Improvement of dynamic load balancing in Cloud Environment

Ginni Bansal, Amanpreet Kaur

Research Scholar, Department of Information Technology, CEC Landran, India

Assistant Professor, Department of Information Technology, CEC Landran, India

ABSTRACT: Dynamic load balancing with decentralized load balancer using PSO technique: Cloud consists of multiple resources and various clients request to the cloud for allocation of shared resources. Each request will be allotted to the virtual machines. In different situation different machines get different load. So to balance the load amongst different virtual machines decentralized load balancer is enhanced using particle swarm algorithm. The main objective is reducing the energy and increasing the throughput in comparison to centralized and simple decentralized load balancer using particle swarm optimization.

KEYWORDS: PSO, Centralized, Decentralized, Energy, Throughput

I. INTRODUCTION

With high flexibility and great retrieval of data as per users' requirements, cloud computing provides numerous services. To handle a very large amount of data several techniques to optimize load and streamline operations are needed to achieve desired performance level for the users. The workload of a processor can be defined as the total time required by the processor to execute all the assigned processes. Load balancing is to ensure that every processor in the system does approximately the same amount of work at any point of time [11]. Load balancing is required so that time of total resource finding can be minimized. As well as rather than having load on all the machines load can be given on all the machines evenly.
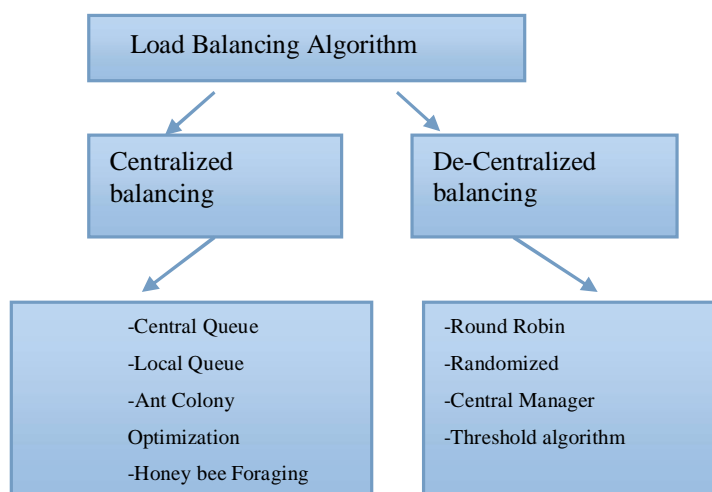

Figure 1. Type of load Balancing

1.1 Centralized load balancing algorithm:
The work load is distributed among the processor at runtime. In this mechanism, master assigns new processes to the slaves based on the new information collected.

Work is central. In non distributed manner one node execute the load balancing algorithm and task of load is shared among them.

Nodes interact in two ways: cooperative and non-cooperative [2].

The main advantage here is, the total load balancing process will get affected, if, one or more node stop working it will just affect the overall performance of system in a certain manner.

In central type, the task of load balancing is done by either single node or group node.

Central load balancing takes two forms: centralized and semi-distributed. In centralized form one node is solely responsible for load balancing of the whole system and other nodes simply interact with the central node.

1.2 Decentralized load balancing algorithm:

It depends on a priori information of the applications and static information about the load of the node.

They do not consider the existing state of system; rather they consider processing power, memory and storage capacity and recently known communication performance.

Distributed algorithms are basically suitable for homogeneous and steady environments. Distributed algorithms always work in master – slave manner, where the performance of any processor is determined before starting the actual execution [3] [4] [5].

1.3 Particle swarm algorithm (PSO) PSO is a swarm based heuristic optimization technique. It is used for identifying the optimal path of solution space. While putting up the load on specific virtual machine for processing of the resources, it moves along all the virtual machines and identifies the optimal machine to put the load. It is one of the mechanisms to identify the optimal V.M, which is load less, available and task map. So the relative energy and time utilization to process the node can be reduced.

Basic Steps for PSO:

1. Initialize population of particles with random position and velocities.
2. Calculate the fitness function value for each and every particle.
3. Compare current particle's fitness value with each particle's fitness value and find Pbest value.

II. LITERATURE SURVEY

In [14], Dr. M.Sridhar et al. defined scheduling is a task performed to get maximum profit to increase cloud computing work load efficiency. For this, resource utilization and managing of load between resources with minimum execution time becomes the main objective. Optimization is the selection of best element (pertaining to specified criteria) from available variable alternatives with the goal to i.e. to accomplish – "maximal output with minimal input". So, a hybrid Particle Swarm Optimization (PSO) is proposed which performs better in execution ratio and average schedule length when it is compared with Max-min scheduling and minimum execution time.

Author Madhurima Rana et al. in [6] discussed Load balancing that ensures no single node will be overloaded and used to distribute workload among multiple nodes improving the system performance and ensuring proper utilization of resources. It also minimizes the time and cost involved in big computing models. To overcome load balancing problem a summary is provided of evolutionary and swarm based algorithms in different environment of cloud. Various soft computing approaches to optimize the load are discussed like Genetic algorithm, Particle swarm optimization, Ant colony optimization, artificial bee colony and other various algorithms. The issues involved in these techniques are listed in a tabular form comparing each other.

The chaos cloud particle swarm optimization algorithm based on the golden section evaluation criteria is presented by Xi Song et al in [4]. Particle swarm is divided into standard particle, chaos-cloud particle and cloud particle using the judge principles based on golden section according to the fitness value. Each population is operated by the different algorithm. An optimal power flow model for Available Transfer Capability (ATC) under the static security constraints is established. The algorithm proposed solves the problems of easily falling into local optimum in basic PSO and the drawback of repeatedly search part of solutions in chaos optimization. It has high accuracy in ATC calculation and can make full use of power resources.

Gulshan Soni et al. discussed the biggest challenge for cloud data centers i.e. how to handle and service the millions of requests that are arriving very frequently from end users efficiently and correctly in [2]. In cloud computing, load balancing is required to distribute the dynamic workload evenly across all the nodes. Load balancing helps to achieve a high user satisfaction and resource utilization ratio by ensuring an efficient and fair allocation of every computing resource. Proper load balancing aids in minimizing resource consumption, implementing fail-over, enabling scalability,

avoiding bottlenecks and over-provisioning etc. "Central Load Balancer" is a load balancing algorithm to balance the load among virtual machines in cloud data center. Results showed that the algorithm can achieve better load balancing in a large-scale cloud computing environment as compared to previous load balancing algorithms.

In [1], Michael Pantazoglou et al. Discussed decentralized approach towards scalable and energy-efficient management of virtual machine (VM) instances that are provisioned by large enterprise clouds. Also, the computation resources of the data center are effectively organized into a hypercube structure. The hypercube seamlessly scales up and down as resources are either added or removed in response to changes in the number of provisioned VM instances. Without supervision from any central components, each compute node operates autonomously and manages its own workload by applying a set of distributed load balancing rules and algorithms. On one hand, underutilized nodes attempt to shift their workload to their hypercube neighbors and switch off. On the other, over utilized nodes attempt to migrate a subset of their VM instances so as to reduce their power consumption and prevent degradation of their own resources, which in turn may lead to SLA violations. In both cases, the compute nodes in our approach do not overload their counterparts in order to improve their own energy footprint. An evaluation and comparative study of the proposed approach provides evidence of its merits in terms of elasticity, energy efficiency, and scalability, as well as of its feasibility in the presence of high workload rates.

Enhancement of the make span of particle swarm optimization based dynamic scheduling in cloud environment is done in [17] by Azade Khalili et.al. Mapping and scheduling the tasks is assigning task to run on the existing resources that helps to maximize utilization and minimize make span. The objective was to optimize task scheduling that uses PSO algo to minimize make span by using different inertia weights. The linear descending inertia weight(LDIW) with an average 22.7% reduction in make span shows best performance.

Jun Zhang et al. proposed a Set-Based PSO approach. It tackles a cloud workflow scheduling problem which enables users to define various Qos constraints like deadline constraint, budget constraint and reliability constraint in [9]. It enables users to specify one preferred Qos parameter as the optimization objective. Defined penalty based fitness functions to address multiple Qos constraints and integrate S-PSO with seven heuristics. A discrete version of Comprehensive Learning PSO algorithm based on S-PSO is implemented.

Geng Yushui et al. in [24] defined data migration which is the key technology to realize the nodes dynamically extensible and elastic load balancing. To reduce migration cost of time is the problem that cloud service providers need to solve. In [18], Hongwei Zhao et al. designed PSO algorithm in order to implement the balanced distribution in Cloud Computing system and to improve the utilization ratio of the resource as well as handling up rate of the system. The system of dynamic dispatching system based on Particle swarm optimization (PSO) for Cloud Computing Environment has been s and implemented after the study on the Cloud Computing.

III. COMPARATIVE ANALYSIS OF PAPERS

Paper Name	Work Undertaken	Constraints
A Set-Based Discrete PSO for Cloud Workflow Scheduling with User-Defined QoS Constraints	A S-CLPSO approach has been designed for the cloud workflow scheduling problem.	To address different QoS factors like reliability, time and cost, seven heuristics are applied to integrate with the SCLPSO approach.
Self-Adaptive Learning PSO-Based Deadline Constrained Task Scheduling for Hybrid IaaS Cloud	An integer programming model is established for the resources allocation problem of an IasS cloud in a hybrid cloud environment.	From cloud providers' perspective, effectively allocating limited resources is important to maximize its profit and guarantee the QoS.
Hybrid Particle Swarm Optimization Scheduling for Cloud Computing	Hybrid Particle Swarm Optimization (PSO) is proposed for scheduling in cloud. The hybrid PSO performs better compared to Max Min Scheduling	PSO performs well in global search but not so well in local search.

Cloud Data Migration Method Based On PSO Algorithm	To cloud storage systems, data migration is key technology to realize the nodes dynamically extensible and elastic load balancing.	It is a test framework designed to help users understand the different cloud computing, database performance.
A Study on Load Balancing in Cloud Computing Environment Using Evolutionary and Swarm Based Algorithms	It provides a pool of shared resources to the users available on the basis of pay as you go service, means users pay only for those services which are used by him according to their access times.	Summary of evolutionary and swarm based algorithms which will help to overcome load balancing problem in different environment of cloud.

IV. CONCLUSION

In this paper extensive load balancing is done based on PSO using decentralized load balancing technique. On taking up the decentralized load balancing by PSO technique the aim is achieved. According to the technique before distributing the load there will be a check on different virtual machines identifying which machine is best machine to take up the load. So according to PSO best state in whole solution space is chosen. Best machine is in terms of more processing power and more buffer space and has competitively less load for the resource allocation or even can take more loads due to large capability. In result, it can increase the throughput of the system and more energy efficient. Main goal is to have load balancing and distributing the load on each machine for better utilization of the resources.

REFERENCES

1. Michael Pantazoglou, Gavriil Tzortzakis, and Alex Delis, "Decentralized and Energy-Efficient Workload Management in Enterprise Clouds", in press, IEEE 2015.
2. Gulshan Soni and Mala Kalra, "A Novel Approach for Load Balancing in Cloud Data Center", IEEE International Conference on Intelligent Computing and Integrated Systems (ICISS), Guilin,vol.14, pp. 807-812, 2014.
3. Cristian Mateos, Elina Pacini & Carlos Garc Garino, An ACO-inspired algorithm for minimizing weighted flow time in cloud-based parameter sweep experiments, 2013.
4. Hongsheng Su, Ying Qi and Xi Song, "The Available Transfer Capability Based On a Chaos Cloud Particle Swarm Algorithm ", IEEE ninth International Conference on Natural Computation (ICNC), vol 13, pp.574-579, 2013.
5. Rajkumar Buyya,"A Particle Swarm Optimization-based Heuristic for Scheduling Workflow A", Cloud Computing and Distributed Systems Laboratory, Department of Computer.
6. Madhurima Rana, Saurabh Bilgaiyan and Utsav Kar, "A Study on load balancing in cloud computing environment using evolutionary and swarm based algorithms", IEEE International Conference on Control, Instrumentation, Communication and Computational Technologies, vol.14, pp. 245-250, 2014.
7. Pooja Samal and Pranati Mishra, "Analysis of variants in Round Robin Algorithms for load balancing in Cloud Computing", International Journal of Computer Science and Information Technologies, vol.4(3), pp. 416-419,, 2013.
8. Wang Yonggui, Han Ruilian. Study on cloud computing task schedule strategy based on MACO algorithm [J]. Computer Measurement & Control, vol.19 (5), pp.1203- 1211, 2011.
9. Jun Zhang and Wei-Neng Chen, "A Set-Based Discrete PSO for cloud Workflow Scheduling with User-Defined Qos Constraints, IEEE International conference on Systems, Man And Cybernetics, vol.12,pp. 773 – 778, 2012.
10. J. Kennedy and R. Eberhart, Particle swarms optimization In IEEE International Conference on Neural Networks, vol. 4, pp 1942–1948, 1995.
11. Andrew J. Page and Thomas J. Naughton, "Dynamic task scheduling using genetic algorithms for heterogeneous distributed computing", 19th IEEE International Conference on Parallel and Distributed Processing Symposium, pp. 189a, 2012.
12. Akhil Goyal,Bharti, "A Study of Load Balancing in Cloud Computing using Soft Computing Techniques", International Journal of Computer Applications, 2014.
13. Zhanghui Liu and Xiaoli Wang," A PSO-Based Algorithm for Load Balancing in Virtual Machines of Cloud Computing Environment", Advances in Swarm Intelligence Lecture Notes in Computer Science ", pp 142-147, 2012.
14. Dr. M.Sridhar and Dr. G..Rama Mohan Babu , " Hybrid Particle Swarm Optimization scheduling for Cloud Computing " , IEEE International Advance Computing Conference, vol. 15, pp 1196-1200, 2015.
15. Zehua Zhang, Xuejie Zhang, "A Load Balancing Mechanism Based on Ant Colony and Complex Network Theory in Open Cloud Computing Federation", 2nd IEEE International Conference on Industrial Mechatronics and Automation, pp. 240-243, 2010.
16. Kun Li, Gaochao Xu, Guangyu Zhao, Yushuang Dong, Dan Wang, "Cloud Task scheduling based on Load Balancing Ant Colony Optimization", Sixth IEEE Annual ChinaGrid Conference, pp. 3-9, 2011.
17. Azade Khalili and Seyed Morteza Babamir, "Makespan Improvement of PSO-based Dynamic Scheduling in cloud environment", 23rd IEEE Iranian Conference on Electrical Engineering (ICEE), vol.15, pp.613-617, 2015.
18. Hongwei Zhao and Wang Chenyu, "A Dynamic Dispatching Method of Resource based on Particle swarm optimization",10th Web Information System and Application Conference for Cloud Computing Environment, vol.13, pp.351-354,2013.

19. Sung-Soo Kim, Ji-Hwan Byeon, Hongbo Liu, Ajith Abraham and Seán McLoone, "Optimal job scheduling in grid computing using efficient binary artificial bee colony optimization", Soft Computing, Springer, vol. 17, No. 5, pp. 867-882,2013.
20. Marco Dorigo, Gianni Di Caro Luca and M. Gambardella, "Ant Algorithms for Discrete Optimization", Artificial Life, Massachusetts Institute of Technology, pp. 137-172, 1999.
21. Nidhi Jain Kansal and Inderveer Chana, "Cloud Load Balancing Techniques:A Step Towards Green Computing", International Journal of Computer Science Issues, vol.9, No.1, pp.238-246, 2012.
22. Qinghai Bai, "Analysis of Particle Swarm Optimization Algorithm", Computer and Information Science, vol. 3, No. 1, pp. 180-184, 2010.
23. Particle Swarm Optimization, "http://en.wikipedia.org/wiki/Particle_swarm_optimization"
24. Geng Yushui and Yuan Jiaheng, "Cloud data migration method based on PSO algorithm", 14th International Symposium on Distributed Computing and Applications for Business Engineering and Science, pp.143-146, 2015.
25. Salim Bitam, "Bees Life Algorithm for Job Scheduling in Cloud Computing",Proceedings of The Third International Conference on Communications and Information Technology, pp. 186-191,2012.
26. Kaur, AmanPreet, and Dheerendra Singh. "enlightening the cloud computing domain." international journal of computers & technology 4.2b2 (2013): 470-476.
27. Kaur, AmanPreet, and Kanika Minhas. "Mitigation of High Rate Shrew DDOS Attack" International Journal of Modern Computer Science (IJMCS) Volume 3, Issue 3, October, 2015.
28. Kaur, AmanPreet, and Dr. Dheerendra Singh. "Exploring Data Security Issues and Scrutinizing them in Cloud Computing Environment" International Journal of Modern Computer Science (IJMCS) Volume 8 • Number 1 June-Dec 2014.