Asian Journal of Electrical Sciences
ISSN: 2249-6297 (P) Vol.3 No.2, 2014, pp.27-33

DOI: https://doi.org/10.51983/ajes-2014.3.2.1930

# Survey on Scheduling Algorithms in Cloud

**A.Sathya Sofia[1] and M. Backialakshmi[2]**
PSNA College of Engg and Tech, Dindigul, Tamil Nadu, India
E-mail:backialakshmidgl@gmail.com
(Received 10 July 2014; Revised 1 August 2014; Accepted 2 September 2014; Available online 9 September 2014)

***Abstract* - Cloud computing is a general term used to describe a new class of network based computing that takes place over the internet. The primary benefit of moving to Clouds is application scalability. Cloud computing is very beneficial for the application which are sharing their resources on different nodes. Scheduling the task is quite a challenging in cloud environment. Usually tasks are scheduled by user requirements. New scheduling strategies need to be proposed to overcome the problems proposed by network properties between user and resources. New scheduling strategies may use some of the conventional scheduling concepts to merge them together with some network aware strategies to provide solutions for better and more efficient job scheduling. Scheduling strategy is the key technology in cloud computing. This paper provides the survey on scheduling algorithms. There working with respect to the resource sharing. We systemize the scheduling problem in cloud computing, and present a cloud scheduling hierarchy.**
***Keywords:* Scheduling, Cloud computing, Resource allocation, Efficiency.**

## I.INTRODUCTION

The latest innovations in cloud computing are making our business applications even more mobile and collaborative, similar to popular consumer apps like Facebook and Twitter. As consumers, we now expect that the information we care about will be pushed to us in real time, and business applications in the cloud are heading in that direction as well.

Cloud computing models are shifting. In the cloud/client architecture, the client is a rich application running on an Internet-connected device, and the server is a set of application services hosted in an increasingly elastically scalable cloud computing platform. The cloud is the control point and system or record and applications can span multiple client devices. The client environment may be a native application or browser-based; the increasing power of the browser is available to many client devices, mobile and desktop alike.

Robust capabilities in many mobile devices, the increased demand on networks, the cost of networks and the need to manage bandwidth use creates incentives, in some cases, to minimize the cloud application computing and storage footprint, and to exploit the intelligence and storage of the client device. However, the increasingly complex demands of mobile users will drive apps to demand increasing amounts of server-side computing and storage capacity.

### *A.Cloud Architecture*

The Cloud Computing architecture comprises of many cloud components, each of them are loosely coupled. We can broadly divide the cloud architecture into two parts: Front End refers to the client part of cloud computing system. It consists of interfaces and applications that are required to access the cloud computing platforms, e.g., Web Browser**.** Secondly, Back End refers to the cloud itself. It consists of all the resources required to provide cloud computing services. It comprises of huge data storage, virtual machines, security mechanism, services, deployment models, servers, etc.

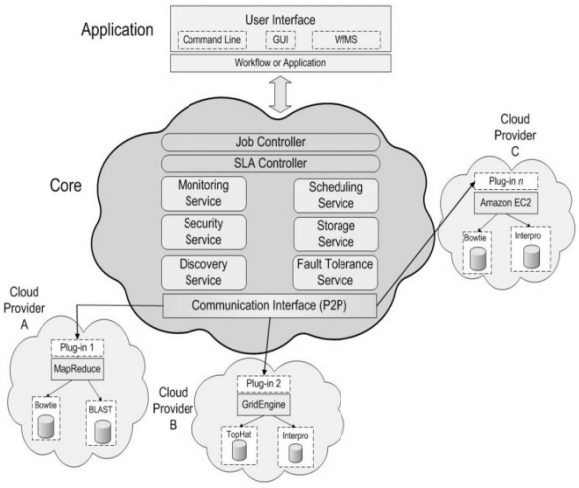


Fig 1.Cloud Architecture

### *1. Resource Allocation*

Resource Allocation is all about integrating cloud provider activities for utilizing and allocating scarce resources within the limit of cloud environment so as to meet the needs of the cloud application. It requires the type and amount of resources needed by each application in order to complete a user job. The order and time of allocation of resources are also an input for an optimal resource allocation.

An important point when allocating resources for incoming requests is how the resources are modeled. There are many levels of abstraction of the services that a cloud can provide for developers, and many parameters that can be optimized during allocation.

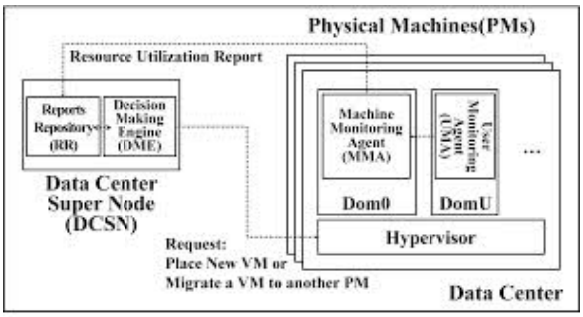


Fig 2. Schematic Representation

The modeling and description of the resources should consider at least these requirements in order for the resource allocation works properly. Cloud resources can be seen as any resource (physical or virtual) that developers may request from the Cloud. For example, developers can have network requirements, such as bandwidth and delay, and computational requirements, such as CPU, memory and storage.

When developing a resource allocation system, one should think about how to describe the resources present in the Cloud. The development of a suitable resource model and description is the first challenge that a resource allocation must address. An resource allocation also faces the challenge of representing the applications requirements, called resource offering and treatment. Also, an automatic and dynamic resource allocation must be aware of the current status of the Cloud resources in real time. Thus, mechanisms for resource discovery and monitoring are an essential part of this system. These two mechanisms are also the inputs for optimization algorithms, since it is necessary to know the resources and their status in order to elect those that fulfill all the requirements.

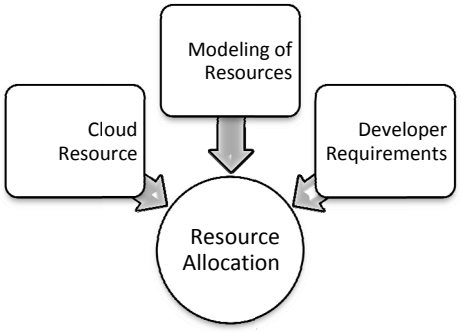


Fig 3. Allocation of Resources

## II. SCHEDULING ALGORITHMS

*1.* ***A green energy-efficient scheduling algorithm using the DVFS technique for cloud datacenters****:*
*Chia-Ming Wu et al, Ruay-Shiung Chang, Hsin-Yu Chan, 2014*

The dynamic voltage and frequency scaling (DVFS) technique can dynamically lower down the supply voltage and work frequency to reduce the energy consumption while the performance can satisfy the requirement of a job. There are two processes in it. First is to provide the feasible combination or scheduling for a job. Second is to provide the appropriate voltage and frequency supply for the servers via the DVFS technique.

This technique can reduce the energy consumption of a server when it is in the idle mode or the light workload. It satisfies the minimum resource requirement of a job and prevent the excess use of resources. The simulation results show that this method can reduce the energy consumption by 5%–25%.

***2. A new multi-objective bi-level programming model for energy and locality aware multi-job scheduling in cloud computing:***
*Xiaoli Wang, Yuping Wang, Yue Cui, 2014*

This programming model is based on Map Reduce to improve energy efficiency of servers. First, the variation of energy consumption with the performance of servers is taken into account. Second, data locality can be adjusted dynamically according to current network state; last but not least, considering that task-scheduling strategies depend directly on data placement policies.

This algorithm is proved much more effective than the Hadoop default scheduler and the Fair Scheduler in improving servers' energy efficiency.

*3.****Cost-efficient task scheduling for executing large programs in the cloud:***
*Sen Su a, Jian Li a, Qingjia Huang a, Xiao Huang a, Kai Shuang a, Jie Wang b,2013*

The cost efficient task-scheduling algorithm using two heuristic strategies . The first strategy dynamically maps tasks to the most cost-efficient VMs based on the concept of Pareto dominance. The second strategy, a complement to the first strategy, reduces the monetary costs of non-critical tasks. This algorithm is evaluated with extensive simulations on both randomly generated large DAGs and real-world applications. The further improvements can be made using new optimization techniques and incorporating penalties for violating consumer-provider contracts.

*4.* ***Priority Based Job Scheduling Techniques In Cloud Computing: A Systematic Review:***
*Swachil Patel, Upendra Bhoi,2013*

Job scheduling in cloud computing mainly focuses to improve the efficient utilization of resource that is bandwidth, memory and reduction in completion time .There are several multi-criteria decision-making (MCDM) and multi-attribute decision-making (MCDM) which are based on mathematical modeling. This PJSC is based on Analytical Hierarchy Process (AHP).

A modified prioritized deadline based scheduling algorithm (MPDSA) is proposed using project management algorithm for efficient job execution with deadline constraint of user's jobs. MPDSA executes jobs with closest deadline time delay in cyclic manner using dynamic time quantum.

There are several issues related to Priority based Job Scheduling Algorithm such as complexity, consistency and finish time.

***5. CLPS-GA: A case library and Pareto solution-based hybrid genetic algorithm for energy aware cloud service scheduling***
*Ying Feng[b], Lin Zhang[a], T.W. Liao,2014.*

On the basis of classic multi-objective genetic algorithm, a case library and Pareto solution based hybrid Genetic Algorithm (CLPS-GA) is proposed to solve the model. The major components of CLPS-GA include a multi-parent crossover operator (MPCO), a two-stage algorithm structure, and a case library. Experimental results have verified the effectiveness of CLPS-GA in terms of convergence, stability, and solution diversity.

***6. Scheduling Scientific Workflows Elastically for Cloud Computing:***
*Cui Lin, Shiyong Lu, 2011*

It proposes the SHEFT algorithm (Scalable-Heterogeneous-Earliest-Finish-Time algorithm) to schedule workflows for a Cloud computing environment. SHEFT is an extension of the HEFT algorithm which is applied for mapping a workflow application to a bounded number of processors.

We schedule these workflows by the HEFT and SHEFT algorithms, and compare workflow makespan by the two algorithms as the size of the workflows increases.

***7. Job scheduling algorithm based on Berger model in cloud environment:***
*Baomin Xu a, Chunyan Zhao b, Enzhao Hua, Bin Hu c,d, et al., 2011*

The Berger model of distributive justice is based on expectation states. It is a series of distribution theories of social wealth. Based on the idea of Berger model, two-fairness constraints of job scheduling are established in cloud computing. The job scheduling is implemented in a cloud Sim platform.

The proposed algorithm in this paper is effective implementation of user tasks, and with better fairness. In future enhancement it deals with build a fuzzy neural network of QoS feature vector of task and parameter vector of resource based on the non-linear mapping relationship between QoS and resource.

***8.Efficient dynamic task scheduling in virtualized data centers with fuzzy prediction:***
*Xiangzhen Kong a,n, ChuangLin a, YixinJiang a, WeiYan a, XiaowenChu et al.,2011*

The general model of the task scheduling in VDC is built by MSQMS-LQ, and the problem is formulates an optimization problem with two objectives: average response time and availability satisfaction percentage. Based on the fuzzy prediction systems, an on-line dynamic task scheduling algorithm named SALAF is proposed. The experimental results show that the proposed algorithm could efficiently improve the total availability of VDCs while maintaining good responsiveness performance.

Considering the cost of consolidation, there exists an optimal consolidation ratio in a VDC that may be related to the hardware resource and the workload, which is an issue in it.

***9. Policy based resource allocation in IaaS cloud:***
*Amit Nathani a, Sanjay Chaudharya, Gaurav Somani et al.,2012*

Haizea uses resource leases as resource allocation abstraction and implements these leases by allocating Virtual Machines (VMs). An approximation algorithm is proposed in which minimize the number of allocated resources which need to be reserved for a batch of tasks. When swapping and preemption both fails to schedule a lease, the proposed algorithm applies the concept of backfilling.

The results show that it maximizes resource utilization and acceptance of leases compared to the existing algorithm of Haizea. Backfilling has a disadvantage of requiring more preemption, which increases overall overhead of the system.

***10.Honey bee behavior inspired load balancing of tasks in cloud computing environments:***
*Dhinesh Babu L.D. a*, P. Venkata Krishnab et al.,2013*

HBB-LB aims to achieve well balanced load across virtual machines for maximizing the throughput. It proposes a load balancing technique for cloud computing environments based on behavior of honey bee foraging strategy. Honey bee behavior inspired load balancing improves the overall

throughput of processing and priority based balancing focuses on reducing the waiting time for the task on a queue of VM.

A task removed from overloaded VM has to find a suitable under loaded.It has two possibilities, either it finds the VM set which is a Positive signal or it may not find the suitable VM i.e a negative signal.

HBB-LB is more efficient with lesser number of task migrations when compared with DLB and HDLB techniques. This algorithm can be extended further by considering the Qos factors in it.

***11. Morpho: A decoupled Map Reduce framework for elastic Cloud computing:***
*Lu Lu, Xuanhua Shi , Hai Jin, Qiuyue Wang, Daxing Yuan, Song Wu, 2014*

To address the problems of frequently loading and running HDFS in virtual clusters and downloading and uploading data between virtual clusters and physical machines, Morpho uniquely proposes a decoupled Map Reduce mechanism that decouples the HDFS from computation in a virtual cluster and loads it onto physical machines permanently. Morpho also achieves high performance by two complementary strategies for data placement and VM placement, which can provide better map and reduce input locality **.**

Evaluation is done using two metrics, job execution time and Cross-rack data transfer amount .Nearly 62% speedup of job execution time and a significant reduction in network traffic is achieved by this method.

***12.CCBKE - Session key negotiation for fast and secure scheduling of scientific applications in cloud computing:***
*Chang Liu et al., Xuyun Zhanga, Chi Yangb, Jinjun Chena,2013*

Cloud Computing Background Key Exchange (CCBKE), a novel authenticated key exchange scheme that aims at efficient security-aware scheduling of scientific applications. This scheme is designed based on the commonly-used Internet Key Exchange (IKE) scheme and randomness-reuse strategy. The data set encryption technique used are block cipher, AES, in Galois Counter Mode (GCM) with 64 k tables, Salsa20/12 and stream cipher.

This scheme improve the efficiency by dramatically reducing time consumption and computation load without sacrificing the level of security. This scheme can be extended in future to improve the efficiency of symmetric-key encryption towards more efficient security-aware scheduling.

***13. A Ranking Chaos Algorithm for dual scheduling of cloud service and computing resource in private cloud:***
*Yuanjun Laili a, Fei Tao a, Lin Zhang a,*, Ying Cheng a, Yongliang Luo a, Bhaba R. Sarker b,2013*

The combination of Service Composition Optimal Selection (SCOS) and Optimal Allocation of Computing Resources (OACR) is known as dual scheduling. For addressing large-scale Cloud Services and Computing Resources (DS-CSCR) problem, a new Ranking Chaos Optimization (RCO) is proposed.

In RCO algorithm ,individual chaos operator was designed, then a new adaptive ranking selection was introduced for control the state of population in iteration. Moreover, dynamic heuristics were also defined and introduced to guide the chaos optimization.

Performances in terms of searching ability, time complexity and stability in solving the DS-CSCR problem is optimal with the use of RCO algorithm but the design of heuristic function for specific problems in the dynamic heuristic operator is complex and hard though.

***14. Analysis and Performance Assessment of CPU Scheduling Algorithms in Cloud using Cloud Sim:***
*Monica Gahlawat, Priyanka Sharma,2013*

This paper analyzes and evaluates the performance of various CPU scheduling in cloud environment using Cloud Sim. Shortest job first and priority scheduling algorithms are beneficial for the real time applications. Because of these algorithms the clients can get precedence over other clients in cloud environment.

Here it deals only with the three algorithms such as FCFS, SJF and priority scheduling. This survey can also be extended for other adaptive and dynamic algorithms suited the virtual environment of cloud.

***15. An Algorithm to Optimize the Traditional Backfill Algorithm Using Priority of Jobs for Task Scheduling Problems in Cloud Computing:***
*Lal ShriVratt Singh, Jawed Ahmed, Asif Khan,2014*

This paper proposes an efficient algorithm 'P-Backfill' which is based on the traditional Backfill algorithm using prioritization of jobs for achieving the optimality of scheduling in cloud systems. The dynamic meta scheduler will deploy the arriving jobs using P-Backfill algorithm to utilize the cloud resources efficiently with less waiting time.

P-Backfill starts the execution of the jobs according to their priority status. It also uses the pipelining mechanism in order to execute multiple jobs at a time. The P-Backfill algorithm is more efficient than other other traditional algorithms such as traditional Backfill, FCFS, SJF, LJF and Round Robin algorithms since it selects the jobs according to their priority levels.

***16. Efficient Optimal Algorithm of Task Scheduling in Cloud Computing Environment:***
*Dr. Amit Agarwal, Saloni Jain,2014*

An optimized algorithm for task scheduling based on genetic simulated annealing algorithm is proposed. Here Qos and response time is achieved by executing the high priority jobs (deadline based jobs) first by estimating job completion time and the priority jobs are spawned from the remaining job with the help of Task Scheduler.

Three scheduling algorithm First come first serve, Round robin scheduling and is generalized priority algorithm. In FCFS resource with the smallest waiting queue time and is selected for the incoming task. Round Robin (RR) algorithm focuses on the fairness. the tasks are initially prioritized according to their size such that one having highest size has highest rank in general prioritized algorithm. The experimental result shows that general prioritized algorithm is more efficient than FCFS and Round Robin algorithm.

***17. Comparative Based Analysis of Scheduling Algorithms for Resource Management in Cloud Computing Environment:***
*C T Lin et al.,2013.*

The resource scheduling in this paper is based on the parameters like cost, performance, resource utilization, time, priority, physical distances, throughput, bandwidth, resource availability.

The scheduling algorithm based on cost factor includes deadline distribution algorithm, backtracking, and improved activity based cost algorithm, compromised time-cost. The algorithm based on the throughput includes Extended Min-Min, modified ant colony optimization. Earliest deadline, FCFS, Round robin is time based

The advantage of this comparative study is that as per the requirements of the consumers and service providers they can select the appropriate class of scheduling algorithms for different types of services required. This study may further be used for optimization of different algorithms for better resource management in cloud computing environment.

***18. Fairness As Justice Evaluator In Scheduling Cloud Resources - A survey:***
*Anuradha1, S. Rajasulochana,2013*

Fairness in scheduling improves the efficiency and provides optimal resource allocation. Fairness constraint proposed by Berger model plays an important role in determining the fair allocation of resources by means of justice evaluation function. An efficient scheduler should provide fair allocation of resources in a way it ensures no task is starving for resources.

The heuristic algorithms are present for both static mapping and dynamic mapping. QoS based heuristic algorithms for static mapping are min-min algorithm, max-min algorithm, opportunistic load balancing, and suffrage heuristics. dynamic scheduling includes immediate mode heuristic algorithms and batch mode heuristic algorithms.

Backfilling algorithms are used to overcome the problem of starvation and waiting time. Backfilling strategy may/may not schedule the jobs based on priority which is both its advantage as well as disadvantage.

***19. A Survey Of Various Qos-Based Task Scheduling Algorithm In Cloud Computing Environment:***
*Ronak Patel, Hiren Mer,2013*

QoS is the collective effort of services performance, which determines the degree of the satisfaction of a user for the services.It is expressed in completion time, latency, execution price, packet loss rate, throughput and reliability.

Task scheduling algorithm based on QoS-driven in cloud computing (TS-QoS) compute the priority of the task according to the special attributes of the tasks, and then sort tasks based on priority.It solves the starvation problem and follow FCFS principle.

***20. Resource management for allocation infrastructure as a Service (IaaS) in cloud computing: A survey***
*Sunilkumar S.Manvi a, GopalKrishnaShyam et al.,*

This paper focuses on some of the important resource management techniques such as resource provisioning, resource allocation, resource mapping and resource adaptation. The common issues associated with IaaS in cloud systems are virtualization and multi-tenancy, resource management, network management, data management, APIs, interoperability.

The performance metrics are used to compare different works under resource management techniques. The metrics considered are reliability, deployment ease, Quality of Service, delay and control overhead.

## III. EXPERIMENTAL RESULTS

From these various scheduling techniques we choose the effective task scheduling algorithm. The algorithm is implemented with the help of simulation tool (Cloud Sim) and the result obtained reduces the total turnaround time and also increase the performance**.** This algorithm deals with the parameters like throughput, makespan and cost.

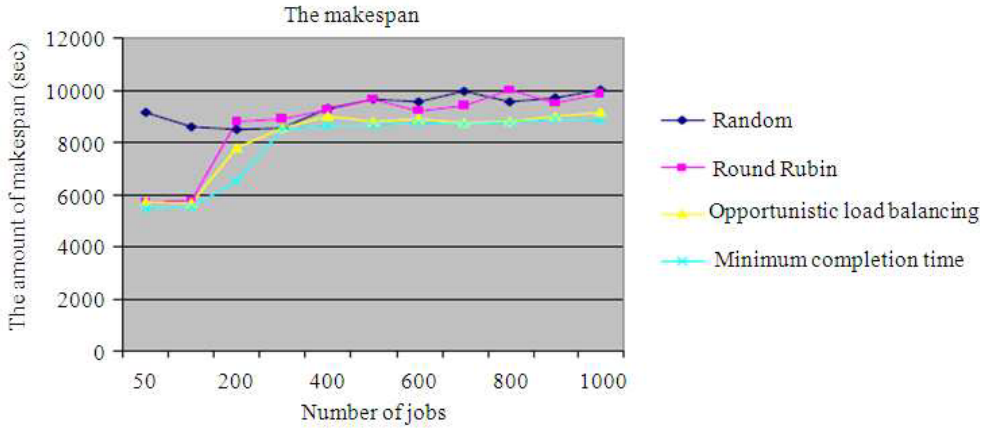


Fig 4. Makespan Vs Jobs

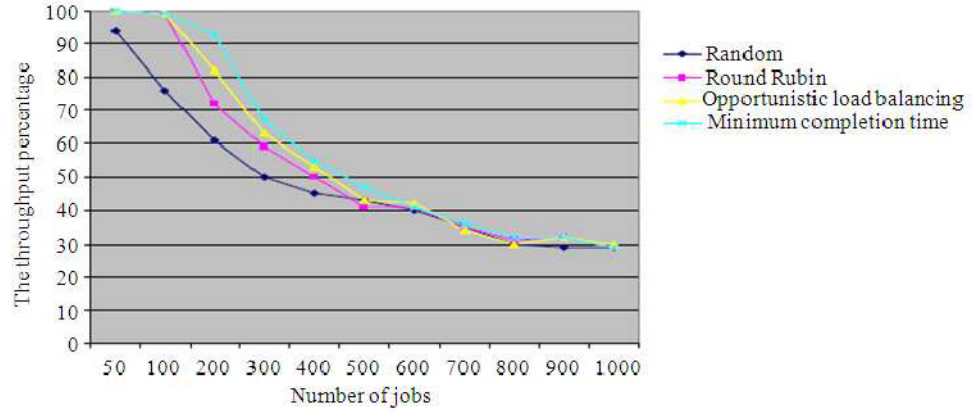


Fig 5. Throughput Vs Jobs

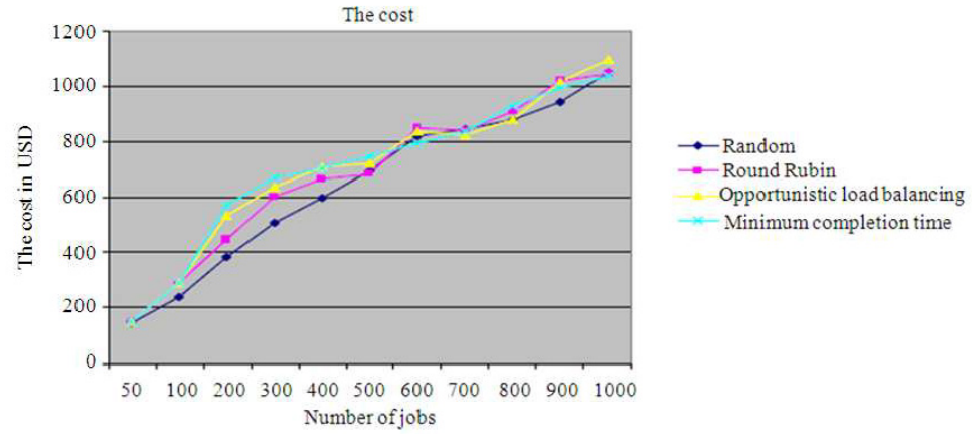


Fig 6. Cost Vs Jobs

Thus the experimental results show that the scheduling algorithms enhance the makespan as well as the throughput of the resources in the cloud environment.

The cloud service providers are those who provide cloud service to the end users. Each CSP promote various scheduling techniques based on their compatibility and availability. The comparison of various CSP and the scheduling algorithm used by their organization is being comprised as below.

TABLE 1 COMPARISON OF CSP'S

| Cloud Service Providers | Open Source | Scheduling Algorithms |
|---|---|---|
| Eucalyptus | Yes | Greedy first fit and Round robin |
| Open Nebula | Yes | Rank matchmaker scheduling, preemption scheduling |
| Rackspace | Yes | round robin, weighted round robin, least connections, weighted least connections |
| Nimbus | Yes | Virtual machine schedulers PBS and SGE |
| Amazon EC2 | No | Xen ,swam, genetic |
| RedHat | Yes | BFS ,DFS |
| lunacloud | Yes | Round robin |

## IV. CONCLUSION

In this paper, we have studied about the problems in scheduling and also about various kinds of scheduling algorithms.

The scheduling algorithm for the datacenter should be chosen based on the requirements of datacenter and the kind of data they store in it. We have analysed the relation between the data that hits the datacenter as well the scheduling algorithm which is required to promote resource allocation in the cloud datacenters. This survey has provided us a crystal clear idea about the wide dimensions of scheduling resources and their functions.

## REFERENCES


[1] Chia-Ming Wu et al, Ruay-Shiung Chang, Hsin-Yu Chan :"A green energy-efficient scheduling algorithm using the DVFS technique for cloud datacenters", Science Direct, Future Generation Computer Systems 37 (2014) 141–147.

[2] Xiaoli Wang, Yuping Wang, Yue Cui: "A new multi-objective bi-level programming model for energy and locality aware multi-job scheduling in cloud computing", Science Direct, Future Generation Computer Systems 36 (2014) 91–101.

[3] Sen Su a, Jian Li a, Qingjia Huang a, Xiao Huang a, Kai Shuang a, Jie Wang b: "Cost-efficient task scheduling for executing large programs in the cloud", Science Direct, Parallel Computing 39 (2013) 177–188.

[4] Swachil Patel, Upendra Bhoi:"Priority Based Job Scheduling Techniques In Cloud Computing", International Journal of Scientific and Technology Research, Volume 2, Issue 11, November 2013, ISSN 2277-8616.

[5] Ying Feng[b], Lin Zhang[a], T.W. Liao:"CLPS-GA: A case library and Pareto solution-based hybrid genetic algorithm for energy aware cloud service scheduling", Science Direct, Applied Soft Computing 19 (2014) 264–279.

[6] Cui Lin, Shiyong Lu:"Scheduling Scientific Workflows Elastically for Cloud Computing", IEEE 4th International Conference on Cloud Computing, 2011.

[7] Baomin Xu a, Chunyan Zhao b, Enzhao Hua, Bin Hu c,d, et al:"Job scheduling algorithm based on Berger model in cloud environment", Science Direct, Advances in Engineering Software 42 (2011) 419–425.

[8] Xiangzhen Kong a,n, ChuangLin a, YixinJiang a, WeiYan a, XiaowenChu et al., :"Efficient dynamic task scheduling in virtualized data centers with fuzzy prediction", Journal of Network and Computer Applications 34 (2011) 1068–1077.

[9] Amit Nathani a, Sanjay Chaudharya, Gaurav Somani et al.:"Policy based resource allocation in IaaS cloud", Science Direct, Future Generation Computer Systems 28 (2012) 94–103

[10] Dhinesh Babu L.D. a*, P. Venkata Krishna b et al.,:"Honey bee behavior inspired load balancing of tasks in cloud computing environments",Science Direct, Applied Soft Computing 13 (2013) 2292–2303.

[11] Lu Lu, Xuanhua Shi , Hai Jin, Qiuyue Wang, Daxing Yuan, Song Wu:"Morpho: A decoupled MapReduce framework for elastic cloud computing", Science Direct, Future Generation Computer Systems 36 (2014) 80–90.

[12] Chang Liu et al., Xuyun Zhanga, Chi Yangb, Jinjun Chena,2013:"CCBKE - Session key negotiation for fast and secure scheduling of scientific applications in cloud computing", Science Direct ,Future Generation Computer Systems 29 (2013) 1300–1308.

[13] Yuanjun Laili a, Fei Tao a, Lin Zhang a,*, Ying Cheng a, Yongliang Luo a, Bhaba R. Sarker b:" A Ranking Chaos Algorithm for dual scheduling of cloud service and computing resource in private cloud", Science Direct ,Computers in Industry 64 (2013) 448–463.

[14] Monica Gahlawat, Priyanka Sharma:"Analysis and Performance Assessment of CPU Scheduling Algorithms in Cloud using Cloud Sim", International Journal of Applied Information Systems (IJAIS) – ISSN : 2249-0868, Volume 5 – No. 9, July 2013.