Dynamism in Cloud Resource Allocation : A Survey

S. Lakshmanan*, Dr. D. Ravindran

Department of Computer Science, St. Joseph's College, Trichirapalli-2, Tamil Nadu , India

ABSTRACT

Cloud Computing is a one of the distributed technology. This technology only paid and use, it provides the various service models on the internet. Such as Iaas, Paas, Saas, it did with scalability, availability, flexibility of resource through virtualization. Virtualization is one of the important technologies of this environment. This model focuses the various resource allocations on a virtual machine with dynamically based on user requirement. It is emulated the various issues and challenges to the virtual machine for resource allocation, that is loading balancing, green computing, optimization, scheduling, on-demand service, migration, etc. In this survey, focus the various technologies and algorithm used in existing dynamically allocation of the different virtual machine, advantage, simulated mechanisms, parameters, and key issues in this environment.

Keywords: Cloud Computing, Resource Allocations, Virtual Machine, Overload Avoidance, Green Computing

I. INTRODUCTION

Cloud computing is a storing and retrieving and maintain the user data using over the internet instead of your CPU. This environment allocation the resource computing and would like to save the energy amount of the data center. Where the resource can be dynamically moving the virtual machine. The resource allocation one of the challenges to cloud computing. For scalability increase or decrease the resource usage on demand as processing needs require. The resource application can increase the resource, when there is high user demand and scale down a resource when there is low demand [1]. As the key characteristic of resource management, service of scheduling, Enhancing the overall system performances. It is performed the Iaas,Paas,Saas service models, with the public cloud. Is based on the resource allocation can be achieved the two levels. **Level 1**: The application such that file host can be uploaded to the cloud. It is assigning the request to the physical computer. **Level 2:** The application receives the various incoming requests, should be assigned to the specific application [2].The resource allocation consists of four main functions, such as scheduling, code to transfer, data transmission and monitoring [3].
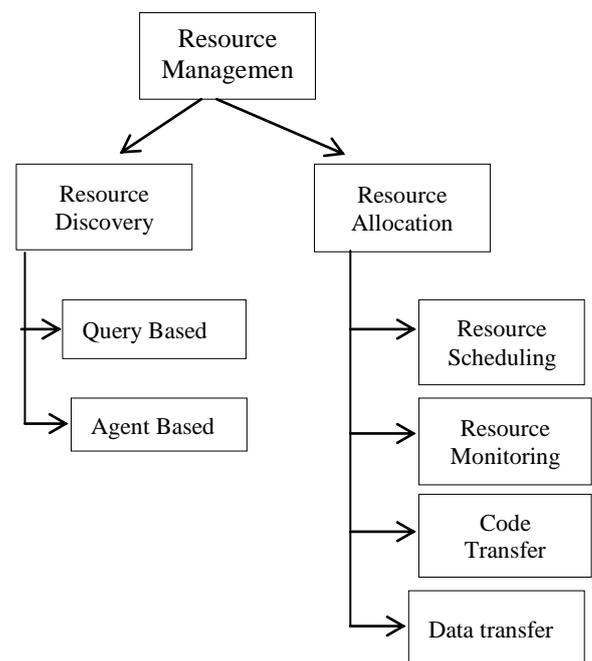

Figure 1. Resource Allocation Stage

The above technique overcomes the fault tolerance, scalability of the resource under the RM organizations.

Resource allocation strategies refer to cloud service providers and allocation, resource with limited of an application for the cloud environment as per need, the strategies of multiple task allocation, exact time, policy, virtual machine, Gossip protocol, utility function, hardware resources dependence, auction ,application, SLA [4].

II. METHODS AND MATERIAL

Related Works

Chee Shin Yeo et al., [5] proposed the penalties to enhance the utility based on SLA of the cluster. This model contained four QoS parameters such that doodling type, deadline, budget, penalty rates, it is describing the LibraSLA. LibraSLA related to Admission control and Resource allocation algorithm. This Simulated to compare better than Libra followed by more job deadlines and minimizing the penalties. This supports the SLAs in cluster level resource allocation for service-oriented grid computing

David L, et al., [6] propose the Heuristics Robust Resource allocation in parallel with distributed via. This technology introduced two phases greedy heuristic (BASIC) and iterative heuristic (Steady State Genetic Algorithm). Both algorithms minimize the routing time. It is based on the min-min algorithm, whereas applied to steady state GA. This algorithm solved the machine fails, increase the system load, estimated the base of resource with data transmission

Daniel warneke et al., [7] propose the Nephele, the Nephele is the first framework of data processing the center. It is used to dynamically allocate and reallocate resources, scheduling tasks, execution of the Iaas cloud. The task can be assigned different type virtual machine, which is initialized and terminated during the job execution with automatic. These virtual machines based on the framework compare the data processing better Hadoop. It can be extended MapReduce. These help to reduce the processing cost and improve the overall resource utilization for consequentially.

Sharma et al., [8] propose the bidding model. This model consists of three stages, stage1: Grid in the client-side algorithm, stage2: Resource Allocation, stage 3: Resource provider. Above the algorithm resolved the

resource management in grids. It is referred to mapping the different type resource (bandwidth, processing power).After the resource allocating resource, it can check itself another resource requesting or not with the queue order manner. This algorithm main goal decreases the number of tasks and increases the existing approach, resource utilization.

Alexander Ser et al., [9] proposed the Adaptive resource allocation predictive model. This model monitoring the service request, system changes, allocation, resource it's based on the SOA. The SOA performance is depending upon to resource to service. The service request and an efficient resource allocation can be found the optimal are related to the correlation between correlation and hypothesis testing.

Zhangxi Lin et al., [10] propose the economic-based approach to resource allocation in web service and pricing for web service. The web service providing two models. **Profit maximization:** these focuses on the adopted, charges organization, user prices, and services, our own benefit of considering user utilities**. Welfare maximization model:** The organization adoptive, own benefit and set price with other utilities. It is based on the consumer consumes the network bandwidth, storage throughput, QoS, CPU Time. This approach improved the utilization of resource, increase benefit of cloud service providers and consumers.

Norman Bobroff et al., [11] propose the MFR and static algorithm. The MFR algorithm accessed the horizontal axis, static algorithm accessed vertical axis at specified the capacity of the violation rate presented the physical server to the virtual machine. It indicated the demand changes, migrates between these both machine hosts applications allocated. To minimize the physical machine workload used the technique as Time series forecasting and bin packing heuristic both algorithms is combined. These algorithms used to reduce the resource consumption and SLA violation.

Timothy wood et al.,[12] introduced new technology as automated black box and gray box, this technique is automatically monitoring system, hotspot detection, new mapping, load the migration. The black box important duties make the decision on within VM application resident. This both mechanism applied to sandipiper system. It implements hotspot detection algorithms.

These hotspot mitigation algorithms how much migration to allocate. The hotspot detection component monitoring, proofing engine usages gather on various virtual and physical servers. **Unobtrusive Black Box monitoring**: It is responsible for each virtual server processor, network, memory, usage, and finds the VM total resource usage. **Gray box monitoring**: It is used to lightweight monitor the each virtual server. In Linux monitoring demanded the interface together CPU, network, memory usage.**Profile Generation:** It received the report on the resource from each nucleus. It maintains user history; generate to report on the virtual machine to physical machines. Also maintained the CPU utilization. Network bandwidth, swap rates, memory, service time, drop rate and increment the request rate. **Hotspot Detection:** It responsible for signaling a need for VM migration with SLA violation. It is performed physical servers with a black box.

Monika choudhary et al., [13] designed the Greeding deadline based and cost based scheduling algorithm. It is based on the framework as follows, **Task Grouping:** It is a collection of compositions based on certain behavior. This behavior framework constraint can be a deadline or minimize the cost. Once the job enter into the group it can be judged the priority and scheduling order. It combined several tasks with reducing the cost of communication. **Prioritization:** It is important to task elements, this element order for the task scheduling parameters. In this framework short deadline based and priority vise executed. The deadline based task is a rearranged or ascending order for executing with minimum cost constrain. The cost based tasks are executed in descending order. This approach makes it maximize the profit and minimize the cost base of the VM. **Greedy Allocation:** The dynamic heterogeneous scheduling resource environment best suitable for greedy algorithms. It is resolving the job scheduling problem. **Deadline constrained based_ Greedy algorithm:** the Greedy algorithm is used to improve the competition of task. It is used to minimize the turnaround task from the individual task. Below the equation used to find the turnaround time.

$$\text{Turnaround Time} = \frac{\text{Resource waiting time} + \text{Task length}}{\text{Proc. Power of Resource}} \quad (1)$$

After calculating turnaround time for each resource, status can be updated.

Minimum cost based Greedy algorithm:

$$\text{Cost of task} = (\text{Task Length}/\text{Porc. Power of Resource}) * (\text{Resource Cost}) \quad (2)$$

According to above the equation resource selected with a minimum cost of sequential order. Both equations are simulated using CloudSim tool.

Ying Song et al., [14] has proposed the two-tier on_demand resource allocation mechanism,this mechanism consists of local and global resource allocation. It is based on feedback from an on-demand capacity of application also implemented a Xen_based prototype. Evaluated VM_based shared a platform with above mention technique. The set of algorithms for dynamic resource allocation with according to the various time-sharing resource demanded and quality of hosted application, whereas improving the performance of the critical application.

Yogitachawla et al., [15]Propose the dynamically optimized cost-based task scheduling algorithm. The cost based task scheduling algorithms used for cloud user and dynamically optimized resource allocation beneficial to cloud service providers. These can improve the communication cost ratio, grouping the user task before resource allocation. Here resource cost calculated as below the equation used.

Resource cost= (RAM of the virtual machine*Cost per memory)+(Size of virtual machine +Cost per storage) (3)

The VM is selected dynamically on the basis of cost, processing power and reduce the processing sequence of VMs, it can help CloudSimtool. These used to minimize the cloud user and service provider processing time for grouping and dynamic optimization are combined.

P.Malathi et al., [16] introduced the new method is SPAR. These methods are partitioning the hotspot, cold spot. The hotspot solves the connected with VM and if any PM hot thresholds for resource utilization, it can be reduced the load balancing with the job. The cold spot solves the rate of resource utilization on green computing. Here, using the concept of "skewness". This

can be improving the overall resource utilization, implemented and evaluated for dynamic resource allocation with multiple VM based on the end user. This achieves the green computing and load balancing with various end users.

B.Selvi et al., [17] proposed the Lopsidedness algorithm. This algorithm measures the unevenness of utilizing multiple resources on a server. It is combined the different type of workload and improve the server resource utilization .It described the **Hot and Cold Points:** It evaluates the resource allocation based on the future of VMs. The server defines the hotspot utilization by the resource threshold. If the server is not a hotspot, the temperature is zero. The Cold spots to define the utilization of all of the resource below a cold threshold. This indicates the idle state of the server and turns off to save energy on the server actively one VM running otherwise inactive. **Hotspot Reduction:** This aim eliminated the all the hotspot, otherwise keep their low possible temperature. It is sorting the list of hotspot with descending temperature. Also, reduce the VM temperature, the number of the dynamic server running low loads without performance, overload avoidance, save the energy system.

G. Sadasiva Prasad et al., [18] designed the mobile cloud and dynamically allocation transport sector. This technique applied to resource allocation for the different time and different routers. It is covering the speed of vehicles and distance of each route with mobile communication devices and traffic manager. It simulates the mathematical model for it finds real time solutions.

Saraswathi At et al.,[1] proposed the VM allocation model based on the job characterization and reconfigure virtual resource with dynamically. It increases the resource utilization. The jobs select low and high priority order.The job associated with **Cancellable:** It is not guaranteeing the job deadline and resumed later. This job can be scheduled after arrival time. **Suspendable:** It is guaranteeing the execution of the job. But not specified the deadline for a job. These can be scheduled any time of flexible time. **Non-Preemptable:** The entire request cannot be preempted. It is simulated using CloudSim tools, the jobs are allocated to a virtual machine on FCFS basis. Also, check the deadline for the job of low or high priority, the length of job completion.

Ravi Kumar U.Ighare et al., [19]Propose the threshold based dynamic resource allocation. It automates resource allocation and remapping the VM to PM. This method allocation policy consisted of three stages **Virtual machines select:** set the two threshold **Upper utilization threshold**: This can be maximizing the resource used in threshold servers. It is free to avoid SLA violation identifies the Hotspot .**Lower utilization thresholds**: This can be minimizing the resource utilization from a server with identifying the cold spot. Both utilization resources keep the hosts of these two thresholds. **Virtual Machine Distribution:** These algorithms detect the Hotspot and coldspots when migration is emulated. It is calculated as over and under by power consumption of VM. **Virtual machine Consolidation:** VM used for a minimum number of a host. This algorithm migrates the VM cold spot and turns off a server. It reduces the number of the server without degrading performance. It is simulated using CloudSim tool, provides the graphical user interface. The above the three stages achieved the overload and consolidated hosts.

Ruiu P et al., [20] proposed the DAViS, Dynamic Allocation of Virtual storage. It is based on the prototyping systems. The optimization resource utilization through dynamically and resize the storage to increase automatically. Below specified parameters have used this mechanism.**Time spent for attaching a new value:** Attached the guest's volume and time needed to execute all instruction. **The free space of the volume after the threshold is executed:** It depends on upon the threshold value and the total disk size.

III. RESULTS AND DISCUSSION

Compare the Existing Algorithm Strategies

Table I. The following table summarizes scheduling method, parameters, and other factors.

SNO	Resource Allocation Algorithm/ technique	Parameters	Simulated Tools	Advantage
1	[11] MFR and static algorithm	Time, frequency, capacity of VM Memory	Mathematical Model	To reduce the resource consumption and SLA violation

No	Technique	Parameters	Model/Tool	Objective
				completion.
2	[7]Nephel with Hadoop	Time, Memory, capacity of resource	Mathematical Model	To reduce the processing cost and improve the overall resource utilization for consequential.
3	[12]automated black box and gray box with sandipiper system	Network bandwidth, swap rate, memory, service time, drop rate	Mathematical Model	Sandpiper is capable of eliminating simultaneous hotspots involving multiple resources.
4	[13]Greeding deadline based and cost based scheduling algorithm	Time ,cost ,size of task	CloudSim	To reduce the cost of communication and improve the competition of task
5	[14]Novel two-tiered on-demand resource allocation mechanism	Time, resource cost	K-VM-1-PM and K-VM-N-PM Model	To improve the performance of critical application
6	[15] Dynamically optimized cost-based task scheduling algorithm.	Cost, time, Memory size	CloudSim	To improve the communication cost ratio, grouping the user tasks before resource allocation
7	[16]SPAR with Scheduling Algorithms	Memory, cloud usage, execution time	Mathematical Model	It can be reduced the load balancing of the job
8	[17] Lopsidedness algorithm	CPU utilization , Time	Mathematical Model	To reduce the VM temperature, the number of dynamic server running low load without performance, overload avoidance, save the energy system
9	[18] mobile cloud and dynamically allocation transport sector	Speed and distance	Mathematical Model	It finds real-time solutions
10	[1]Priority scheduling algorithm	Time, length of job , Disk size	CloudSim	To check the deadline for the job of low or high priority, the length of job
11	[19] threshold based dynamic resource allocation	Size, time ,cost	CloudSim	To achieve the overload and consolidated hosts
12	[20] Dynamic Allocation of Virtual storage	Time ,Space, Volume	Mathematical Model	To optimize the storage of VMs and storage resources, through autonomic and Resizing the storage of VMs occupation

IV. CONCLUSION

The resource allocation delaying the service request .The performance guarantee is an important issue of the web application. In the paper ,allocation of resource roles, domain knowledge, mathematical foundation, the deal of simulation, the rule of parameters for an order to handle the resource management problem, In comparison, summarize the based on the model, challenges, parameters, tools.

V. REFERENCES

[1]   Saraswathi AT a, Kalaashri.Y.RA b, Dr.S.Padmavathi c" Dynamic Resource Allocation Scheme in Cloud Computing", Procedia Computer Science ,47,PP. 30 – 36,2015

[2]   N.Krishnaveni, G.Sivakumar "Survey on Dynamic Resource Allocation Strategy in Cloud Computing Environment," International Journal of Computer Applications Technology and Research, Vol. 2, No- 6,PP. 731to- 737, 2013.

[3]   M.B. Qureshi et al..," Survey on Grid Resource Allocation Mechanisms", Springer March 2013

[4]   V.Vinothina, Dr.R.Sridaran, Dr.PadmavathiGanapathi" A Survey on Resource Allocation Strategies in Cloud Computing", (IJACSA) International Journal of Advanced Computer Science and Applications, Vol.3, No.6, June2012

[5]   Chee Shin Yeo and RajkumarBuyya" Service Level Agreement based Allocation of Cluster Resources: Handling Penalty to Enhance Utility"

[6] David L. Janovy, Jay Smith, Howard Jay Siegel, and Anthony A. Maciejewski1" Models and Heuristics for Robust Resource Allocation in Parallel and Distributed Computing Systems,"

[7] Daniel Warneke and Odej Kao "Exploiting Dynamic Resource Allocation for Efficient Parallel Data Processing in the Cloud," Ieee Transactions On Parallel And Distributed Systems, January 2011.

[8] Sharma et al., "Dynamic Resource Allocation in Grid Computing"International Journal of Advanced Research in Computer Science and Software Engineering 4(2), February - 2014, pp. 423-426

[9] Alexander Serebrenik and Natalia Sidorova" Predicting Service Request Rates for AdaptiveResource Allocation in SOA"

[10] ZhangxiLin, Huimin Zhao, SathyaRamanathan," Pricing Web Services for Optimizing Resource Allocation – An Implementation Scheme"

[11] Norman Bobroff, AndrzejKochut, Kirk Beaty "Dynamic Placement of Virtual Machines for Managing SLA Violations".

[12] Timothy Wood, PrashantShenoy, ArunVenkataramani, and MazinYousif "Black-box and Gray-box Strategies for Virtual Machine Migration"

[13] Monika Choudhary ,Sateesh Kumar Peddoju"A Dynamic Optimization Algorithm for Task Scheduling in Cloud Environment, " International Journal of Engineering Research and Applications , Vol. 2, No- 3, May-Jun 2012, pp.2564-2568

[14] Ying Song, Yuzhong Sun and Weisong Shi "A Two-Tiered On-Demand Resource Allocation Mechanism for VM-Based Data Centers," Ieee Transactions On Services Computing, Vol. 6, No. 1, January-March 2013.

[15] YogitaChawla, MansiBhonsle "Dynamically optimized cost based task Scheduling in Cloud Computing" International Journal of Emerging trends & technology in computer science,Vol. 2,No-3, May – June 2013.

[16] P. Malathi1, R. KanagaSelvi"Dynamic Resource Allocation for Green Cloud Computing,"International Journal of Research In Computer Applications And Robotics, Vol.2 Issue.1, Pg.: 139-144 January 2014

[17] B. Selvi, C. Vinola, Dr. R. Ravi "Efficient Allocation of Resources in Cloud Server Using Lopsidedness," International Journal of Computer Science and Mobile Computing, Vol.3, No-4, April- 2014, pg. 1007-1012.

[18] G.Sadasivaprasad, Dr.K.Rajagopal, Dr.K.Prahladarao "Dynamic Resource Allocation In Road Transport Sector Using Mobile Cloud Computing Techniques," International Journal of Mechanical Engineering and Technology (Ijmet), Vol. 5, No- 11, Nov.2014, PP. 17-25.

[19] RavikumarU.Ighare and R.C. Thool "Threshold based Dynamic Resource Allocation using Virtual Machine,"International Journal of Current Engineering and Technology, Vol.5, No.4, Aug. 2015

[20] Ruiu P, CaragnanoG.GragliaL"Automatic Dynamic Allocation of Cloud Storage for Scientific Applications," 2015 Ninth International Conference on Complex, Intelligent, and Software Intensive Systems