

III PROPOSED METHODOLOGY

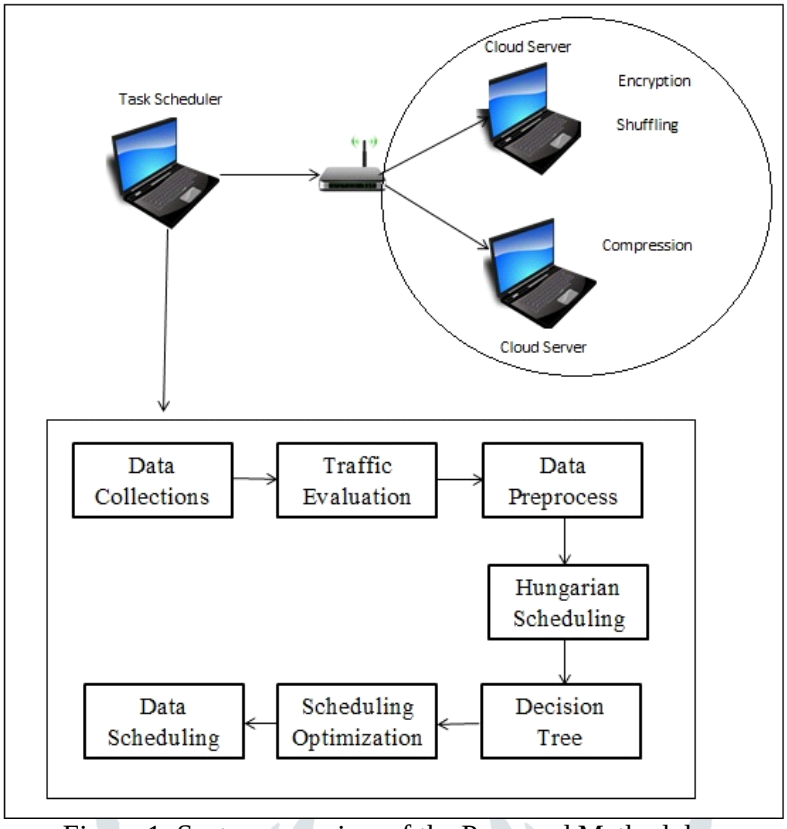

Figure 1: System overview of the Proposed Methodology

The proposed methodology for Scheduling Multi resource jobs in the cloud is depicted in figure 1. The steps that are involved in this process are broadly described in the below mentioned steps.

Step 1: Traffic Evaluation- To deploy the model of scheduling, multi resource jobs in the cloud proposed system uses the three Laptops. Where a folder containing with numerous types of files is fed to the system of a laptop and that is called as the Data generator and Task Scheduler. Whereas the other two laptops are called as the cloud servers or Job Performers.

The task scheduler needs to estimate the traffic in between itself and with the other two cloud servers. For this purpose the task server measure the timing in between the span of before sending a "KNOCK" message and after receiving "THANKS" message from both the servers on specific UDP port numbers. This time is measured in terms of milliseconds which is further used in the task of data process.

Step 2: Data Preprocessing- In the designed proposed model three jobs are decided and deployed in the two servers. First one is Data Compression, Second one is Encryption and the third one is Data Shuffling. The Task Scheduler measures the time in between the two cloud servers for these three jobs.

So three sets of time in milliseconds are measured for the three allocated jobs for the three performing threads. This time in milliseconds is arranged in the form of a matrix as shown in the figure 2.

	ST1	ST2	ST3
L1	175	179	155
L2	188	155	172
L3	156	155	161

Figure 2: Performance Time Matrix

Here each of the rows indicates the three trials for each of the Thread, that are nominated to the position of the column, This forms a matrix that is called as the performance Time Matrix $PT_{MAX.}$ In Figure 2 ST1,ST2 and ST3 are the threads that are handling the resource allocation task. Whereas the L1,L2 and L3 are the Measured time set for the Jobs of Data Shuffling, Data Encryption and Compression.

Once the Time is evaluated, then Performance time matrix is stored in an object of the matrix. Then a cluster of data is being formed for the three different tasks at the Job Scheduler Server End. These clusters are formed based on the jobs for which the data is being allocated in the cloud server's end. The cluster of .txt extension files is subject to the job of Shuffling and Encryption. Whereas the all the files are put in a cluster for the job of Compression.

Step 3: Hungarian Scheduling and Decision Making - This is the core part of the proposed model, Where a decision is being is taken to transmit the pattern of the data clusters to the cloud servers. So that the optimization in resource Scheduling can be obtained and this process contains 5 steps as described below. An abstract genetic algorithm is being used to take the decision of scheduling the resources for the jobs as described below.

Initial Population - Here the formed performance time matrix is taken as the input, based on which the scheduling is going to be decided in the coming sections.

Fitness Function - Here for the created matrix a performance time is evaluated based on the Hungarian model to call as the fitness function. This can be performed using the following Algorithm 1.

ALGORITHM 1: Hungarian Time Estimation

//Input: No. of Scheduler Threads C_r
// No. of Clusters C_l
//Time set T_{set}= ($C_{r1}C_{l1}$, $C_{r2}C_{l2,}$ $C_{r3}C_{l3}$ ..... $C_{rn}C_{ln}$)
//Output: total time
Function: Hungarian _Model(C_r, C_l, T_{set})

- Start
- total time =0
- ***for*** i=0 to size of
- ***for*** j=0 to size of
- ***if*** j=0 then
- ***if*** is engaged
- =+

Where [T: Time]

- ***else***
- ***if*** is engaged
- ***if*** >
- =++

Where[W_t = Waiting Time]

- total time=total time+
- end ***for***
- end ***for***
- ***return*** total time
- stop

Selection - Here the fitness function that is time, is selected if it is less than or equal to the past generation as the most optimized time.

Cross Over - Here the generated initial population matrix is changed according to rows that eventually indicates the clusters based on the permutation of the cluster numbers with the thread sequence. This crossover matrix is selected in the next step of genetic algorithm , i.e. Mutation.

Mutation - Here Cross over matrix is again evaluated for the fitness function and perform the selection process till it yields the best pattern of cluster commitment for the resource scheduling in the cloud for the given jobs.

Step 4: Scheduling Optimization - Here in this step the clusters are selected based on the obtained pattern numbers from the Hungarian model. Then these clusters are loaded to respective threads to route the data to the cloud servers. This leads to optimize the time taken for scheduling the multi resource jobs.

Step 5: Allocated Jobs - The proposed model is dealt with three fixed jobs that are allocated to the two cloud servers which are described as below.

Encryption - A cloud server is allocated to perform the job of encryption using the AES algorithm for the resources of .txt files. Here a secured key is being created to encrypt the data before storing it on the cloud server.

Shuffling - A Cloud server, which is given a job of shuffling performs it on the .txt file resources. To deal with this the textual characters are shuffled based on the randomly generated indices in between the range of the zero to character length.

Compression- Here Cloud server uses the Huffman Compression technique to compress the files on all formats of the data to store in the cloud server.

IV RESULT AND DISCUSSIONS

The proposed methodology for the scheduling of Multi-Resource jobs in the Cloud through the utilization of the Hungarian Algorithm has been implemented in real-time through the following hardware. The wirelessnetwork capabilities were handled by a DLink Wireless Router, connecting three laptops together.Each of the laptops has a standard configuration where it is powered by a Core i5 processor and assisted with 6 GB of physical memory utilized as RAM (Random Access Memory). All of the three laptops used for the deployment of this technique were running the Windows as their Operating System. As the presented system was coded using the Java Programming language, the model utilized the NetBeans as IDE (Integrated Development Environment).

The technique has been deployed successfully and has generated some surprising results. Therefore, to evaluate the results being generated, the data transfer rate experiment was conducted to qualitatively ascertain the dominance of the proposed methodology in comparison to that of [16]. The authors in [16] have performed various simulations for achieving the results, whereas our system has been implemented in real-time.

The synthetic traffic used to evaluate the results has not been used this implementation as the proposed methodology is deployed in real-time and uses the real-time traffic to correlate the data transfer rates from the study done in [16]. The TCP congestion control mechanism is evaluated which is based on the parameters of linear bandwidth. These parameters have been extracted from their respective systems and tabulated in table 1 given below.

Time (in Seconds)	New Reno Method (No. of Bytes)	Updated New Reno Method (No. of Bytes)	Hungarian Method (No. of Bytes)
1.02237	0.9909269	1.0020404	1.2358797
5.00649	3630.979172	3666.548178	3895.254576
10.0115	6966.665845	8358.763655	8553.972969
25.0543	17026.35237	22608.85813	25017.49578
50.0462	32286.29365	43248.17203	45652.27344
75.0173	48992.35601	52787.55067	61468.78304

Table 1: Data Transfer Rate for Varying time (in seconds)

The Tabulated data depicting the data transfer rates are then utilized to construct a graph. This graph gives us an indication of the comparable performance difference between the system proposed in [16] and our methodology. It is clearly evident that our proposed system for resource scheduling in the Cloud platform through the Hungarian Algorithm yields a higher data transfer rate in comparison to [16]. This is due to the fact that a lot of bandwidth gets piled up which contributes to the congestion and negatively impacts the performance and speed of the transmission are reduced in comparison to the Hungarian Algorithm based system. There is a significant improvement in the performance which can be validated through the observation of the Figure3 given below.

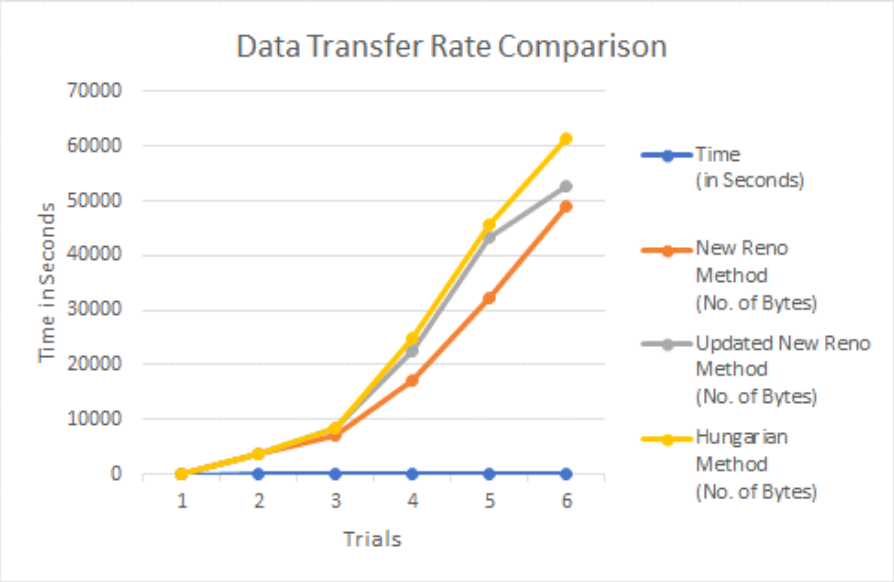

Figure 3: Data Transfer Rate Comparison

V CONCLUSION AND FUTURESCOPE

The network congestion is always a big worry in the datacenters and cloud infrastructures because of its huge amount of data movement in to and fro form within their infrastructure. Some methodologies do exist which are dealing with only bigger resource files to maintain the network traffic. Resource scheduling always plays a vital role in the business to tackle the network congestion very efficiently. The proposed model uses the three laptops and given one laptop with the role of resource scheduler server. And another two laptops are given the role of Job Performer servers or storage Servers. The data in the form of different extension of files are fed to the system of resource Scheduler where a traffic congestion is being estimated with the cloud servers. These estimated time is being used to schedule the resources for the cloud servers using the optimized Hungarian model with the abstract decision tree with the help of genetic algorithm. The comparative results clearly show that the proposed model achieves better results in scheduling by allocating more data in less time.

In the Future this model can be deployed in real time cloud to handle massive data in terms of TB.

REFERENCES

[1] F. Li and L. Zhang, "A production-based scheduling model for complex products in cloud environment", 5th International Conference on Enterprise Systems, 2017.

[2] H. Wei and F. Meng, "A Novel Scheduling Mechanism for Hybrid Cloud Systems", IEEE 9th International Conference on Cloud Computing, 2016.

[3] M. Sarvabhatla et al, "A Dynamic and Energy Efficient Greedy Scheduling Algorithm for Cloud Data Centers", IEEE International Conference on Cloud Computing in Emerging Markets, 2017.

[4] A. Panneerselvam et al, "Hyper-Heuristic MapReduce Workflow Schedulingin Cloud", Proceedings of the Second International conference on I-SMAC (IoT in Social, Mobile, Analytics, and Cloud), I-SMAC, 2018.

[5] B. Ravandi et al, "A Self-Learning Scheduling in Cloud Software Defined Block Storage" IEEE 10th International Conference on Cloud Computing, 2017.

[6] N. Sooezi et al, "Scheduling Data-Driven Workflows in Multi-Cloud Environment", IEEE 7th International Conference on Cloud Computing Technology and Science, 2015.

[7] X. Zhao et al, "A Cloud Computing Platform for FY-4 Based onResource Scheduling Technology", International Conference on Advanced Cloud and Big Data, 2016.

[8] Yang Cui, Zhang Xiaoqing, "Workflow Tasks Scheduling Optimization Based on Genetic Algorithm in Clouds", the 3rd IEEE International Conference on Cloud Computing and Big Data Analysis, 2018.

[9] X. Zhu et al, "Real-Time Tasks Oriented Energy-AwareScheduling in Virtualized Clouds", IEEE Transactions on Cloud Computing, 2013.

[10] A.Rezaeian et al, "A Budget Constrained Scheduling Algorithm for Hybrid Cloud Computing Systems Under Data Privacy", IEEE International Conference on Cloud Engineering, 2016.

[11] M. Rekik et al, "A Context-Based Scheduling Approach for Adaptive Business Process in the Cloud", IEEE International Conference on Cloud Computing, 2014.

[12] R. Patel et al, "Scheduling of Jobs based on Hungarian Method in Cloud Computing", International Conference on Inventive Communication and Computational Technologies, 2017.

[13] N. Dange et al, "Scheduling of Task in Collaborative Environment using Mobile Cloud", International Conference on Global Trends in Signal Processing, Information Computing and Communication, 2016.

[14] S. Akintoye et al, "Optimization of Virtual Resources Allocation in Cloud Computing Environment", IEEE Africon 2017 Proceedings, 2017.

[15] Z. Mann, "Multicore-aware virtual machine placement in cloud data centers", IEEE Transactions on Computers, 2016.

[16] Ramratan Ahirwal ,Ganesh Lokhande , Yogendra Kumar Jain " TCP Congestion Control through Bandwidth Estimation Mechanism in MANET " Volume 2– No.4,www.ijais.org,May 2012