RESEARCH ARTICLE OPEN ACCESS

Survey of Optimization of Scheduling in Cloud Computing Environment

Er.Mandeep kaur[1], Er.Rajinder kaur[2], Er.Sughandha Sharma[3]
Research Scholar[1 & 2] Department of Computer Science and Engineering, Assistant Professor[3]
CGC College of Engineering, Gharuan
Punjab-India

ABSTRACT
Cloud computing environments facilitate applications by providing virtualized resources that can be provisioned dynamically. The advent of Cloud computing as a new model of service provisioning in distributed systems, encourages researchers to investigate its benefits and drawbacks in executing scientific applications such as workflows. There are a mass of researches on the issue of scheduling in cloud computing, most of them, however, are about workflow and job scheduling eared to be lagging behind the rapid developments in this field. This paper is the first systematic review of peer-reviewed academic research published in this field, and aims to provide an overview of the swiftly developing advances in the technical foundations of cloud computing and their research efforts. Structured along the technical aspects on the cloud agenda, we discuss lessons from related technologies advances in the introduction of protocols, interfaces, and standards. Techniques for modelling and building clouds and new use-cases arising through cloud computing.

Keywords:- SAAS, PAAS, IAAS, NAAS, Cloudsim, CIS

I. INTRODUCTION

With the rapid development of storage technologies, processing and success of internet, computing resources have become more powerful and cheaper than even before. This technological fashion lead to the realization of new computing model called cloud computing in which resources are provided as utilities which can be leased or released by various users through internet in an on demand fashion. Thus cloud computing is recently emerged as a new technology that hosting and delivering new services via internet.

Cloud Computing is a type of computing that involves a large number of computers that are connected in a network such as internet. We can also say that it is a synonym of distributed computing. In which there are large number of computers that are operating and managed at the same time. Cloud computing is a development of grid computing, parallel computing and distributed computing. Generally cloud refers to the Software, Infrastructure, and platform that are sold 'as a service' over the internet. There are various number of cloud networks like public cloud, private cloud and hybrid cloud. Cloud computing is basically a combination of two things that is Online application and Online Storage. Gmail is an excellent example. If you are using the various social networking sites like Gmail, yahoo etc. then you are cloud computing user. These software applications are not installed on your computer but you are using them over the internet. Similarly if you are iphone user and you have enabled icloud then your apps, videos and photos are backed up or stored by the computer managed by the apple and the data will be transfer to that computer by the internet. There are various cloud computing examples like Amazon, Google, Oracle Cloud and SalesForce etc. Amazon EC2 (Amazon elastic compute cloud) enables the different cloud users to launch and manage the various server instances using the application programming interface (API) or available tools or utilities. EC2 provides the ability to create the instances at multiple locations. Apache Hadoop is an open source software framework for large scale processing and storage for data sets.

Now Cloud computing reached at the point where it can take place of your entire Operating system. The underlying concept of cloud computing date back to 1950's when large mainframe computers were used and often referred to as 'Static Terminals or computers' because they were used for communication but had no internal processing. In early 2008, Eucalyptus became the first open source API (application programming interface) compatible platform for deploy the private clouds. The main enabling technology for cloud computing is virtualization that hides the heterogeneity of cloud resources and generalize the physical infrastructure that is a rigid component and makes it available as a soft component that is easy to use and manage. The virtualized server is called virtual machine. There are various service models exists in cloud computing which are saas, paas, iaas and Naas.

- ***(i)*** **Software as a service (saas):** It is referred to as on-demand software and usually price on pay-per-use basis. The users have right to access an application software and databases.
- ***(ii)*** **Platform as a service (paas):** In this type of models, cloud provider delivers a platform that typically including the programming language, operating

system and database. The application developers can develop and run their software on a cloud platform without any cost of buying the underlying hardware or software.

***(iii)* Infrastructure as a service (iaas):** In the most basic, providers of Iaas offer computers- physical or virtual machines or various resources. In short, infrastructure as a service provide the infrastructure to develop the applications by the application developers.

***(iv)* Network as a service (naas):** This is the least common model where the user is provided the network connectivity services such as VPN or bandwidth usage on demand.

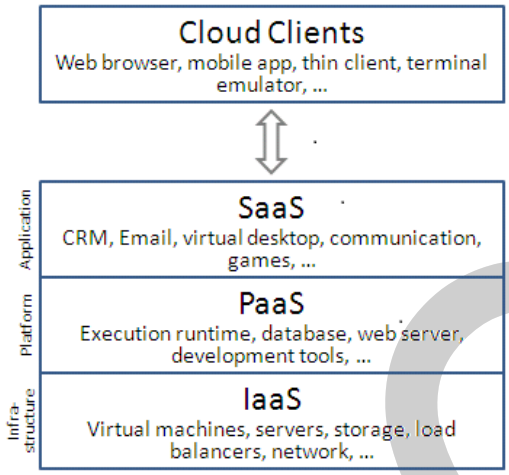

Fig1. Service models of cloud computing

Since cloud computing is heterogeneous pool of resources so scheduling plays a major role in this field. Scheduling tasks to the cloud resources is an important step in cloud source management. Scheduling or job scheduling is a task of assigning the system resources to the various tasks that are waiting for the CPU time and emerged in a queue. The system must decide that which particular job took first give it the CPU time for processing, So that all the jobs can executed in fair and efficient manner. Also fairness in scheduling is the important criteria that provides the resource allocation in an optimized way and improves efficiency.

As the cloud scale expanding, scheduling is still an issue to resolve. An efficient scheduling mechanism should meet the Qos parameters and enhance resource utilization. The main home to the computation power and storage is the Data centre that is central to cloud computing and provides the thousands of devices like servers, routers and switches. Data centre compress of basic layers which are core layer, aggregation layer and access layer. In the access layer the servers are physically connected in the rack form. There are 20 to 40 servers per rack. The aggregation layer provides the functions like domain service, load balancing and more. And the last layer i.e. core layer that provides connectivity to multiple aggregate switches with no single point of failure.

II. PROPOSED WORK

The proposed method of job scheduling in this paper is that to more optimize the way of scheduling by various machine learning algorithms so that there will be a linear or non linear mapping of tasks to resources and scheduling by enhancing the Berger model theory of job scheduling, in which the concept of distributive justice is used. Distributive justice is based on expectation states. Expectation states are used to present the justice or injustice by distributing the resources under various circumstances.

The proposed scheduling algorithm is in cost effective manner with short make span and meets the various Qos parameters like bandwidth, utilization rate and time.

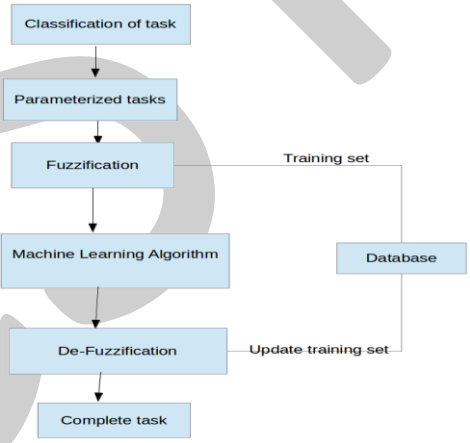

Fig 2: Data flow diagram of proposed methodology

Cloud computing provides the number of advantages to the end users and business of all sizes. The major advantage is that you don't need to have the knowledge to develop and maintain the infrastructure. The others are cost efficiency, convenience and continuous availability, backup and recovery, scalability and performance, increase storage capacity. But if there are positive sides then there exists the negative effects also which are related to security and privacy in cloud, limited control and flexibility and increased vulnerability. But despite of its disadvantages, cloud computing remains strong.

There is a hope of mitigation of disadvantages and advantages will further grow. Since cloud computing seems to have made IT field little bit easier.

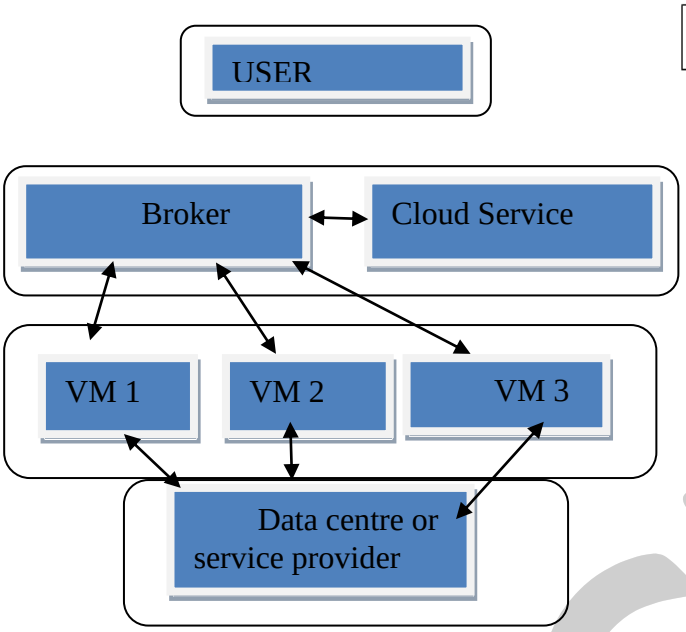

Fig 3 Scheduling in clouds

III. CLOUDSIM

Simulation is a technique where a program models the behavior of the system (CPU, network etc.) by calculating the interaction between its different entities using mathematical formulas, or actually capturing and playing back observations from a production system. Cloudsim is a framework developed by the GRIDS laboratory of university of Melbourne which enables seamless modelling, simulation and experimenting on designing cloud computing infrastructures.

3.1 Cloudsim Data Flow

Each data centre entity registers with the Cloud Information Service registry (CIS). CIS provides database level match-making services; it maps user requests to suitable cloud providers. The DataCenterBroker consults the CIS service to obtain the list of cloud providers who can offer infrastructure services that match application's quality of service, hardware, and software requirements. In the case match occurs the broker deploys the application with the cloud that was suggested by the CIS . Cloudsim data flow is shown in Figure 3.1

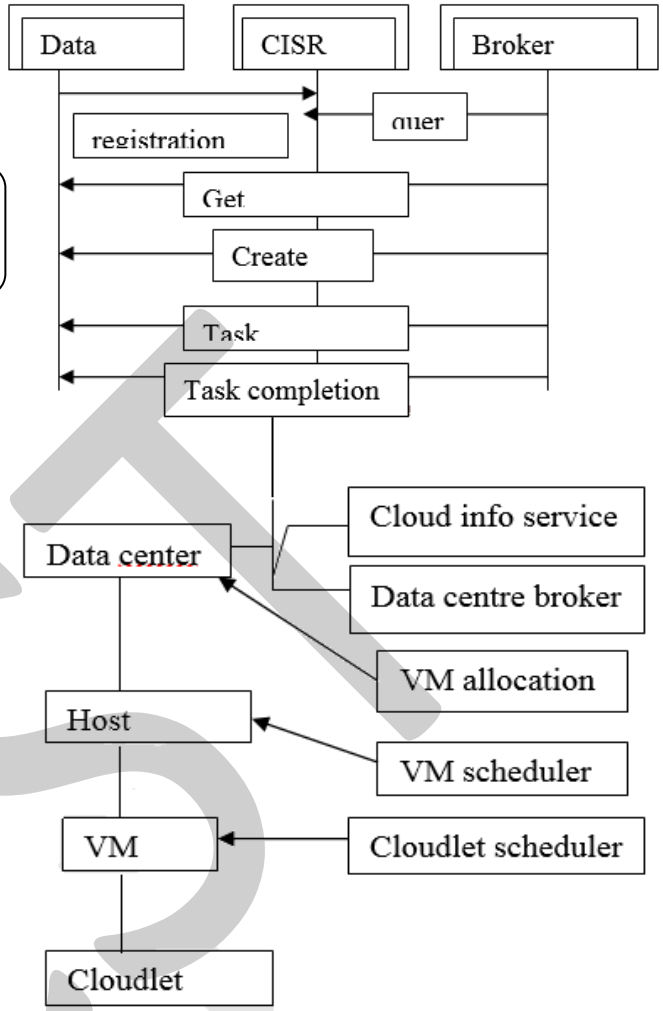

Figure5. Main parts of cloud sim related to our experiment

IV. RELATED WORK

jia ru. Introduces an effective scheduling algorithm, which attempts to maximize the cloud resource utilization, improve the computation ratio and reduce makespan, overhead and delay in cloud based systems by using the method of analysis of different scheduling algorithms which can be adopted in cloud based systems and simulate these using cloud sim. Then evaluate both of their advantages and disadvantages. Then propose an improved scheduling algorithm and verify the proposed algorithm in cloud sim. Study MapReduce framework and analyze the mapreduce scheduling algorithms. The improved mapreduce scheduling policy can improve the utilization rate of resources and reduce workload of nodes. In this paper various scheduling mechanisms are used like on line scheduling problem, batch scheduling problem and mapreduce. The author first analyze the different scheduling policies and improve some scheduling algorithms. Then analyze the Mapreduce algorithm and optimize it. The proposed algorithm is an attempt to reduce

network delay, maximize the utilization and achieve minimum waiting time.

B. Anuradha S. Rajasulochana introduces that Since cloud has a heterogeneous pool of resources, scheduling plays a vital role in cloud computing. It is all about executing the current jobs under given constraint. Fairness in scheduling is an important criterion that improves the efficiency and provide optimal resource allocation. Scheduling tasks to cloud resources is the main step in cloud management. Thus the main aim of scheduling algorithm is to assign tasks to available resources in cost effective manner. Efficient scheduling mechanism should meet the Qos parameters and should enhance resource utilization. With the cloud scale expansion fairness in scheduling and resource allocation is still an issue.

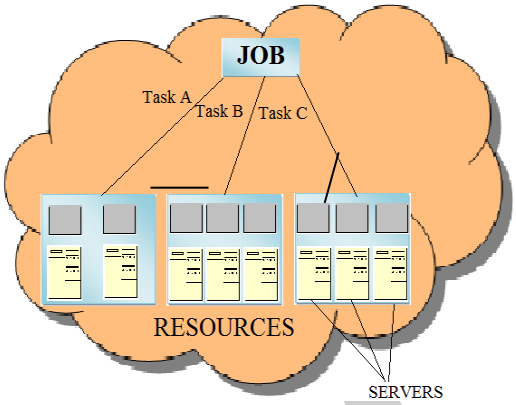

Figure 6: Scheduling tasks to available resources.

Fairness is a challenging factor when the tasks waiting for the resources are to be mapped. The task resource mapping can be compared to the commodity economy model similar to pay per use basis. Fairness can be dealt with user as well as resources. From the user perspective fairness is to provide user satisfaction and from the resource perspective it is effective resource utilization. The author used many approaches for static as well as dynamic mapping of resources. In future the resources can be effectively scheduled by leveraging techniques like scheduling and resource provisioning, task scheduling, scheduling and load balancing where efficiency and fairness will be the end goal.

Baomin Xu, Chunay Zhao, Enzhao Hu, Bin Hu. The paper proposed for the first time an algorithm of job scheduling based on Berger model. In the job scheduling process, the algorithm establishes dual fairness constraint. The first constraint is to classify user tasks by QoS preferences, and establish the general expectation function in accordance with the classification of tasks. In cloud computing, you need to face a variety of tasks from different users. With the number increase of users, cloud-scale expansion and so on, the key issues of cloud computing are to ensure an equitable opportunity for the use of resources. The aim is to enable the user needs to get more satisfaction. The Berger model of distributive justice is based on expectation states. In order to be able to map the theory of distributive justice in Berger model to resource allocation model in cloud computing, it is need to carry on the task classification, fairness function definition of user tasks, the task and resource parameterization, the task and resource mapping, and etc. Through the expansion of Cloud Sim platform, job scheduling algorithm based on Berger model is implemented. The validity of the algorithm is verified on the extended simulation platform. By comparing of simulation results with the optimal completion time algorithm, the proposed algorithm in this paper is effective implementation of user tasks, and with better fairness.

Jinhua Hu, jianhua Gu, Tianhai Zhao. In view of the current load balancing in VM resources scheduling, this paper presents a scheduling strategy on VM load balancing based on genetic algorithm. Considering the resources scheduling in cloud computing environment and with the advantage of genetic algorithm, this method according to historical data and current states computes in advance the influence it will have on the whole system when the current VM service resources that need deploying are arranged to every physical node, and then it chooses the solution which will have the least influence on the system after arrangement. In this way, the method achieves the best load balancing and reduces or avoids dynamic migration thus resolves the problem of load imbalancing. In genetic algorithm, fitness function is the criterion for the quality of the individuals in the population. It directly reflects the performance of the individuals.

Yogita Chawla, Mansi Bhonsle stated that Cloud computing is based on the concepts of distributed computing, grid computing, utility computing and virtualization. It is a virtual pool of resources which are provided to users via Internet. Cloud computing service providers' one of the goals is to use the resources efficiently and gain maximum profit. This leads to task scheduling as a core and challenging issue in cloud computing. This paper gives different scheduling strategies and algorithms in cloud computing which helps to understand the wide tasks of various scheduling options to select one for a given environment.

Dr. Ajay jangra, **Tushar Saini.** Cloud computing is a most recent new computing paradigm where applications, records and IT services are provided over the Internet. Cloud computing has come out to be an interesting way of changing the whole computing Schedulers for cloud computing determine on which processing resource jobs of a workflow should be allocated. Scheduling theory for cloud computing is in advance a lot of awareness with increasing popularity in this cloud area. The received tasks are grouped on the basis of data and requested resources by the task and prioritized. Resource selection is done on the basis of its cost and turnaround time. Task selection is on the basis of a priority formula. This way of resource selection and task selection

gives better results over sequential scheduling. This paper will give the way for the future findings related to scheduling techniques.

Ranjana saini , Indu stated that Cloud computing is based on the concepts of distributed computing, grid computing, utility computing and virtualization. It is a virtual pool of resources which are provided to users via Internet. This leads to job scheduling as a challenging issue in cloud computing. In this research a discussion towards the resource management of virtual machines in cloud and how to make resources more efficiently available to clients is provided. The main focus is on job scheduling. In this present work, there is a parametric analysis performed to identify the requirement of process migration and according to this analysis the migration will be performed on these processes. The effectiveness of the work is identified in terms of successful execution of the processes within the time limits.

REFERENCES

[1]"Job Scheduling algorithm using Berger model in Cloud Environment ", Baomin Xu, Chunyan Zhao, Enzhao Hu, Bin Hu,, Elsevier in Advances in Engineering Software, Vol. 42 , No. 7, Pp. 419-425, 2011

[2]"Particle swarm optimization scheme to solve resource-constrained scheduling problem". Chen, R. M.; Wu, C. L.; Wang, C. M. & Lo, S. T. (2010). Expert systems with applications, Vol. 37, No. 3 (March 2010), *pp. 1899-1910, ISSN:* 0957-4174.

[3] "A Berkeley view of cloud computing", M. Armbrust, A. Fox, R. Grifth, A. D. Joseph, R. Katz, A. Konwinski,G. Lee, D. Patterson, A. Rabkin, I. Stoica, and M. Zaharia. Above the clouds Technical report, University of California at Berkeley, February 2009.

[4]"A Survey of Proposed Job Scheduling Algorithms in Cloud Computing Environment" Rohit O. Gupta, Tushar Champaneria, Computer Science & Engineering, L.D.College of Engineering, India,Volume 3, Issue 11, November 2013.

[5] "Modeling and Simulation of Scalable Cloud Computing Environments and the CloudSim Toolkit:Challenges and Opportunities",Rajkumar Buyya, Rajiv Ranjan, Rodrigo N. Calheiros, in The 2009 International Conference on High Performance Computing and Simulation, HPCS2009, pp:1-11.

[6] "An Optimistic Job Scheduling Strategy based on QoS for Cloud Computing", Huang Q.Y., Huang T.L IEEE International Conference on Intelligent Computing and Integrated Systems (ICISS), Guilin, pp. 673-675, 2010.

[7]"Research issues in Cloud Computing",V.Krishna Reddy, B. Thirumala Rao , LSS Reddy. Global Journal Computer Science & Technology Vol. 11, no. 11, June 2011,pp.70-76.

[8] "Tasks scheduling optimization for the cloud computing systems" Sandeep Tayal, international journal of advanced engineering sciences and technologies vol no. 5, issue no. 2, 111 – 115.

[9] Li Jianfeng, Peng Jian. Task scheduling algorithm based on improved genetic algorithm in cloud computing environment[J].JournalofComputerApplications,2011,1(31):184~186

[10] "Job Scheduling Model for Cloud Computing Based on Multi-Objective Genetic Algorithm", Jing Liu , Xing-Guo Luo, Xing-Ming Zhang3, Fan Zhang and Bai-Nan Li. International Journal of Computer Science Issues, Vol. 10, Issue 1, No 3, January 2013.

[11] "Fairness As Justice Evaluator In Scheduling Cloud Resources: A Survey" B. Anuradha, S. Rajasulochana. International Journal of Computer Engineering & Science, Nov. 2013. ©IJCES ISSN: 22316590