Hybrid Particle Swarm Optimization based Load Balancing in Cloud Computing

[1]Dr.M.Thillainayaki, [2]Mr.D.Shyam Sundar, [3]Ms.N.Saranya, [4]Mr.M.Vijayakumar, [5]Ms.R.Mahadevi
[1,2,3,4,5]Assistant Professor, Nehru Arts and Science College, Coimbatore, India,
[1]thillainayakim@gmail.com, [2]ntshyam@gmail.com, [3]nsaranyanataraj@gmail.com,
[5]mahadevi.ramasamy@gmail.com

Abstract - Cloud computing is known as a provider of dynamic services using very large scalable virtualized resources over the Internet. Clouds consist of a collection of virtualized resources, which include both computational and storage facilities that can be provisioned on demand, depending on the users' needs and users are charged on a pay-per-use basis. Scheduling and load balancing is a troublesome assignment in cloud environment. A workflow is a model that structures a machine-meaningful portrayal of business conduct varying from that of code. Cloud computing helps user applications dynamically provision as many compute resources at specified locations. Workflow scheduling plays a vital role in the workflow management. A cloud workflow system is a type of service which facilitates the automation of distributed applications based on the novel cloud infrastructure.

Key Words: Cloud Computing, Infrastructure, Load Balancing, Scheduling, Virtualized, Workflow

I. INTRODUCTION

Cloud computing refers to the delivery of computing resources over the Internet. Cloud services allow individuals and businesses to use software and hardware that are managed by third parties at remote locations. Instead of keeping data on your own hard drive or updating applications for your needs, use a service over the Internet, at another location, to store your information or use its applications. The cloud computing model allows access to information and computer resources from anywhere that a network connection is available.

II. SERVICE MODELS

Software as a Service (SaaS): Consumers purchase the ability to access and use an application or service that is hosted in the cloud. A benchmark example of this is Salesforce.com, where necessary information for the interaction between the consumer and the service is hosted as part of the service in the cloud.

Platform as a Service (PaaS): Consumers purchase access to the platforms, enabling them to deploy their own software and applications in the cloud. The operating systems and network access are not managed by the consumer, and there might be constraints as to which applications can be deployed.

Infrastructure as a Service (IaaS): Consumers control and manage the systems in terms of the operating systems, applications, storage, and network connectivity, but they do not control the cloud infrastructure.

III. PSO CONCEPT

A swarm intelligence based algorithm finds a solution to an optimization problem in a search space.

- Proposed solution exists in the form of a fitness function.
- The swarm is typically modeled by particles in multidimensional space that have a position and a velocity.
- A Particle is a candidate solution in the population and represents a task.
- Particles fly through hyperspace.
- An iterative process to improve candidate solutions is set in motion. The particles iteratively evaluate the fitness of the candidate solutions.

IV. GA-BASED TASK SCHEDULING FOR CLOUD COMPUTING

The GA-based task scheduling model and illustrate its design. This work explains how to make an optimal task schedule and compose elements of the GA scheduling function. As mentioned, a cloud user reaches a SLA with a cloud provider to process a task. A SLA document includes user requirements like time and budgetary constraints of the task, which indicate acceptable deadline and payable budget of the cloud user. QoS attributes like response time and throughput can be comprised in a SLA document besides time and budgetary constraints. A cloud provider has to consider user requirements and virtual machine information before allocating tasks from to virtual machines.

 DOI : 10.35291/2454-9150.2020.0067

Following distributed computing, parallel computing, grid computing, utility computing, etc., the computer industry and academia put forward cloud computing model, which achieves generalization and commercialization of these previous models in some sense. Cloud computing, the long-held dream of computing, has the potential to transform a large part to makes software even more attractive as a service and shaping the way hardware is designed and purchase. No doubts it would increasingly change the way people live and work.

Cloud computing can be defined as "a type of parallel and distributed system consisting of a collection of interconnected and virtualized computers that are dynamically provisioned and presented as one or more unified computing resources based on service-level agreements established through negotiation between the service providers and consumers". With the advancement of the Cloud, there are new ways of opening up on how applications can be built on the internet. There are various cloud service providers who are willing to provide large scaled computing infrastructure at a cheaper price and provide the infrastructure services in a very effective manner which the users can scale up or down at will. There are also large scaled software systems such as social networking sites and e-commerce applications gaining popularity today which can benefit greatly by using such cloud services to minimize costs and improve service quality to the end users.

V. COMPUTING TECHNIQUES

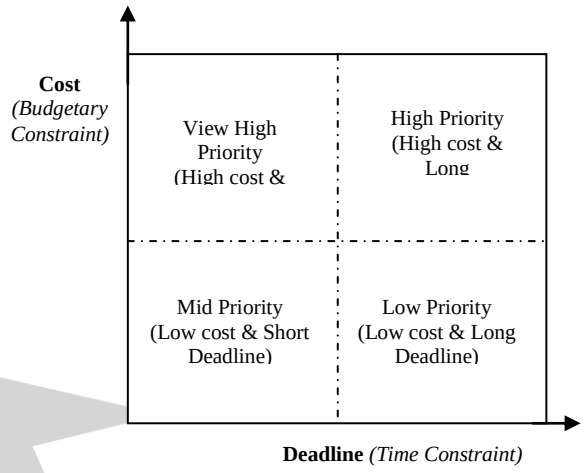
The standard PSO process is outlined as follows:
Step 1: Initialize a population of particles with random positions and velocities in a *D*-dimensional search space.
Step 2: Update the velocity of each particle.
Step 3: Update the position of each particle.
Step 4: Map the position of each particle into the solution space and evaluate its fitness value according to the desired optimization fitness function. Simultaneously update the *pbest* and *gbest* positions if necessary.

Step 5: Loop to Step 2 until the termination criterion is met, usually after a sufficient good fitness or a maximum number of iterations.

Encoding and Initiation

A chromosome ch_k indicates the task allocation information, i.e. a task schedule. k is from 1 to z, which denotes the number of chromosomes in a population. In consideration of user satisfaction and provider's profit, the task scheduler determines where to allocate each task every scheduling cycle. A chromosome ch_kconsists of α[i] and β[i], which indicates the information of task processing and virtual machine allocation. The length of a chromosome is the same as the number of inputted tasks.

The encoding operation to express a task schedule as a chromosome is as follows. A set of inputted tasks is sorted before the creation of chromosomes. Unlike existing distributed computing environments, cloud users pay for computing services in person. Accordingly, tasks received from cloud users with high cost have to be allocated to virtual machines faster than others. In this work, classify takes into four groups by time and budgetary constraints as shown in Figure.



Restart & Stop Condition

If the best fitness value of the current population does not reach the minimum fitness threshold, chromosomes with high fitness values in the current population fill the half of the next population. And, the rest of the next population is filled with chromosomes generated randomly. This restart operation helps the GA scheduling function to prevent local optimum and explore various search places.

Finally, if particle gets stagnated at any place in solution space, then it is shifted to new location. The pseudo code of HPSO is presented.

1. Set dimension of particle as number of ready tasks
2. Initialize particles position and velocity vectors randomly
3. By using SPV rule, find discrete vector $S_{ik}=\{S_{i1},S_{i2},\ldots S_{in}\}$ from the continuous position vector $X_{ik}=\{ X_{i1}, X_{i2}\ldots X_{in}\}$ for each particle and then map discrete vector S_{ik} elements of each particle into resource vector $R_{ik}=\{ R_{i1}, R_{i2}\ldots R_{in}\}$
4. Evaluate the fitness of each particle
5. If fitness value is better than personal best pbest, update pbest by current fitness value. Select the best particle as gbest.
6. Update velocity and position of each particle through If a particle gets stagnated for predetermined number of iteration, then particle is shifted to new location
7. If maximum iteration is reached or stopping criteria is satisfied, then stop. Otherwise repeat from step 3

 DOI : 10.35291/2454-9150.2020.0067

A further drawback is that stochastic approaches have problem-dependent performance. This dependency usually results from the parameter settings in each algorithm. The different parameter settings for a stochastic search algorithm result in high performance variances. In general, no single parameter setting can be applied to all problems. Increasing the inertia weight (w) will increase the speed of the particles resulting in more exploration (global search) and less exploitation (local search) or on the other hand, reducing the inertia weight will decrease the speed of the particles resulting in more exploitation and less exploration. Thus finding the best value for the parameter is not an easy task and it may differ from one problem to another. Therefore, from the above, it can be concluded that the PSO performance is problem-dependent. The problem-dependent performance can be addressed through hybrid mechanism. It combines different approaches to be benefited from the advantages of each approach.

The comparison of scheduling algorithm using parameters

Scheduling Algorithm	Resource Matrix	Dead Line (%)
PSO	128 x 8	85.7
Hybrid PSO/GA		93.76
PSO	256 x 16	78.17
Hybrid PSO/GA		82.23
PSO	512 x 32	73.5
Hybrid PSO/GA		76.3
PSO	1024 x 64	75.2
Hybrid PSO/GA		79.73

A novel hybrid PSO-GA method that aims to increase swarm diversity when a stagnation condition is detected is proposed. This article aims to solve job scheduling problems, and compare it with other techniques. In this research, a new optimized scheduling hybrid algorithm was presented named HPSOGA. HPSOGA is one independent Task scheduling to decrease Make span and minimize the missed tasks. Hybrid Particle Swarm Optimization (HPSO) based scheduling heuristic to balance the load across the entire system while trying to minimize the makespan of a given task sets. The experimental results showed Effectiveness of the proposed combined approach in finding optimal solutions.

VI. CONCLUSION

Security in cloud computing must be approached cautiously. Many organizations are moving to the Cloud, so the impact of potential breaches gets higher and higher. This thesis has provided and overview of cloud computing, including the proposed dynamic scheduling algorithm, PSO and hybrid PSO-GA techniques. The focus of the thesis is to understand the actions and workflow scheduling algorithms in cloud computing. The cloud computing has multiple service and deployment models with their own security concerns. The problem of task scheduling in cloud computing environment is evaluated. PSO algorithm and Genetic algorithm are most famous algorithms for scheduling tasks in distributed systems. In order to improve the performance of standard algorithm the modified PSO algorithm is suggested, in which hybridization is merged into standard PSO algorithm for generating initial population in order to minimize makespan. This algorithm can be used in cloud computing environment for efficient scheduling of tasks on existing resources, so that completion time of tasks become minimized.

REFERENCES

[1] Ercolani and Giuseppe. 2013. Cloud Computing Services Potential Analysis. An integrated model for evaluating Software as a Service. Cloud Computing: 77-80.

[2] Chung and Hui-chun. 2013. Particle Swarm Optimization for Workflow Scheduling in Cloud Computing Environments.

[3] Fehling, Christoph, Frank Leymann, Ralph Retter, Walter Schupeck, and Peter Arbitter. 2014. Cloud Computing Patterns: Fundamentals to Design, Build, and Manage Cloud Applications. Springer Publishing Company, Incorporated.

[4] Garg, Saurabh Kumar, Steve Versteeg, and Rajkumar Buyya. 2013. A framework for ranking of cloud computing services. Future Generation Computer Systems 29, no. 4: 1012-1023.

[5] Hao, Liang, Gang Cui, Ming Cheng Qu and Wen De Ke. 2014. Resource Scheduling Algorithm of Cloud Computing for Energy Optimization Base on Virtual Machine Migration. In Advanced Materials Research, vol. 926, pp. 3232-3235.

[6] Javanmardi, Saeed, Mohammad Shojafar, Danilo Amendola, Nicola Cordeschi, Hongbo Liu and Ajith Abraham. 2014. Hybrid Job Scheduling Algorithm for Cloud Computing Environment. In Proceedings of the Fifth International Conference on Innovations in Bio-Inspired Computing and Applications IBICA 2014, Springer International Publishing, pp. 43-52.

[7] Li, Taoshen and Xixiang Zhang. 2014. On the Scheduling Algorithm for Adapting to Dynamic Changes of User Task in Cloud Computing Environment. International Journal of Grid & Distributed Computing 7, no. 3.

[8] Liang, Helan, Yanhua Du, and Sujian Li. 2013. An improved genetic algorithm for service selection under temporal constraints in cloud computing. In Web Information Systems Engineering–WISE 2013. Springer Berlin Heidelberg, pp. 309-318.

[9] Liu, Jing, X. Luo, X. Zhang, Fan Zhang and Bn Li. 2013. Job scheduling model for cloud computing based on multi-objective genetic algorithm. IJCSI International Journal of Computer Science Issues 10, no. 1: 134-139.

[10] Poola, Deepak, Saurabh Kumar Garg, Rajkumar Buyya, Yun Yang and Kotagiri Ramamohanarao. 2014. Robust scheduling of scientific workflows with deadline and budget constraints in clouds. In The 28th IEEE International Conference on Advanced Information Networking and Applications (AINA-2014), pp. 1-8.

 DOI : 10.35291/2454-9150.2020.0067