IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 5, Issue 9, September 2016

A Hybrid Approach for Scheduling in Cloud Computing

Navdeep Kaur[1], Mandeep Kaur[2]

Student, Department of Computer Science & Engineering, Rayat& Bahra, Mohali, India [1]

Assistant Professor, Department of Computer Science & Engineering, Rayat& Bahra, Mohali, India[2]

Abstract: Cloud computing is emerging as a potential player in field of Distributed computing. Therefore enhancement in scheduling algorithm is the major research area which is being under focused by various researchers. Efficient scheduling algorithm tends to increase the potential of any Cloud Service Provider (CSP) infrastructure. In order to increase the performance of scheduling capabilities, this work has been focused on improving Map e-K algorithm with the combination of worst fit scheduling algorithm. The Map e-K loop provides monitoring, analysing, planning & execution phases & under the planning phase we have used latency matrix that will create plan to allocate job request to VMs. Also, in this work thebinary search technique has been used to find appropriate VMs for incoming job request from the userThis hybrid version of Map e-K is successfully tested& shown under observed results.

Keywords: Cloud, Map e-K, Binary Search, latency Matrix.

I. INTRODUCTION

Cloud Computing is a model for permitting omnipresent, suitable, on-demand service access to a common group of configurable Computing resources (e.g., networks, servers, storage and applications) that can be quickly provided and released with least management struggle or Cloud provider dealings [38]. The vision of key characteristics of Cloud environment is shown in Figure 1.1.

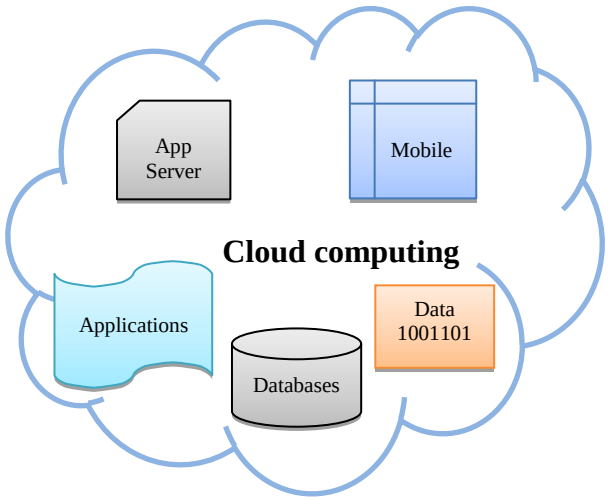

Figure 1: Vision of Key Characteristics of Cloud Environment.

A. Elements of Cloud Computing

There are seven main elements of Cloud Computing, classified on economic, architectural and strategic elements [36]:

I. Utility Pricing:Cloud computing is well-defined by its usage based billing model. Cloud provider contributes to this as Cloud consumers of the platform use computing and storage facilities on requirement and pay based on the consumption, using an Operating Expenses (OPEX) budget and Capital Expenditures (CAPEX).

II. Elastic Resource Capacity: Cloud computing scales computing and storage resources up and down, consumers can add or remove resources immediately and make payment forthe resources a Cloud consumer is consuming.

III. Virtualized Reosurces: Without virtualization Cloud Computing is impossible, not for mysterious technical causes, but for one recognizable business requirement: the requirement of multi-tenancy. In order to take advantage from saving of scale[38].

IV. Managed Operations:Maintenance of software's, data replication and countless other tasks is mandatory to handle mission-critical business applications on a daily basis that becomes the accountability of a third party, according to SLAs[38].

B. Impact of scheduling in cloud computing

Scheduling in cloud computing which incorporates central scheduling, intelligent scheduling and agent negotiated scheduling got many similarities and also differences. First, the biggest difference between cloud computing environment and traditional computing environment is the target of scheduling. ,in traditional computing environment, it mainly schedules processes or tasks so the granularity is small and the transferred data is small; whereas in cloud computing environment the scheduled target is VM resources so the granularity is large and the transferred data is large as well. Second, in cloud computing environment compared with the deployment time of VMs, the time of scheduling algorithm can almost be neglected. This work sees to the equal distribution of

IJARCCE
ISSN (Online) 2278-1021
ISSN (Print) 2319 5940
International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 5, Issue 9, September 2016

hardware resources of VMs in cloud computing environment so that the VM can improve its running efficiency while meeting the QoS needs of subscribers. Scheduling is the one of the most prominent activities that executes in the cloud computing environment. To increase the efficiency of the work load of cloud computing, scheduling is one of the tasks performed to get maximum profit.

This work presents an Hybrid approach for optimizing the utilization of Cloud computing resources as well as the runtime of an application . Within this context, the three main objectives are the allocation of as little computing resources as possible, the minimization of runtime & minimization of SLA violations . The Hybrid approach tackling this challenge on a per-job-basis relies onwell-known concepts from autonomic computing [8], particularly on the MAPE-K loop containing a monitor, analyzer, planner, executor, and knowledge component.

The remainder of this paper is organized as follows: The next section discussing related work followed by a section introducing proposed methodology & results. Finally a conclusion of the work including future work completes this paper.

II. RELATED WORK

This section involves the work done by the various researchers in the field of Scheduling algorithms in cloud computing. In[5], the authors discussed about information retrieval as the most fundamental requirement for any kind of computing application which requires search operation to be performed from databases. They have shown that binary search is more efficient searching technique than linear search. In[1], Best fit & Worst fit dynamic virtual machine scheduling techniques are used for reducing the response time along with efficient & balanced resource utilization. The results shows that Worst fit technique is better than Best fir technique.

In[3], authors have compared various algorithms in cress over & mutation by comparing & analysing mutation & local search algorithms based on particle swarm. In[4], the authors present the self configuring a adaptive framework optimizing resource utilization for scientific applications on top of cloud technologies. The proposed framework self configures the layers by evaluating monitored resources, analysing their state & generating an execution plan on per job basis.In[10],authors discussed existing use cases from grid & cloud computing systems to identify the level of SLA realization in state-of-art system emerging challenges. They have discussed some still open challenges such as scalability, dynamic environmental change, heterogeneity, SLA management automation, and multiple QOS parametrs.

In [7], authors have shown the effective role of latency on broker scheduling algorithm. According to the authors in most of the existing techniques like FIFO which allot the existing requests as per nature of the request but the latency of the user- VM's location is ignored. In[5], the QoS encompasses different parameters like smart job allocation strategy, efficient load balancing, response time optimization, reduction in wastage of bandwidth accountability of overall system. The ETC & MTC metrices help to map the jobs to the appropriate VMs which will reduce the overall response time & waiting time of jobs. In[9], authors presented the priority based algorithm for scheduling virtual machines on physical host in cloud computing environment. In this strategy the requests are ranked according to the profits they can bring.

III.PROPOSED WORK

A novel approach is proposed for VM placement which effectively solves the problems of minimizing the response time, load balancing, and balance resource utilization and SLA Violation in cloud data center.

In this work three dimensional problems are solved:

- Improving response time
- SLA Violations
- Improving Utilization

In this algorithm the hosts are classified according to their resource availability. In general there are three types of performance parameters of any system i.e. CPU, B/W and Memory. The allocation of the VM is typically done by Mape-K loop in which the sensor to sense the VM status and create a plan to allocate job requests to VMs. However after allocation a large number of resources remain under utilized. The algorithm shows the total remaining underutilized resources is accumulated and is allocated to next requests. Let us assume that when request for VM arrived to scheduler place this VM on the satisfying host in the list in which all host have remaining resource capacity in order. In other words if a VM having CPU requirement greater than or equal to Memory requirement and Memory requirement greater than or equal to Bandwidth is placed on the satisfying host in the list, in which all host have remaining resource capacity of CPU greater than or equal to Memory and Memory greater than or equal to Bandwidth. The detail is Shown in Algorithm. To improve the arranging and retrieving system we are putting worst fit and binary search to make retrieving fast.

In the Worst Fit algorithm hosts in each list in datacenter are sorted in descending order according to remaining capacity of resources. When request of a new VM or already running VM for VM placement arrives at the cloud datacenter, VM scheduler find the appropriate list and apply the binary search on the selected list to find host that is the worst fit in remaining resource capacity than the VM requirement capacity in all dimensions.

In Algorithm:

- Firstly select a set of VM's according to the resource demand of the user.
- Max parameter is used to estimate the total no. of VM's in the datacenter.
- Create a Latency Matrix on the Planning Phase of the Mape-K Loop.

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 5, Issue 9, September 2016

- Utility optimization means check the remaining memory, remaining CPU and remaining bandwidth
- If the request = remaining CPU then total remaining power of CPU is greater than to memory and bandwidth.
- If the request = remaining memory then total remaining power of memory is greater than to CPU and bandwidth.
- If the request = remaining bandwidth then total remaining power of bandwidth is greater than CPU and memory.
- After these steps allocate the user's request to VM's according to the size of the request & remaining capacity of the virtual machine.
- Manage the database (knowledge) with the help of the worst fit scheduling & binary search technique, i.e improve the results of the Map e-K.

The worst fit algorithm is a means by which Planner searches for free-Virtual Machine in Datacenter in which it can place the desired Job request. The algorithm selects the largest possible free VM to which job can be allocated. In this system all the monitoring results obtained from virtual machines and will be evaluated for analysis. All the requests are sorted out in descending order according to their size.

A worst fit algorithm is used to planning the further request as the performance of the Virtual Machines are already being put into system. Based on the historical performance all the new requests will be taken into account and submitted for execution to the virtual machines.

IV. RESULTS

Series of experiments were performed; in each experiment number of cloudlets, number of data centers and number of virtual machines in data center was varied. As a result these experiments were to analyze the utilization, response time and SLA Violations.

A. SLA Violations

Obtained the results of the SLA Violations using CloudSim& compare with the existingMape-K algorithm. SLA is a service level agreement between user and CSP(cloud service provider).

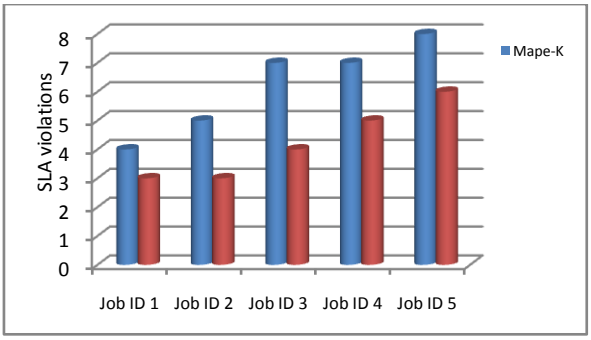

Figure.2: SLA Violations comparison result.

B. Response Time

The Response time is improve using worst fit technique is applied on proposed Mape-K Algorithm that is efficient as compare to existing Mape-k in which searching technique is used to compare the results of the both algorithms.

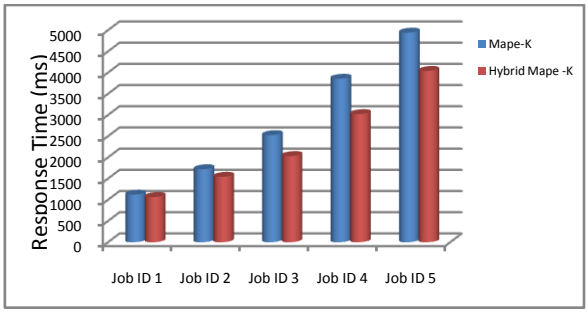

Figure.3: Response Time

C. Utilization

Utilization of Virtual Machines using scheduling & searching techniques. The maximum no. of resources or VMs to be used means maximum utilization, all the VMs in the datacenter are doing the jobs without any delay & check the utilization of the virtual machines & compare the results of the both algorithm's i.e better than existing Map e-K.

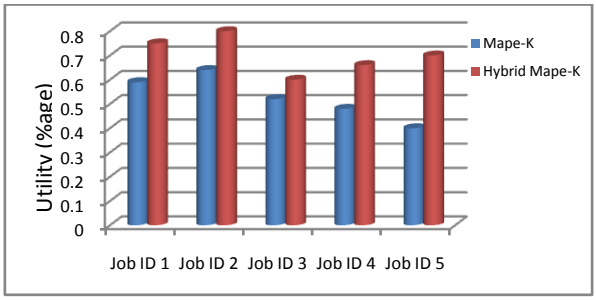

Figure.4: Utilization

V. CONCLUSIONS

In this work, we have discussed about the job allocation to the different VMs inside a Cloud Data Center with the help of the Mape- k loop, which is a classical scheduling policy . It helps to map the jobs to the appropriate VMs, which will reduce the overall response time and waiting time of the jobs and with the introduction of worst case algorithm, we observe a further improvement in performance. A system has been developed in such a way that will help the module to identify intelligently that how many jobs may be served by a single VM at a certain time stamp. This will improve the overall make span of the Cloud Service Provider.

It describes the importance of efficient VM scheduling technique to provide the solution for the problems of VM placement, response time, balance resource utilization and load balancing. The proposed Best Fit and Worst Fit techniques are beneficial for cloud provider as well as cloud user for cost saving . The experimental results show that the Worst Fit technique is better than the Best Fit

technique for different capacities hosts. But both techniques provide better improvements with other traditional techniques. As proposed algorithms are evaluated using the CloudSim simulator toolkit. So run time challenges can be resolved by implementing the proposed algorithms in real world cloud environment. In the proposed scheduling algorithm VM migration strategy is not considered for under utilize host or over utilize host. The evaluation for the response time, balance utilization of resources of a host and load distribution on all the hosts is only based on the VM scheduling techniques. So using the migration strategy for under utilize host and over utilize host, which can provides the measurable improvements for load balancing and can also provides the server consolidation.

ACKNOWLEDGMENT

First of all I want to thank the light of god who guided me throughout the way. Then apart from the efforts from me, I would like to show my greatest appreciation to my guide **Mrs. Mandeep Kaur,** (Assistant professor) for her great efforts of supervising and leading me to accomplish this fine work.

REFERENCES

[1] Rathor, Vijaypal S., R. K. Pateriya, and Rajeev K. Gupta. "An Efficient Virtual Machine Scheduling Technique in Cloud Computing Environment."International Journal of Modern Education and Computer Science 7, no. 3 (2015): 39.

[2] Parmar, Vimal P., and C. K. Kumbharana. "Comparing Linear Search and Binary Search Algorithms to Search an Element from a Linear List Implemented through Static Array, Dynamic Array and Linked List."International Journal of Computer Applications 121, no. 3 (2015).

[3] Arti, Malti Rani, LohitKapoor, " Scheduling strategies in cloud computing" (IJCST) Volume3 Issue 3,May-June 2015.

[4] Koehler, Martin. "An adaptive framework for utility-based optimization of scientific applications in the cloud." Journal of Cloud Computing 3, no. 1 (2014): 1.

[5] Banerjee, Sourav, MainakAdhikari, and UtpalBiswas. "Development of a smart job allocation model for a Cloud Service Provider." In Business and Information Management (ICBIM), 2014 2nd International Conference on, pp. 114-119. IEEE, 2014.

[6] Lin, Weiwei, Chen Liang, James Z. Wang, and RajkumarBuyya. "Bandwidth-aware divisible task scheduling for cloud computing." Software: Practice and Experience 44, no. 2 (2014): 163-174.

[7] Rishu, "Resourceful Utilization of Latency in Virtual Machine Allocation in Cloud Computing," (IJARET) ISSN:2320-6802, Issue IX sep 2014.

[8] Supreeth, S., and ShobhaBiradar. "Scheduling virtual machines for load balancing in cloud computing platform." International Journal of Science and Research (IJSR), India Online ISSN (2013): 2319-7064.

[9] Xiao, Jing, and Zhiyuan Wang. "A Priority Based Scheduling Strategy for Virtual Machine Allocations in Cloud Computing Environment." In Cloud and Service Computing (CSC), 2012 International Conference on, pp. 50-55. IEEE, 2012.

[10] Wu, Linlin, and RajkumarBuyya. "Service Level Agreement (SLA) in utility computing systems." IGI Global (2012).

[11] Bessai, Kahina, Samir Youcef, AmmarOulamara, Claude Godart, and SelminNurcan. "Bi-criteria workflow tasks allocation and scheduling in Cloud computing environments." In Cloud Computing (CLOUD), 2012 IEEE 5th International Conference on, pp. 638-645.IEEE, 2012.

[12] Wu, Hu, Zhuo Tang, and Renfa Li. "A priority constrained scheduling strategy of multiple workflows for cloud computing." In Advanced Communication Technology (ICACT), 2012 14th International Conference on, pp. 1086-1089. IEEE, 2012.

BIOGRAPHIES

Navdeep Kaur is pursuing Masters of Technology in Computer Science Engineering from Rayat& Bahra College Mohali, Punjab (India) She received the degree of Bachelor of Technology in Computer Science Engineering from Amritsar College of Engineering & Technology Amritsar, Punjab (India). Her area of interest is cloud computing.

Mandeep Kaur received her M.Tech degree in computer engineering. She has more than 5 years of teaching experience. Presently she is working as Assistant Professor in Rayat& Bahra College of Engineering & Biotechnology, Mohali, Punjab.