Introduction underutilized. it delivers all services through the

IntroductionCloud computing is the best everdeveloped technology in the IT industry and a new delivery technique for theservices on pay per use basis. Cloud Computing hasbecome one of the current technologies adopted by both industry and academic. It provides a flexible and efficient way to store and retrieve the data. Themain problem is to schedule the incoming request in a way that it should takeleast response time, efficient resource utilization and at the same timeresources should not be underutilized. It delivers all services through theinternet dynamically when user demands, such as operating system, network, storage, software, hardware and resources. These services are classified intothese types: Software as a Service (SaaS), Platform as a Service (PaaS), Infrastructure as a Service (IaaS), and Everything as a Service (Xaas).

The Cloud computing domain is divided into three categoriessuch as Public, Private and Hybrid cloud. Public Cloud: A cloud is called a “ public cloud” when the services are renderedover a network that is open for public use. Private Cloud: This cloud infrastructure operated solelyfor a single organization, whether managed internally or by a third-party, andhosted either internally or externally. Hybrid Cloud: It is a combination of both public as well asprivate cloud. Hybrid cloudis a composition of two or more clouds (private, community or public) 1 This paper has been orchestrated as follows. Section I gives introduction on cloud computing. Section II specifies the loadbalancing in cloud computing. Section III illustrates the literature study onrelevant load balancing algorithms.

Section IV describes performance analysisof load balancing algorithms.

Section V gives the conclusion. II. LOAD BALANCING IN CLOUD

COMPUTING Load balancing is a new technique that provides highresource time and effective resource utilization by assigning the total loadamong the various cloud nodes, side by side it solves the problem of overutilizationand underutilization of virtual machines. Load balancing resolves the problemof overloading and focuses on maximum throughput, optimizing resourceutilization and minimize response time.

It is the pre requirements formaximizing the cloud performance and utilizing the resources efficiently. The utilizationof clouds has been improved by a resource allocation method which has pre-emptibletask execution. The load balancing is an efficient and critical concept incloud computing and it helps to utilize the resources optimally, therebyminimizing the consumption of resources. Thus load needs to be distributed overthe nodes in cloud-based architecture, so that each resource does the equalamount of work at any point of time that is allocated by a load balancer.

Theload balancer determines the various request allocation to different servers. Theload balancer uses various algorithm to determine the server which has tohandle the request. 1In cloud computing, different load balancing algorithmshave been proposed.

The ultimate aim of all these proposals is to achieve highthroughput and minimizing the response time. III- RELATED WORK Round Robin Virtual Machine Load Balancing AlgorithmIt is a verysimple load balancing algorithm that places the newly coming cloudlets on theavailable virtual machines in a circular manner. The major advantage of thisalgorithm is simplicity and easy implementation. The main drawbacks are priorknowledge of user tasks and

system resources. Further, it does not make use ofcurrent state of the system. 2 Throttled Virtual Machine Load Balancing AlgorithmIn this dynamic approach a user submits request to the Data CenterController (DCC). Then, Data Center Controller asks the VM Load Balancer todetermine the appropriate virtual machine that can handle the given workloadeasily.

Throttled VM Load Balancer keeps a virtual machine list and theirstatus (available/busy). If a suitable VM is found on memory space, cores oravailability basis, then throttled VM Load Balancer accept the cloudlet requestand allot the cloudlet request over that virtual machine. Otherwise, client hasto wait in the queue until a suitable VM becomes available.

Among all, it isbest approach for load balancing, since it maintains the present state of allVMs in data center. But the major drawback is that it works properly only ifall VMs in a data center have same hardware configuration. 2 ESCE Virtual Machine load Balancing Algorithm ESCE standsfor Equally Spread Current Execution. It is also called as Active VM LoadBalancing algorithm. This algorithm is based on spread spectrum technique. Itequally distributes the workload among VMs in data center.

A job queue keepsall the cloudlet requests that need the VM for their execution. ESCE VM LoadBalancer (VMLB) also maintains a list of virtual machines. The VM Load Balancercontinuously checks the job queue and VM list. If a VM is found free, thencloudlet request will be allotted over that VM. At the same time, VMLB inspectthe overloaded VMs. If any virtual machine is found overloaded, then VMLB movesome load to an idle or an under loaded virtual machine, so as to reduce someload of overloaded VM. The main

drawback is high computational overhead. 2 An enhanced priority based HTV load balancing algorithm This algorithmperforms an effective and reliable resource allocation of the tasks on theservers in cloud computing environment.

This algorithm considers the threeparameters such as load on the server, current performance of server and timelimit of the tasks. This algorithm computes the load and performance factor ofeach virtual machine and then allocates the incoming task to various virtualmachines according to their time limit and stand-by time to increase thethroughput and performance. 3 Virtual Machineand Physical Machine Categorization Algorithm In general, load balancing approach that amplify thephysical resource utilization and curtail the energy consumption. To calculatethe performance of the approach it is compared with the existing load balancingapproach and judged against the number of migration and energy consumption. Experiment results say that this approach gives better result while it iscompared with the existing load balancing methods. 4 Composite algorithmLoad Balancing is required to properly manage the resources of theservice contributor. Load balancing is a technique to distribute the workloadamong many virtual machines in a Server overthe network to achieve optimal resource consumption, decrease in dataprocessing time, decrease in average response time, and avoid overload. Theobjective of effective, efficient and enhanced composite scheduling algorithmis used to maintain the load and provides efficient resource allocationtechniques.

This Composite approach is applied for load balancing using EquallySpread Current Execution (ESCE) and Throttled algorithms. 5 Distributed dynamic

priority based algorithm It is used forbalancing the load on instances effectively, improves the system consistency, minimizes response time and increases the throughput. Allocating the resourceson virtual machines based on priority achieves the better response time andprocessing time. Load balancing ensures all instances in a node in the networksto do the equal amount of work at any instant of time. Priority based resourceprovision to improve the utilization of resources and reducing response time ofcloud services. 6Burstness-aware loadbalancing algorithmThis algorithm can adapt to the variation inthe request rate by adopting two load balancing algorithms: RR in burst andRandom in non-burst state. Fuzzy logic is used in order to assign the receivedrequest to a balanced VM.

This algorithm has been evaluated and compared withother algorithms using Cloud Analyst simulator. Experimental results show thatthe algorithm improves the average response time and average processing time incomparison with other algorithms. 7 Honey Behavior LoadBalancing TechniqueIn this Technique the high priority tasks are removed fromoverloaded virtual machine and they are allocated to under loaded virtualmachine by considering least numbers of same priorities to those tasks, costeffective virtual machine, and least expected completion time which alsobalances the loads of dependent tasks in pre-emptive manner. The least expectedcompletion time, cost and priority at submission time of that task helps toproduce minimum completion time, reduces waiting time of the tasks andeventually achieves better resource utilization.

8 Weighted based optimized load balancing This approachis for distributing of incoming jobs uniformly among the servers or virtualmachines. The

performance is analyzed using Cloud simulator and compared withexisting Round Robin and EIPR algorithms. Simulation results have demonstratedthat this algorithm has distributed the load uniformly among virtual machines. 9 New enhanced load balancingalgorithmFirst an adaptive strategyhas been devised for load balancing according to the quality of the solutionsfounded by Genetic. Secondly, the enhanced load balancing strategy is combinedwith the setting of other parameters like fitness and the selection of theinitial resource pool provides the significant impact on the performance of thealgorithm. In this paper the new enhanced load balancing algorithm gives thebetter result than the existing genetic algorithm. 10 Tow –level global loadbalancing frameworkA framework for global server loadbalancing of the Web sites in a cloud with two-level load balancing model. Thisframework is intended for adapting an open-source load-balancing system and theframework allows the network service provider to deploy a load balancer indifferent data centers dynamically while the customers need more load balancersfor increasing the availability.

11 Dynamic load balancing algorithms Load balancing algorithms play important role in equalizingload among data centers and in efficient use of computing resources. In thispaper, performance of a dynamic load balancing algorithm has been evaluated bydividing data-centers in different zones. It has been shown that the proposedalgorithm improves the computing efficiency of data-centers and minimizes theresponse time of user’s applications. 12 IV- PERFORMANCE ANALYSISOF LOAD BALANCING ALGORITHMSThe forthcoming part discusses various load balancingalgorithms and shows the results comparatively. Min- min Load

Balancing Algorithm Min-Min is a simple and fast algorithm capable ofproviding improved performance. Min-Min is use the ideal tasks at first whichresults in best and improve the overall makespan.

Assigning small task first isits drawback. Thus, smaller tasks will get executed first, while the largertasks keep on in the waiting stage, which will finally result in poor machineuse. Min-Min exhibits minimum completion time for jobs which are unassigned andlater allocating the jobs with minimum completion time (hence min-min) to anode that is capable of handling it. 13 Max-min Load Balancing Algorithm This algorithm first for all the available tasks aresubmitted to the system and minimum completion time for all of them arecalculated, then among these tasks the one which is having the completion time, the maximum is chosen and that is allocated to the corresponding machine. If ina task set only a single long task is presented then, Max-Min algorithm runsshort tasks concurrently along with the long task. Max-Min is almost identicalto Min-Min, except it selects the task having the maximum completion time andallocates to the corresponding machine.

The algorithm suffers from starvationwhere the tasks having the maximum completion time will get executed firstwhile leaving behind the tasks having the minimum completion time. 13 RASA Algorithm RASA is Resource Aware Scheduling Algorithm it combinesMin-Min and Max-Min algorithm alternatively in order to achieve betterperformance. In RASA resource completion time is calculated and then Min-Minand Max –Min algorithms are applied alternatively. 13 Minimum Makespan algorithm This algorithm first checks the no-more task left thenminimum makespan taks first selected and it compare with the task and migrated. The two tasks are produces the same

makespan time it choose the node withhigher computational resources. 13 PA-LBIMM Algorithm PA-LBIMM priority aware load balancing improved min minalgorithm separate the tasks into G1 and G2groups. The tasks submitted by VIPuser’s or high priority users are considered as group G1 and tasks submitted bylow priority users are considered as group G2. Tasks are scheduled to theresources on the priority basis.

Firstly, for all the tasks in G1, each task isassigned to the VIP category resource by using Min-Min. Then each task in G2group is assigned to all the resources by using Min-Min. Now, load balancingfunction of LBIMM algorithm is executed to load balance all the resources.

14 RPA-LBIMM Algorithm This algorithm RPA-LBIMM recovery priority aware load balancing improved minmin algorithm is also similar to the priority aware load balancing improved minmin algorithm. Here use the recovery policy which helps the cloud scheduler toreschedule the tasks if a resource fails at the time of execution to achievethe minimum makespan. According to this policy, First of all, scheduler looksfor the failed resource. All the tasks that were scheduled by PA-LBIMM toexecute on will be considered as a task set. 14 TABLE 1. List of Load balancing Algorithms in Cloud computing S. No.

Load balancing alg. Parameters Used 1 Min Min Makespan, Resource utilization 2 MaxMin Maekspan, Resource utilization 3 RASA Maekspan , Resource utilization 4 Minimum Makespan Maekspan , Resource utilization 5 PA-LBIMM Maekspan 6 RPA-LBIMM Maekspan Table 1 lists out the various loadbalancing algorithms and the parameters used in the algorithms. All

thesealgorithms use the parameters completion time and resources utilization incloud environment. These algorithms distribute the load based on the number ofuser request and number of available resource in cloud and also considering theproper utilization of resource. TABLE 2. Task Completion Time makespoan (sec.) of the givenalgorithms S. No.

Load balancing alg. Completion time Makespan (sec.) 1 Min Min 30. 35 2 MaxMin 20. 03 3 RASA 11. 45 4 Minimum Makespan 10. 8 5 PA-LBIMM 38.

37 6 RPA-LBIMM 35 Table 2 represents thecompletion time of the task based on the number of nodes in the simulationarea. The makespan of the different algorithms are given. TABLE 3. Resource utilization of the algorithms S. No. Load balancing alg. Resource utilization (%) 1 Min Min 43.

09 2 MaxMin 79. 32 3 RASA 88. 72 4 Minimum Makespan 91. 01 5 PA-LBIMM — 6 RPA-LBIMM — Table 3 represents the resourceutilization of various load balancing algorithms. The average resourceutilization of the different algorithms is given. V – conclusionThe studydeals with various load balancing algorithms. The existing algorithms arestatic, dynamic, composite, and prioritized.

The ultimate purpose of thosealgorithms is to reduce the response time and maximize the resourceutilization. The results of the existingalgorithms are confined to give better result. Still there is plenty of spaceto improve the results to extract best service from cloud service providers. This study also shows the comparative results of the existing load balancingalgorithms based on the parameters such as makespan and resourceutilization.

References1 Reena Panwar Bhawna Mallick, “ A Comparative Study of

Load BalancingAlgorithms in Cloud Computing", International Journal ofComputer Applications (0975 – 8887) Volume 117 – No. 24, May 2015. 2 Mamta Khanchi, Sanjay Tyagi, " An EfficientAlgorithm for Load Balancing in Cloud Computing" International Journal of Engineering Sciences & Research Technology, June, 2016. 3 DivyaGarg, Urvashi Saxena," Dynamic QueueBased Enhanced HTV Dynamic Load Balancing Algorithm in Cloud Computing", International Journal of Innovative Research in Science, Engineering andTechnology, Vol.

5, Issue 1, January 2016. 4 Sachin Soni, Praveen Yadav, " A LoadBalancing Approach to Minimize the Resource Wastage in Cloud Computing", InternationalAdvanced Research Journal in Science, Engineering and Technology, Vol. 3, Issue 3, March 2016. 5 Navtej Singh Ghumman, Rajesh Sachdeva," anefficient approach for load balancing in cloud computing using compositetechniques", International Journal of Research in Engineering and AppliedSciences, volume 6, issue 2 February, 2016.

6 G. Suryadevi, D. Vijayakumar, R. SabariMuthuKumar, Dr. K . G.

Srinivasagan, " An Efficient Distributed Dynamic Load BalancingAlgorithm for Private Cloud Environment", International Journal of InnovativeResearch in Science, Engineering and Technology, Volume 3, Special Issue 3, March 2014. 7 SallyF. Issawi, Alaa Al Halees, Alaa AlHalees, " An Efficient Adaptive Load Balancing Algorithm for Cloud ComputingUnder Bursty Workloads", Engineering, Technology & Applied Science Research Vol. 5, No. 3, 2015. 8 Khushbu Zalavadiya, Dinesh Vaghela , " Honey Bee

Behavior Load Balancing ofTasks in Cloud Computing”, International Journal of Computer Applications (0975 – 8887), Volume 139 –No. 1, April 2016.

9 B.

Nithya Nandhalakshmi , Mahalingam,” Efficient Load Balancing in CloudComputing Using Weighted Throttled Algorithm”, International Journal ofInnovative Research in Computer and Communication Engineering , Vol. 3, Issue 6, June 2015. 10 Er. Pooja Er.

Vivek Thapar,” An Enhanced Virtual Machine Load Balancing Algorithmfor Cloud Environment”, , InternationalJournal of Emerging Research in Management &Technology, ISSN: 2278-9359 (Volume-5, Issue-5), May 2016.

11 Po-Huei Liang1 and Jiann-Min Yang. “ Evaluationof two-level Global Load Balancing Framework in cloud environment” InternationalJournal of Computer Science & Information Technology (IJCSIT) Vol 7, No 2, April 2015.

12 Navtej Singh Ghumman, Rajesh Sachdeva,” anefficient approach for load balancing in cloud computing using compositetechniques”, International Journal of Research in Engineering and AppliedSciences, volume 6, issue 2 February, 2016. 13 Er.

Rajeev Mangla, Er. Harpreet Singh,” Recovery and user priority based loadbalancing in cloud computing”, International Journal of Engineering and Scienceand Research, February 2015. 14 Harish Chandra, Pradeep Semwal, SandeepChopra,” load balancing in cloud computing usinga novel minimum makespanalgorithm”, International Journal ofAdvanced Research in Computer Engineering & Technology , Volume 5, Issue 4, April 2016. Abstract: Cloud computing is a new technology now a days for

providing differentkind of services to the end users. It focuses mainly the dynamic services usingvery large scalable and virtualized resources over the Internet. The allocation of the groupof resources may start a problem of availability of these resources and distributingthe workload of all the VMs among themselves called as load balancing. Load balancing is a mainchallenge in cloud environment.

It helps to distribute the dynamic workloadacross multiple nodes to ensure that no single node is overloaded. It helps inproper utilization of resources . It also improves the performance of thesystem. Load balancing is the process of finding overloaded nodes and thentransferring the extra load to other nodes.

Due to novelty of cloudcomputing field, there is no many standard load balancing algorithm used incloud environment. Hence efficient utilization of resources must be importantand for that load balancing plays a major role to get maximum benefit from theresources. Having understood the vital role of load balancing, this papercovers various load balancing algorithms related to cloud computing and shows acomparative study on load balancing algorithms. Keywords: Cloud Computing, Load Balancing, Virtual Machine D. Suresh Kumar, Research ScholarSchool of Computer Science, Engineeringand Applications Bharathidasan University Tiruchirappalli, Tamil Nadu,