International Journal of Recent Trends in Engineering & Research
ISSN(ONLINE) : 2455-1457
IMPACT FACTOR : 3.344

Cloud resource Allocation: A step ahead Optimized Greedy Prioritization Approach (*OGPA*)

Shagun Agarwal[1] , Shubhankar Saxena2, Thirunavukkarasu K3, Dr. Ajay S. Singh4

[1,2] 2 Yr-B. Tech (CSE),Galgotias University, India
[3]Assistant Professor, SCSE, Galgotias University, India
[4] Professor, SCSE,Galgotias University, India

Abstract--Cloud computing is a new age of technology having long run scope and huge potential. Cloud has made possible to access user application and relevant resources as well as data from anywhere via remote access. Today enterprises hire resources from cloud for business aspect like storage, computing, etc. so as to reduce infrastructure cost and also to gain remote access for applications. Cloud computing, an emerging technology dealing with virtualization, is an on-demand service as it encompasses us with dynamic, flexible, convenient and efficient resource allocation for reliable and guaranteed service in pay-as-per-use manner. Since there are number of requests generated and cloud simultaneously allocates them resources efficiently without compromising the quality of service (QoS). The main problem arises in this scenario is the optimization of resources in order to increase the efficacy of the servers. This paper discusses about various resources allocation schemes and also it lays emphasis on prioritization of resources and greedy particle swarm optimization of resources.

Keywords–virtualization, optimization, resource allocation, prioritization, greedy particle swarm optimization.

I. INTRODUCTION

Cloud computing is a distributed architecture which centralizes server resources on scalable platform providing on demand computing resources and services. Here large number of clients is connected to the network (public and private) so as to utilize the dynamically scalable infrastructure for application, data and file storage. This significantly reduces the capital expenditure by avoiding the cost of infrastructural setup for these services. Cloud computing is a practical approach which experiences direct cost benefits and has transformed experiences direct cost benefits and has transformed data centers from capital intensive setup to variable priced dynamic environment. It provides IT resources in comparatively less capital and ease access. The cloud is not application oriented but is service oriented so it allows on-demand virtualization and pay-per-use utilities. Cloud computing provides infrastructure as a service (IaaS), platform as a Service (PaaS), Software as a Service (SaaS) and business processes as a solution. The cloud is reliable, custom made and provides QoS (Quality of Service). The problems that arises in cloud is low bandwidth, response time and delay but due to huge amount of resources and user requests there is a specific need of handling and optimizing the resources so as to provide better services and computing.

The main advantage of cloud computing for business enterprises is its scalability. Business enterprises can scale up and down as their required instances without investing in setting up their own IT infrastructure including computer hardware, network, computing and also there is no need to worry for licensed software and patch management. This reduces huge amount of capital investment and also management of resources is ease.

In this paper, we have discussed various cloud computing resource allocation technologies which are proficiently reducing the resource wastage and maximizing the profits. Resource allocation is very challenging task as we have to take care of both, user requirements as per their needs as well as server's

performance without violating the SLA. The allocation of resources mainly focuses on assignment of computing power, memory, network bandwidth, storage, etc. and also it has to manage the Quality of Service (QoS) and maintain the Service Level Agreement.

The paper also proposes an algorithm based on prioritization of resources along with greedy particle swarm optimization which applies priority on resources over parameters like time, cost, processor requirements, resource requirements and then introduces the result after prioritization in GPSO (Greedy Particle Swarm Optimization) algorithm to allocate the resources using greedy heuristics approach and particle swarm optimization to optimize configuration based on workloads. The GPSO algorithm initially will be receiving the prioritized information from priority algorithm and then allocation of resources is done by GPSO so as to provide better services and maximizing the profits.

II. LITERATURE SURVEY

Allocation of resources in cloud is done so as to satisfy the demands of client and also to maximize the profit of organization providing cloud services. Here are some of the resource allocation techniques-

Ikki fujiwara et al [1] proposes a resource allocation technique in which the participants can trade off by double sided combinational auction. It proposed a market based resource allocation which is very effective and enables participants to trade future and current services in forward market and the spot market.

In Zuling et al [2], a new resource allocation algorithm, CRAA/FA (cloud resource allocation algorithm via fitness-enabled Auction) which provides market for cloud resources and allows resource agents and service agents to bargain in market via dynamic asking/bidding strategy.

Zhenzhong et al [3] introduces a Statistical based Load Balance method to solve the problem of load imbalancing in cloud environment where data analysis of on-line historical performance is done for forecasting the resource demands of each VM.

In Rasoul Beik [4], green cloud computing technique is delivered so as to minimize energy wastage and also it highlights various greedy approaches to improve the efficiency of CPU. To incorporate energy efficiency designing of Energy-aware allocation layer based on greedy approach is implemented.

In Thangaraj et al [5], is based on predefined allocation policy to allocate resources for Infrastructure as a Service (IaaS). It focuses on deadline sensitive policy for allocation of resources successfully by reducing the request rejection with the help of an open source resource lease manager, Haizea, which acts as a scheduler.

Tomita et al [6], proposed the congestion control method for a cloud computing environment which reduces the size of required resource for congested resource type, instead of restricting all service requests as in the existing networks.

Atsuo Inomata et al [7], has proposed a dynamic resource allocation method based on the load of VMs on IaaS, abbreviated as DAIaS. This method enables users to dynamically add and/or delete one or more instances on the basis of the load and the conditions specified by the user.

III. PRIORITY ALGORITHM OVERVIEW

The resource allocation model based on priority [8] is very efficient as it decides the priority among different users based on various key parameters and cloud administrator plays a key role in deciding the efficiency as he decides the priority among the users. Various parameters are considered such as cost, time, profit, users, number of processors requests, resources assigned, available resources, criteria of selection, user logs, etc. For each task different parameters are considered depending on the type of requests, multiple users submit their requests demanding same resource. This creates trouble for cloud admin to decide whom to allocate the resources so this algorithm helps cloud administrator to set priority among users for allocation of resources effectively. This technique is more efficient as it adds decision taking capability for resource allocation.

TABLEI. PARAMETERS CONSIDERED FOR PRIORITIZATION

No of Users	eg. : 10 users
Servers	eg. : S1, S2, S3
Time to run	eg. : 4 hours
No of processor requests	eg. : 8 Processors
Amount of memory	eg. : 5 GB, 1 TB etc
Time of request	eg. : 2:00 pm
Software to be Used	eg. : Matlab
Job type	eg. : sequencial or parallel
User type	eg. : internal or external

Algorithm for prioritization of the resources:

Algorithm: *To compute and assign the priority for each request based on the threshold value and allocate the service to each request's*

***Step 1**: [Read the client's request data i.e., time, importance, price, node and requested server name] Insert all values into the linked list.*
***Step 2**: [For each request and its task find the time priority value based on predefined conditions] Assign priority value to each task for the client's request.*
t_p[i]=priority value
***Step 3**: [For each request and its tasks find the node priority value based on predefined conditions] Assign priority value to each task for the client request.*
n_p[i]=priority value
***Step 4**: [For each client's input data check whether it is in the threshold value or not]*
if (input value is within the threshold limit and total node <= available node)
[Add respective computed time and node priority value and other parameters like importance and price.]
Sum[k]=t_p[i] + n_p[i] + importance + price
Print—ready to execute.
available node = available node – total node
else if (input value is within the threshold limit)
sum[k]=t_p[i] + n_p[i] + importance + price
print—within the limit but it is in a queue
else
print—exceed the condition
***Step 5**: [Sort the sum[k] values]*
***Step 6**: [Client's request is ready to execute from least value of sum[k]*
Stop

IV. GREEDY PARTICLE SWARM OPTIMIZATION

A hybrid of heuristics greedy search and particle swarm optimization [9] is a new algorithm which is called Greedy Particle Swarm Optimization (GPSO). This algorithm requires more computation but enhances the VM configurations and the particle swarm optimization lets out the greedy algorithm to get trapped in local optima.

- The PSO module sends the candidate shares(particles) and VM configurations to the greedy module.
- The greedy module from PSO module takes particles and VM configurations and returns the updated VM configurations and the corresponding estimated costs for those configurations.

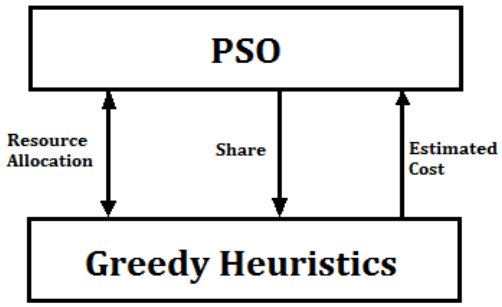
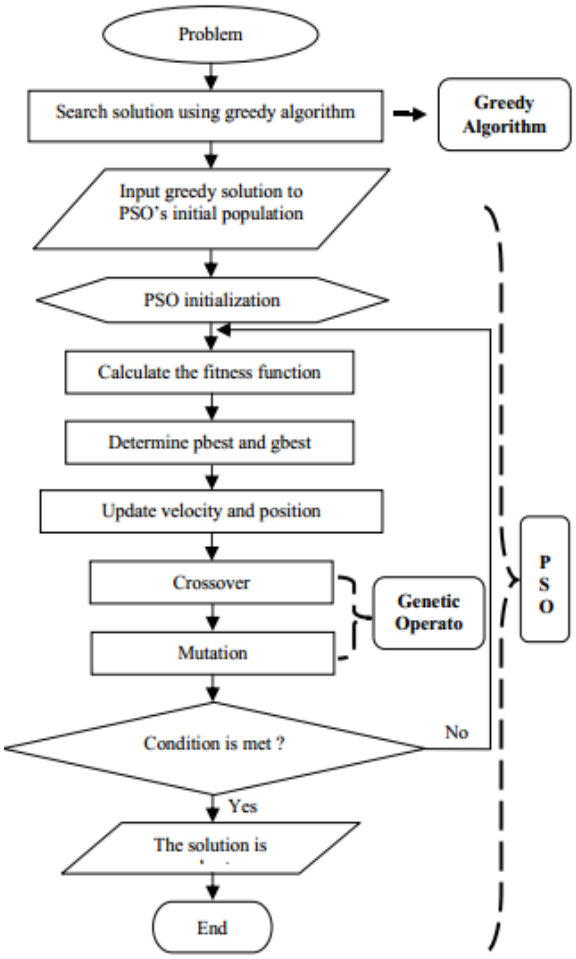


Fig. 1. GPSO module



Flow chart of Greedy Algorithm

V. OPTIMISED GREEDY PRIORITIZATION APPROACH

There are various resource allocation techniques which are implemented and are managing the resources and their allocation. As there are so many users demanding the same resources or computing facilities such as high performance computational environments dealing with scientific simulations and other higher resources as memory, CPU, bandwidth etc. The proposed optimized greedy approach allows cloud a technique of resource allocation where cloud admin can allocate priority to the desired requests on the basis of parameters such as time, processor requests, its importance and price. That data is processed in priority algorithm and is sorted according to the set priority. Now the sorted requests will be automated to greedy particle algorithm where the VM configuration and shares of the resource request and priority are processed and greedy algorithm allocates the resources in an optimized manner so as to fulfill the demands of customer and maximizing the profits.

VI. CONCLUSION

Cloud Computing technology is increasingly being used in business markets and enterprises. In cloud paradigm, an effective and optimized resource allocation strategy is required so as to achieve user satisfaction and maximizing the profits for cloud service providers. This paper gives a brief introduction to some of the resource allocation techniques and emphasize on priority resource algorithm and Greedy Particle Swarm Optimization and also gives a new idea on over a hybrid approach between priority algorithm and GPSO algorithm for maximum optimization of resource. Hence this paper will hopefully motivate the future researchers to come up with better resource allocation techniques.

REFERENCES

1. Ikki Fujiwara, Kento Aida, "Applying Double-sided Combinational Auctions to Resource Allocation in Cloud Computing", 2010 10th Annual International Symposium on Applications and the Internet, ISSN: 978-0-7695-4107-5/10, DOI 10.1109/ SAINT.2010.93.
2. Zuling Kang, Hongbing Wang, "A Novel Approach to Allocate Cloud Resource with Different Performance Traits", IEEE 10th International Conference on Services Computing, 2013, ISSN: 978-0-7695-5026-8/13, DOI 10.1109/ SCC.2013.109.
3. Zhenzhong Zhang, Haiyan Wang, Limin Xiao, Li Ruan, "A Statistical based Resource Allocation Scheme in Cloud", International Conference on Cloud and Service Computing, 2011, ISSN: 978-1-4577-1637-9/11.
4. Rasoul Beik, "Green Cloud Computing: Greedy Algorithms for Virtual Machines Migration and Consolidation to Optimize Energy Consumption in a Data Center", International Journal of Digital Application & Contemporary research (Volume 2, Issue 9, April 2014).
5. Thangaraj P, Soundarrajan S, Mythili A, "Resource allocation policy for IaaS in Cloud computing", International Journal of Computer Science and Management Research, Vol 2, Issue 2, pp 1645 - 1649, February 2013, ISSN 2278-733X.
6. Kazuki MOCHIZUKI and Shin-ichi KURIBAYASHI, Evaluation of optimal resource allocation method for cloud computing environments with limited electric power capacity, 2011 International Conference on Network-Based Information Systems.
7. Atsuo Inomata, Taiki Morikawa, Minoru Ikebe. 2011, Proposal and Evaluation of a Dynamic Resource Allocation Method based on the Load of VMs on IaaS, IEEE 2011.
8. K C Gouda, Radhika T V, Akshatha M, "Priority based resource allocation model for cloud computing", *International Journal of Science, Engineering and Technology Research (IJSETR) Volume 2, Issue 1, January 2013, ISSN: 2278 – 7798.*
9. Fatma A. Omara, Sherif M. Khattab, Radhya Sahal "Optimum Resource Allocation of Database in Cloud Computing", Egyptian Informatics Journal (2014) 15, 1–12.
10. Sudeepa R, Dr. H S Guruprasad, "Resource Allocation in Cloud Computing", International Journal of Modern Communication Technologies & Research (IJMCTR) ISSN: 2321-0850, Volume-2, Issue-4, April 2014.

