International Journal of Scientific Research & Engineering Trends
Volume 5, Issue5, Sept-Oct-2019, ISSN (Online): 2395-566X

A Review on Efficient Approach for Scheduling Tasks in Cloud Computing Environments

Anjali Kumari
Dept. of Computer Science & Engg.
anjali.vnsit2010@gmail.com
Radharaman Institute of Technology
Bhopal,, MP, India

Manoj Lipton
Dept. of Computer Science & Engg.
manojlip@gmail.com
Radharaman Institute of Technology
Bhopal, MP, India

***Abstract* - Nowadays, Cloud computing is widely used in companies and enterprises. However, there are some challenges in using Cloud computing. The main challenge is resource management, where Cloud computing provides IT resources (e.g., CPU, Memory, Network, Storage, etc.) Based on virtualization concept and pay-as-you-go principle. The aim of this proposed algorithm is to minimize the completion time and cost of tasks, and maximize resource utilization. The performance of this proposed algorithm has been evaluated using Clouds toolkit. It schedules the entire task within a particular deadline while increasing the system performance. The cloud-sim simulator is used for evaluating the task scheduling like file uploading. In this module we have to initialize the performance of this proposed algorithm has been evaluated using Cloud Sim toolkit. It schedules the entire task within a particular deadline while increasing the system performance. Here we propose a VM Placement aims to distribute the dynamic workload smoothly to all the hosts in the cloud to gain an improvement in both the utilization of Resources and the speed of execution time. It allocates the incoming tasks to all available VMs. In order to achieve balancing and avoid congestion.**

***Keywords*- Cloud Computing, Network, Storage, VM, IT resources.**

I. INTRODUCTION

Cloud computing has changed the traditional large-scale computational environment by making computing resources available on a pay-per-use basis and provides a new direction for network-based applications by enabling sharing of services. The encouraging of cloud computing is to aggregate heterogeneous distributed sources to solve complicated industrial and scientific issues. The highest concentration of cloud-computing systems is about sharing Resources in large-scale multi organizational cooperation and their usage in new applications. To achieve this, an efficient scheduling system is a vital part for cloud Computing.

The dynamic and heterogeneous nature of cloud sources lead to the increased complexity of scheduling algorithms. Therefore, deterministic algorithms may not have enough efficiency to solve this issue. In this study, a new solution is presented to improve dynamic scheduling in cloud environments by combining greedy and max–min scheduling methods. The most important features of the proposed method include reduction of completion of the last task, reduction of total waiting time, observing of load balance and back up of data dynamic operations. The results of the authors' simulations show performance improvement in comparison with greedy and max–min algorithms. People use cloud computing to access a network of computers that are in a concealed state. What makes this concept scalable and efficient is the workflow management system [6]. Resources are installed on service providers' servers instead of the users 'systems. The service requested by the user for varied applications can be directed toward any data centre for any cloud server. Directing and routing of a service request to various servers are based on cloud management policies depending on server load, proximity to the database etc. It can be satisfied by providing optimized task scheduling algorithms which adapt the advantages of existing algorithms while considering the dynamism and scalability characteristics of cloud environments.

There is no exact definition of cloud but we can define cloud in various ways and by considering various means. Cloud computing is Internet-connected mode of supercomputing. It is a type of shared infrastructure, which simply puts the huge system pools together by using various means; distributed virtualization etc. It gives users a variety of storage; Networking and computing resources in the cloud computing environment via Internet users put a lot of information and accesses a lot of computing power with the help of its own computer. to R.Buyya that defines the cloud as " Cloud is a parallel and distributed computing system which basically consist of a collection of inter-connected and virtualized computers that are provisioned dynamically and presented as one or more than one unified computing

resources based on service-level agreement (SLA) established through negotiation between the service providers of cloud and users [1]. Cloud computing is a large-scale distributed computing model, which depends on the economic size of the operator of cloud that is abstract, virtualized and dynamic.

The main content of cloud computing is to manage computing power, storage, various kind of platforms and services which assigned to the external users on demand through the internet. Cloud computing is a rapidly emerging computation paradigm with the goal of freeing up users of cloud from the management of hardware, software, networks and data resources and shifting these burdens to cloud service providers[2]. Clouds provide a very large number of resources, including platforms for computation, data centers, storages, Networks, firewalls and software in form of services. At the same time it also provides the ways of managing these resources such that users of cloud can access them without facing any kind of performance related problems. Cloud Computing Services are divided into three classes, according to the abstraction level and the service model of providers, namely:

- Infrastructure as a Service
- Platform as a Service
- Software as a Service

Distribution, virtualization and elasticity are the basic characteristics of cloud computing. Virtualization is one of the main features of cloud. Most of the software and hardware have provided support to virtualization. We can perform virtualization on many factors such as hardware, software, storage and operating system, and manage them in cloud platform.

1. Overview Task Scheduling- Cloud consists of a number of resources that are different with one other via some means and cost of performing tasks in cloud using resources of cloud is different so scheduling of tasks in cloud is different from the traditional methods of scheduling and so scheduling of tasks in cloud need better attention to be paid because services of cloud depends on them. Task scheduling plays a key role to improve flexibility and reliability of systems in cloud.

The main reason behind scheduling tasks to the resources in accordance with the given time bound, which involves finding out a complete and best sequence in which various tasks can be executed to give the best and satisfactory result to the user. In cloud computing, resources in any form i.e. cups, firewall, network are always dynamically allocated according to the sequence and requirements of the task, subtasks. So, this leads task scheduling in cloud to be a dynamic problem means no earlier defined sequence may be useful during processing of task. The reason behind the scheduling to be dynamic is that because flow of task is uncertain, execution paths are also uncertain and at the same time resources available are also uncertain because there is a number of tasks are present that are sharing them simultaneously at the same time. The scheduling of tasks in cloud means choose the best suitable resource available for execution of tasks or to allocate computer machines to tasks in such a manner that the completion time is minimized as possible. In scheduling algorithms list of tasks is created by giving priority to each and every tasks where setting of priority to different tasks can be based on various parameters. Tasks are then chooses according to their priorities and assigned to available processors and computer machines which satisfy a predefined objective function [12].

1. Scheduling Types- Static scheduling schedule tasks in known environment i.e. it already has the information about complete structure of tasks and mapping of resources before execution, estimates of task execution/running time. Dynamic scheduling must depend on not only the submitted tasks to cloud environment but also the current states of system and computer machines to make scheduling decision.

General View of Task Scheduling Cloud computing uses virtualization technique for mapping the resources of cloud to the virtual machine layer, implement the user's task, so the task scheduling of cloud computing environment achieve at the applications layer and the virtual layer of resources. Scheduling is nothing but the mapping of tasks and resources in accordance with some certain principles for achieving the desired goal. Cloud computing paradigm simplifies the mapping of tasks to resources; the required resources together form to be virtual machines (VMs), In consequence, user does not actually control the physical storage of their data, which results in the separation of ownership and management of data. The CSP can freely access and search the data stored in the cloud. Meanwhile the attackers can also attack the CSP server to obtain the user's data. The above two cases both make users fell into the danger of information leakage and data loss.

II. RELATED WORK

Cloud computing has changed the traditional large-scale computational environment by making computing resources available on a pay-per-use basis and provides a new direction for network-based applications by enabling sharing of services. The encouraging of cloud computing is to aggregate heterogeneous distributed sources to solve complicated industrial and scientific issues. The highest concentration of cloud-computing systems is about sharing resources in large-scale multi organisational cooperation and their usage in new applications. To achieve this, an efficient scheduling system is a vital part for cloud computing. The dynamic and heterogeneous nature of cloud sources lead to the increased complexity

of scheduling algorithms. Therefore, deterministic algorithms may not have enough efficiency to solve this issue. In this study, new solutions presented to improve dynamic scheduling in cloud environments by combining greedy and max–min scheduling methods. The most important features of the proposed method include reduction of completion of the last task, reduction of total waiting time, observing of load balance and back up of data dynamic operations. The results of the authors' simulations show performance improvement in comparison with greedy and max–min algorithms.

III. LITERATURE SURVEY

Cloud computing is a large-scale distributed computing model, which depends on the economic size of the operator of cloud that is abstract, virtualized and dynamic. The main content of cloud computing is to manage computing power, storage, various kind of platforms and services which assigned to the external users on demand through the internet. Cloud computing is a rapidly emerging computation paradigm with the goal of freeing up users of cloud from the management of hardware, software, networks and data resources and shifting these burdens to cloud service providers[2]. There are different scheduling algorithms available for cloud computing. Two major parameters are task length and deadline considered in the proposed approach. There are some other parameters which tempt the utilization of resources and scheduling of tasks.

Improved Cost Based Algorithm [4] this algorithm ameliorates the traditional cost-based scheduling algorithm for making appropriate resource allocation. The tasks are grouped as per processing power of resources. Earliest Feasible Deadline First [5] the principle of this algorithm is very simple to understand. In this scheduling algorithm, the task having the shortest deadline is getting scheduled. Earliest Deadline First (EDF) is a dynamic scheduling. Whenever a scheduling event occurs (end of a task, the release of the new task.) then the queue will be searched for the process that is closest to its deadline, the found process will be the next that is going to be scheduled for execution.

Leila Zohrati [1] Cloud computing has changed the traditional large-scale computational environment by making computing resources available on a pay-per-use basis and provides a new direction for network-based applications by enabling sharing of services. The encouraging of cloud computing is to aggregate heterogeneous distributed sources to solve complicated industrial and scientific issues. The highest concentration of cloud-computing systems is about sharing resources in a large-scale multi organizational cooperation and their usage in new applications. To achieve this, an efficient scheduling system is a vital part for cloud computing. The dynamic and heterogeneous nature of cloud sources lead to the increased complexity of scheduling algorithms. Therefore, deterministic algorithms may not have enough efficiency to solve this issue. In this study, new solutions presented to improve dynamic scheduling in cloud environments by combining greedy and max–min scheduling methods. The most important features of the proposed method include reduction of completion of the last task, reduction of total waiting time, observing of load balance and back up of data dynamic operations. The results of the author's simulations show performance improvement in comparison with greedy and max–min algorithms.

Jian Shen [2] Data sharing in cloud computing enables multiple participants to freely share the group data, which improves the efficiency of work in cooperative environments. However, how to ensure the security of data sharing within a group and how to efficiently share the outsourced data in a group manner are formidable challenges. Note that key agreement protocols have played a very important role in secure and efficient group data sharing in cloud computing.

In this paper, by taking advantage of the symmetric balanced incomplete block design (SBIBD), we present a novel block design-based key agreement protocol that supports multiple participants, which can flexibly extend the number of participants in a cloud environment according to the structure of block design. Based on the proposed group data sharing model, we present general formulas for generating the common conference key K for multiple participants. Note that by benefiting from the (v,k+1,1)-block design, the computational complexity of the proposed protocol linearly increases with the number of participants and the communication complexity is greatly reduced. In addition, the fault tolerance property of our protocol enables the group data sharing in cloud computing to withstand different key attacks, which is similar to Yi's protocol.

Yang Chen[3] give full play to the big data in the cloud computing environment application advantages has become an important information service mode of the era of internet paper with information service quality evaluation as the main line, using fuzzy comprehensive evaluation method to analyze a set of cloud computing environment information service quality evaluation process, at the same time, using the case of project construction example analysis, the evaluation of cloud computing environment based on large data information service quality has important reference value.

1. A Priority based Job Scheduling Algorithm [4] in this algorithm the approach is presented for job scheduling by using mathematical calculations. In this algorithm for scheduling the priority is considered and

each job request for resources with some priority. This paper discusses issues related to the algorithm such as consistency, complexity and makes span i.e. finish time. As per author make span can be reduced further by improving the algorithm.

2. A Priority based Scheduling Strategy for VM Allocation [11] the objective of proposed algorithm is to gain more benefits to the service providers since the current resources are not adequate to process all the requests. This paper proposes a priority algorithm to find the best choice. This technique can increase the benefits than applying typical FCFS strategy. If more information can be made available, e.g. the regular pattern of the usage the algorithm can be improved further.

3. Greedy Based Job Scheduling Algorithm [8] this algorithm focuses on QoS, as the cloud computing is a business oriented service. The goal of the algorithm is to reduce completion time and to give faster solution to scheduling problem. This algorithm classifies tasks based on QoS and then as per the task category, the appropriate function is assigned. The results of the algorithm are compared with other existing algorithm that are, the algorithm based on Berger model and existing scheduling strategy of Cloud Sim tool. Improved Priority based Job result of the Hybrid Cuckoo algorithm is compared with FIFO, Genetic Algorithm and Cuckoo algorithms.

4. Credit Based Scheduling Algorithm [6] Cloud providers have limited amount of resources, and are thus obliged to strive to maximize utilization. In this paper scheduling algorithm introduced which are based on user priority and task length. Credits are assigned to task length and task priority. And by calculating the credit value the jobs are scheduled to be executed. The proposed algorithm works more efficiently than previous methods. The proposed scheme can be enhanced so as to consider other parameter like a deadline.

5. Improved Max-Min Task Scheduling Algorithm [7] In paper results are not obtained, but they only reviewed the Max-Min algorithm and proposed Improved Max-Min algorithm. The Improved Max-Min algorithm overcomes the anomalies of Max-Min algorithm. The proposed approach considers the user priority. The Improved Max-Min algorithm reduces more make span than traditional Max-Min algorithm.

6. Priority Based Earliest Deadline First Scheduling Algorithm [9] In this method there are two task scheduling algorithms are used, one is Earliest Deadline First and other is priority based scheduling algorithm. This algorithm focuses on resource allocation and memory utilization. The given approach will reduce the execution time of pre-empted tasks and they can schedule efficiently. This algorithm overcomes the waiting time problem of pre-empted tasks. The waiting queue is introduced which processes the pre-empted task Output information generated by simulations. There is no exact definition of cloud but we can define cloud in various ways and by considering various means. Cloud computing is Internet-connected mode of supercomputing. It is a type of shared infrastructure, which simply puts the huge system pools together by using various means; distributed, virtualization etc. It gives users a variety of storage, networking and computing resources in the cloud computing environment via Internet, users put a lot of information and accesses a lot of computing power with the help of its own computer. According to R.Buyya that defines the cloud as "Cloud is a parallel and distributed computing system which basically consist of a collection of inter-connected

Table 1.Summarization of Various Scheduling Technique.

Algorithms	Algorithms Objective	Algorithms Objective Scheduling	Scheduling Parameters	Future Work Tool	Tool
A Priority based Job Scheduling Algorithm [11]	To reduce makespan	Priority Based Job Scheduling, Multiple Criteria Decision Making Mode	Priority	To gain less finish time.	-Cloud Sim
A Priority based Scheduling Strategy for VM Allocation [10]	To maximize benefits of the service provider and improve resource utilization	Priority Based VM Scheduling	Priority of Jobs	-	-
Greedy Based Job Scheduling Algorithm [8]	To improve QoS and reduce make span	Greedy Algorithm Based Job Scheduling	Time Preference, Bandwidth Preference, Expectation Time Preference, JEF Function and function result	Improvement can be done in the proposed algorithm to reduce completion time and	-

				to gain more fairness	
Credit Based Scheduling Algorithm [6]	To improve resource utilization and reduce makes pan. Cloud Sim	Credit Based Scheduling	User Priority and Task Length	The proposed algorithm can be enhanced further considering deadline as a scheduling parameter	Clouds
Priority Based Earliest Deadline First Scheduling Algorithm [9]	To reduce the average waiting time further. Cloud Sim	Priority Based Scheduling	Priority and Deadline of tasks	Waiting time can reduce further	Cloud Sim

and virtualized computers that are provisioned dynamically and presented as one or more than one unified computing resources based on service-level agreement (SLA) established through negotiation between the service providers of cloud and users [1]. Cloud computing is a large-scale distributed computing model, which depends on the economic size of the operator of cloud that is abstract, virtualized and dynamic. The main content of cloud computing is to manage computing power, storage, various kind of platforms and services which assigned to the external users on demand through the internet. Cloud computing is a rapidly emerging computation paradigm with the goal of freeing up users of cloud from the management of hardware, software, networks and data resources and shifting these burdens to cloud service providers[2]. Clouds provide a very large number of resources, including platforms for computation, data centers, storages, Networks, firewalls and software in form of services. At the same time it also provides the ways of managing these resources such that users of cloud can access them without facing any kind of performance related problems. Cloud Computing Services are divided into three classes, according to the abstraction level and the service model of providers, namely:

- Infrastructure as a Service
- Platform as a Service
- Software as a Service

Distribution, virtualization and elasticity are the basic characteristics of cloud computing. Virtualization is one of the main features of cloud. Most of the software and hardware have provided support to virtualization. We can perform virtualization on many factors such as hardware, software, storage and operating system, and manage them in cloud platform.

IV. PROPOSED APPROCH

The proposed system is introduced to overcome all the disadvantages that arise in the existing system. Our proposed system will schedule the entire task with in a particular deadline while increasing the system performance.

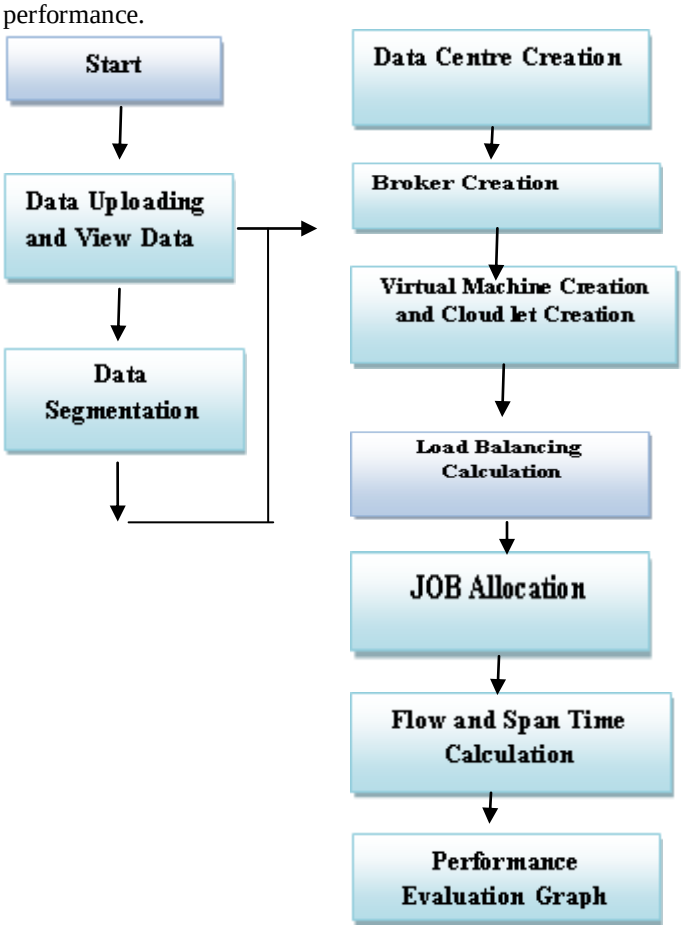

Fig.1 Flow Chart of System.

Propose a VM Placement aims to distribute the dynamic workload smoothly to all the hosts in the cloud to gain an improvement in both the utilization of Resources and the speed of execution time. It allocates the incoming tasks to all available VMs. In order to achieve balancing and avoid congestion, the proposed algorithm allocates tasks to the least loaded VM and prevents the allocation of tasks to a VM when the variation of this VM processing time from average processing time of all VMs becomes more than or equal to a threshold value. This leads to a reduction of the overall response time and the processing time of hosts.

Modules

- Choose File and Read
- Change File Content and Update
- Upload the Data (File)
- Segment the Data
- Data Centre Creation
- Broker Creation
- Virtual Machine Creation
- Job Creation
- Load Balancing Calculation
- Virtual Machine Allocation
- Flow and Completion Time Calculation
- Performance Analysis Report

1. Data centre- It is set of hosts or servers which provide infrastructure service. Data center configuration can be heterogeneous or homogeneous resources.

2. Hosts- It is the physical entity that is resource to tasks.
Job: Job is the task

3. Service Broker- it decides which VM will provide service to request.

4. VM allocation-These policies in Cloud Sim model the allocation of resources to tasks. Load balancing between the virtual machine is done by Cloud Simulator.

We have to analyze which virtual machine has highest space for resource. For achieving load balancing we assign job to virtual machines. Then initialize all ram size broker id, job id and host id for the particular virtual machine. Finally we have to analyses the work load of the jobs which is in the job. We also analyses the following in Bar chart.

- Work Load
- Flow
- Cost

The cloud-sim simulator is used for evaluating the overall performance of the task scheduling.

V. CONCLUSION

The proposed algorithm will be targets to minimize completion time and cost, and maximize resource utilization. All the tasks are effectively scheduled with in a particular deadline. The performance of the proposed system is highly increased while reducing the task completion time and increasing the system resource utilization The File is uploaded for scheduling the task. The Task is considered as a file uploading. All the files are uploaded into the cloud server within the particular time limit. The files are segmented into number of blocks and then processed. The cloud-sim simulator is used for evaluating the task scheduling like file uploading. In this module we have to initialize Task scheduling is one of the most famous problems in cloud computing so; there is always a chance of modification of previously completed work in this particular field. The researchers at their own time performed their work according to their knowledge space and after some time their work had been carried out some other people. During scheduling they had considered various techniques and applied constraints but as the cloud computing is too vast that they had not been able to capture all aspects at the same time but they mentioned these facts that there is a chance of modification of algorithms and which part has to be modified.

REFERENCES

[1]. Leila Zohrati, Maryam Nooraei Abadeh , Elham “flexible approach to schedule tasks in cloud computing environments” ISSN 1751-8806 Received on 16th January Revised doi 10.1049/iet-sen.2017.0008 The Institution of Engineering and Technology 2018.

[2]. Jian Shen Tianqi ,Zhou Debiao He Yuexin Zhang Xingming Sun Yang Xiang “Block Design-based Key Agreement for Group Data Sharing in Cloud Computing” IEEE Transactions on Dependable and Secure Computing Year: 2017 Early Access Article Publisher: IEEE

[3]. Xu Xiao-tao ,Yang Chen Dai ,Guang-hua Ma Hua-“long Information service quality evaluation study of cloud computing environment based on big data” "2017 IEEE 2nd International Conference on Cloud Computing and Big Data Analysis (ICCCBDA) Year: 2017

[4]. P. Geetha ,C.R. Rene Robin “A comparative-study of load-cloud balancing algorithms in cloud environments” 2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS) Year: 2017

[5]. Sruti Basu , Prasant Kumar, Pattnaik “A consistency preservation based approach for data-intensive cloud computing environment” 2017 8th International Conference on Computing, Communication and Networking Technologies (ICCCNT) Year: 2017

[6]. Thomas, Antony, G. Krishnalal, and VP Jagathy Raj. "Credit Based Scheduling Algorithm in Cloud Computing Environment." Procedia Computer Science 46 (2015): 913-920.

[7]. Bhavisha Kanani, Bhumi Maniyar, "Review on Max-Min Taskscheduling Algorithm for Cloud Computing", Journal Of Emerging Technologies And

Innovative Research (JETIR), Volume 2, Issue 3,March 2015

[8]. Li, Ji, Longhua Feng, and Shenglong Fang. "An greedy-based job scheduling algorithm in cloud computing." Journal of Software 9.4 (2014): 921-925.

[9]. Gupta, Gaurav, et al. "A simulation of priority based earliest deadline first scheduling for cloud computing system." Networks & Soft Computing (ICNSC), 2014 First International Conference on. IEEE,2014.

[10]. xiao, Jing, and Zhiyuan Wang. "A Priority Based Scheduling Strategy for Virtual Machine Allocations in Cloud Computing Environment." Cloud and Service Computing (CSC), 2012 International Conference on. IEEE, 2012.

[11]. Ghanbari, Shamsollah, and Mohamed Othman. "A priority based job scheduling algorithm in cloud computing."Procedia Engineering 50 (2012): 778-785.