LOAD BALANCING IN CLOUD COMPUTING USING BIOGEOGRAPHY BASED OPTIMIZATION

Kamaljeet Mangat
Department of Computer Science & Engineering
PURCITM, Mohali
kamalmangat@yahoo.co.in

Guntaas Kaur
Department of Computer Science & Engineering
PURCITM, Mohali
guntaas.0510@gmail.com

ABSTRACT- Cloud computing depends on sharing of assets to accomplish intelligibility and economies of scale, like a utility over a network. In this various types of job scheduling and load balancing approaches has been used for this purpose. Various tasks on cloud have to be allocated to different resources for prevention of slow computation process at cloud. In the previous work various algorithms and optimization approaches have been used for the purpose of minimizing of make spans. The main purpose of minimization of make span is to compute all the tasks allocated, in less computation time and with less error. The genetic algorithm, Particle swarm optimization approaches has been used for the purpose of optimization of make spans of scheduled tasks. This approach provides not too good results as compare to particle swarm optimization approach. To overcome this limitation a nature based optimization approach BBO (Biography Based Optimization) is used for optimization of make spans developed by various algorithms.

Keywords: Cloud Computing, Load Balancing, BBO, PSO, GA

I. INTRODUCTION

Cloud computing depends on sharing of assets to accomplish intelligibility and economies of scale, like a utility (like the power matrix) over a network. At the establishment of cloud computing is the more extensive idea of met foundation and imparted administrations. Cloud computing, or in more straightforward shorthand simply "the cloud", likewise concentrates on expanding the viability of the imparted resources. Cloud resources are normally imparted by different clients as well as reallocated by every interest. This can work for allocating resources to clients. Case in point, a cloud PC office that serves European clients amid European business hours with a particular application (e.g., email) may reallocate the same assets to serve North American clients amid North America's business hours with an alternate application (e.g., a web server). This methodology is used to boost the utilization of computing power accordingly decreasing natural harm.

A. CHARACTERSTICS

Cloud computing has a number of characteristics that distinguishes it from other computing paradigms. These characteristics can be categorized as essential characteristics and common characteristic. The NIST has identified five essential characteristic fourteen and eight common characteristics of cloud computing. The essential characteristics are:

- **On Demand Self-Service**: allows for provisioning of computing resources automatically as needed.
- **Broad Network Access**: access to cloud resources is over the network using standard mechanisms provided through thin or thick clients in a heterogeneous manner. For example through Smartphone's, mobile phones and laptop computers.
- **Resource Pooling**: the vendors' resources are capable of being pooled to serve multiple clients using a multi-tenant model, with different physical and virtual resources in a dynamic way. The pooling and assigning of resources is done based on the changing needs of clients or consumers. Example of resources include; computation capabilities, storage and memory.
- **Rapid Elasticity**: allows for rapid capability provisioning, for quick scaling out and scaling in of capabilities. The capability available for provisioning to the client seems to be unlimited and that it can be purchased as demanded.
- **Measured Service**: allows monitoring, control and reporting of usage. It also allows for transparent between the provider and the client. In conjunction with the essential characteristics as identified by NIST, there are other cloud computing characteristics.

Deployment models

- **Private cloud:** Private cloud will be cloud framework worked singularly for a solitary association, whether managed inside or by an outsider, and facilitated either inside or remotely. Undertaking a private cloud project obliges a critical level and degree of engagement to virtualize the business environment, and requires the association to reconsider choices about existing assets. At the point when done right, it can enhance business, however every project in the undertaking raises security issues that must be addressed to prevent serious vulnerabilities Self-run server farms are by and large capital serious. They have a critical physical foot shaped impression, obliging assignments of space, equipment, and environmental controls. These resources must be revived occasionally, bringing about extra capital uses.
- **Public cloud:** A cloud is known as an "public cloud" when the administrations are rendered over a system that is open for open use. Open cloud administrations may be free. Saas is an extensive open cloud. For the most part, open cloud administration suppliers like Amazon AWS, Microsoft and Google own and work the base at their server farm and access is by and large through the Internet. AWS and Microsoft likewise offer direct interface administrations called "AWS Direct Connect" and "Purplish blue Express Route" individually, such associations clients to buy or lease a private connection with a peering point offered by the cloud provider.
- **Hybrid cloud:** Half breed cloud is an arrangement of two or more clouds (private, group or open) that stay particular elements however are bound together, offering the advantages of different sending models. Half and half cloud can likewise mean the capacity to associate collocation, managed and/or dedicated services with cloud resources.

II. LOAD BALANCING

load balancing distributes appropriates workloads over various processing resources, for example, PCs, a PC group, system joins, focal transforming units or circle drives. Load balancing aims to streamline resource utilization, maximize throughput, minimize reaction time, and keep away from over-load of any single resource. Utilizing multiple segments with burden adjusting rather than a solitary segment may build unwavering quality through repetition. Burden adjusting more often than not includes committed software or hardware, for example, a multilayer switch or a Domain Name System server process. Load balancing is a technique which varies from direct holding in that heap adjusting partitions movement between system interfaces on a system attachment (OSI model layer 4) premise, while channel holding infers a division of activity between physical interfaces at a lower level, either every bundle (OSI model Layer 3) or on an information join (OSI model Layer 2) premise.

Load balancing is the procedure which verifies that each processor inside the framework or each node in the system consume equivalent measure of power and give or take equivalent measure of work at any moment of time. The load can be distinguished as information transferring limit, CPU load or system delay.

Parameters

Computation Time: it is a total time required for all jobs.

Job Execution Time: time required by individual job or can be said that execution time per job.

Computation Cost: it is total cost which further depends upon the computation time and effort.

Make Span: it contains Gantt chart which is a type of bar chart which illustrates the project schedule.

III. BIOGEOGRAPHY-BASED OPTIMIZATION

This section refers to application of biogeography-based optimization in solving problems and surveying different parts of this algorithm. The basis of BBO algorithm is based on two main parts: Migration and Mutation.

- **Migration** involves two main processes immigration and emigration. Immigration and emigration are affected by various factors such as distance of an island to the nearest neighbour, size of the island, habitat suitability index (HSI) etc. HSI involves various factors such as rainfall, vegetation, climate etc. These factors favour the existence of species in a habitat. Habitats those are well suited for the residence of biological species will be having high HSI. A habitat with high HSI will be occupied

with large number of species, so will be having high emigration rate and low immigration rate (since the habitat is nearly saturated with species). Similarly a habitat with low HSI will be having small number of species. This idea is used in BBO for carrying out migration. In BBO, as in other optimization algorithms, initially a large number of candidate solutions are generated randomly for the problem under consideration. Associated with each solution there will be a HSI. Each solution generated is considered as a habitat. Each solution or habitat is a collection of suitability index variables (SIVs). Suitability index variables indicate the suitability of the habitat to which it belongs. High HSI habitat is analogous to good solution and low HSI habitat is analogous to poor solution. Through migration high HSI solutions share a lot of features with poor solutions and poor solutions can accept a lot of features from good solutions. Relationship between species count, immigration rate, and emigration rate is shown in the figure, where I refers to the maximum immigration rate, E is the maximum emigration rate, S0 is the equilibrium number of species and Smax is the maximum species count. The decision to modify each solution is taken based on the immigration rate of the solution.

- **Mutation:** Another important process in this optimization technique is mutation. Mutation is the sudden drastic change made to the HSI of any habitat due to certain cataclysmic events. Mutation increases the diversity among the population. Each candidate solution‘s’ is associated with a mutation probability. Sudden changes in climate of one habitat or other incidents will cause the sudden changes in HSI of that habitat. In BBO algorithm, this situation can be model in the form of sudden changes in value of SIV. Each member of one habitat has its own probability. If this probability is too low, then this solution has high chance to mutate. In the same manner, if probability of a solution is high that solution has a little chance to mutate. Consequently, solutions with high HSI and low HSI have a little chance to development a better SIV in the next iteration. Unlike high HSI and low HSI solutions, medium HSI solutions have a greater chance to development better solutions after mutation procedure.

Features of BBO:

- An efficient algorithm for optimization.
- Doesn’t take unnecessary computational time.
- Good in exploiting the solutions.
- Solutions doesn’t die at the end of each generation like other optimization algorithms.

.

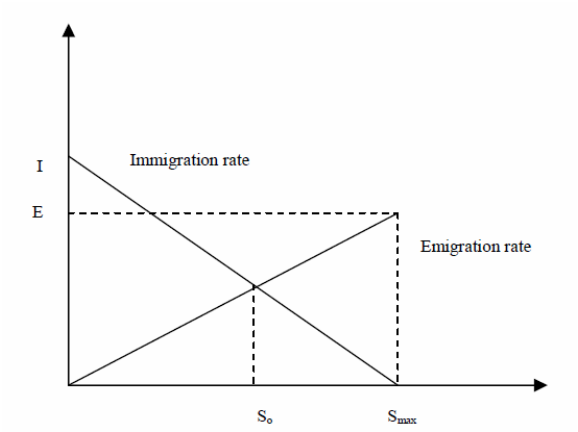
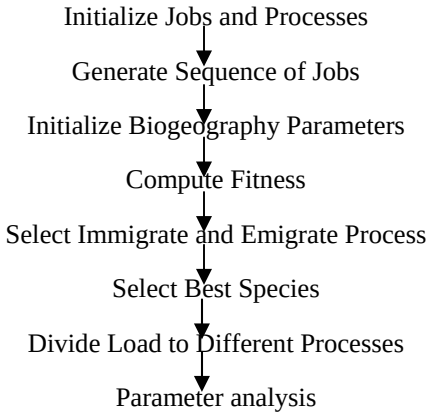


Fig 3.1 Relation between Immigration & Emigration

M(s) = Mmax (1-Ps)/Pmax
where Mmax is a user-defined parameter, Ps is the species count of the habitat; Pmax is the maximum species count. Mutation is carried out based on the mutation probability of each habitat by replacing an SIV from the habitat with another randomly generated SIV.
Immigration rate Ri can be defined as
Ri = I (1-F(s)/n)
Emigration rate Re can be defined as
Re = E (F(s)/n)

Flow of Work:



IV. PROBLEM FORMULATION

Load balancing in cloud computing is main tasking that has to be managed for optimum utilization of the cloud environment. In this various types of job scheduling and load balancing approaches has been used for this purpose. Various tasks on cloud have to be allocated to different resources for prevention of slow computation process at cloud. In the previous work various algorithms and optimization approaches has been used for the purpose of minimizing of makespans. The main purpose of minimization of make span is to compute all the tasks allocated, in less computation time and with less error. The genetic algorithm, Particle swarm optimization approaches has been used for the purpose of optimization of makespans of scheduled tasks. This approach provides not too good results as compare to particle swarm optimization approach. To overcome this limitation a nature based optimization approach BBO (Biography Based Optimization) is used for optimization of makespans developed by various algorithms.

V. METHODOLOGY

In the Purposed work various phases have to be used for the development of the load balancing system in the cloud computing environment. These different phases have to be done for the completion of purposed work.

- Load balancing has been done by using dividing different tasks into number of jobs so that they can be allocated to different resources for processing to complete in less computation time.
- In cloud computing scenario no. of tasks has to be assigned on various processes to handle load on the cloud. These tasks have been divided into sets and the dependency checking is done for prevention of dead lock state or to prevent demand of various extra resource allocations.
- Makespan has been developed on the basis of the allocation. This makespan has to be optimized Biography based optimization that performs the work on the basis principle of biography base means to say that environment conditions available in the scenario. These conditions are no of processor, priorities of tasks and no. of tasks available.

VI. RESULTS AND DISCUSSION

Table 1 and Table 2 represent number of jobs and number of virtual machines corresponding to the system response time in micro second for the genetic algorithm and BBO.

Number of Jobs	Virtual Machines	System Response Time (micro-sec)	
		GA	BBO
100	70	7.14	5.56
150	70	18.89	15.64
200	70	18.40	16.56
500	70	30.15	27.17

Table 1. System response time for 70 virtual machines and jobs

Number of Jobs	Virtual Machines	System Response Time (micro-sec)	
		GA	BBO
100	100	14.77	12.71
150	100	16.58	14.32
200	100	20.09	15.97
500	100	28.18	33.82

Table 2. System response time for 100 virtual machines and jobs

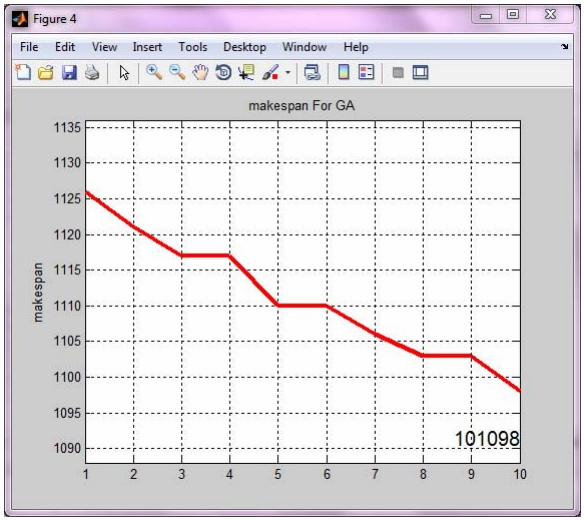


Fig. 6.1 Makespan for 100 jobs and 70 virtual machines

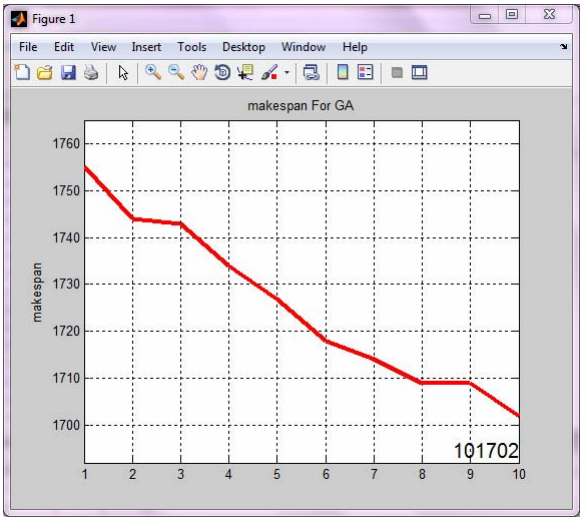


Fig. 6.3 Makespan for 200 jobs and 70 virtual machines

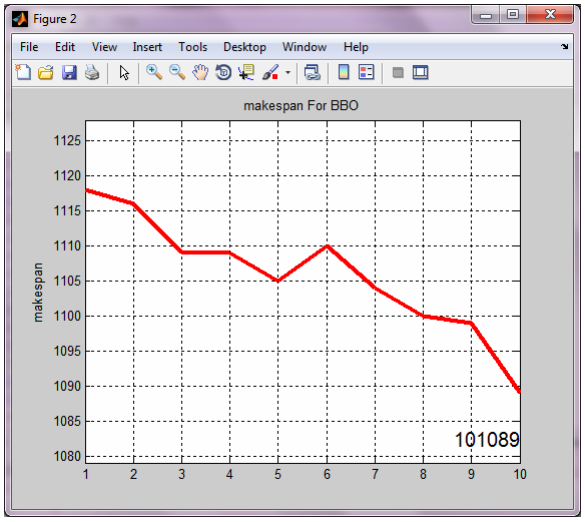


Fig. 6.2 Makespan for 100 jobs and 70 virtual machines

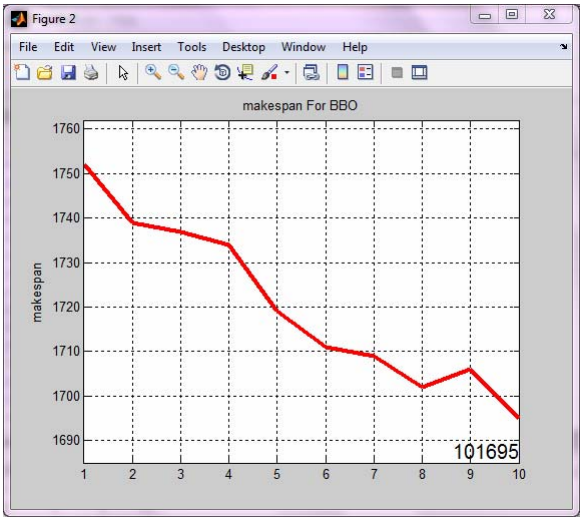


Fig. 6.4 Makespan for 200 jobs and 70 virtual machines

VII. CONCLUSION AND FUTURE SCOPE

Cloud computing is a model for enabling ubiquitous network access to a shared pool of configurable computing resources. Load balancing in cloud computing is main tasking that has to be managed for optimum utilization of the cloud environment. In this various types of job scheduling and load balancing approaches has been used for this purpose. The genetic algorithm approach has been used for the purpose of optimization of make spans of scheduled tasks. This approach provides not too good results as compare to particle swarm optimization approach. To overcome this limitation a nature based optimization approach BBO (Biography Based Optimization) is used for optimization of make spans developed by various algorithms. We got various types of parameters & on the basis of these parameters we conclude that our system gives us better results.

The interest of scientific community in BBO has risen sharply in recent years. The exploration capability of BBO makes it attractive for solving many complex problems in various fields. BBO has been proved to be very efficient in solving many NP hard problems, that is the problems for which event the best known algorithms have exponential time complexity. BBO is also applicable to multi objective optimization and constrained optimization. Extensive evaluations on various BBO algorithms have proved their efficiency in various fields. The BBO algorithm can be explored further by incorporating more features from the theory of island biogeography. Various approaches from other optimization algorithms can also improve BBO.

ACKNOWLEDGEMENT

I would like to begin with expressing my gratefulness to my supervisor, Mrs. Kamaljeet Mangat, who has provided me with her invaluable academic guidance and support during this research. I really appreciate all the valuable time she has spent in helping me and discuss various problems on the way.

REFERENCES

[1] Suraj Pandey, Linlin Wu, Siddeswara Mayura Guru, Rajkumar Buyya, "A Particle Swarm Optimization-based Heuristic for Scheduling Workflow Applications in Cloud Computing Environments" in 24th IEEE International Conference on Advanced Information Networking and Applications, DOI 10.1109/AINA.2010.31.

[2] Ratan Mishra and Anant Jaiswal, "Ant colony Optimization: A Solution of Load balancing in Cloud", in (IJWes T) Vol.3, No.2, April 2012.

[3] Ammu P K , Sivakumar K C, Rejimoan R, "Biogeography-Based Optimization - A Survey', International Journal of Electronics and Computer Science Engineering, 2012, vol. 2, issue 1, pp. 154-160.

[4] Wang Meihong, Zeng Wenhua, "A comparison of four popular heuristics for task scheduling problem in computational grid" in Proceedings of the 7th IEEE Symposium on Advances in Parallel and Distributed Systems, 2013, West Lafayette, pp.330-335.

[5] Jianfeng Zhao and Hongze Qiu, "Genetic Algorithm and Ant Colony Algorithm Based Energy-Efficient Task Scheduling", in Third International Conference on Information Science and Technology March 2013, pp 23-25.

[6] Yongqiang Gao , Haibing Guan , Zhengwei Qi , Yang Hou , Liang Liu , "A multi objective ant colony system algorithm for virtual machine placement in cloud computing" in Journal of Computer and System Sciences 79 (2013), pp 1230–1242.

[7] R. Angel Preethima , Margret Johnson, "Survey on Optimization Techniques for Task Scheduling in Cloud Environment", in IJARCSSE, Volume 3, Issue 12, December 2013.

[8] Tingting Wang, ZhaobinLiu , Yi Chen, Yujie Xu, Xiaoming Dai, "Load Balancing Task Scheduling based on Genetic Algorithm in Cloud Computing", in 12th International Conference on Dependable, Autonomic and Secure Computing, 2014 IEEE.

[9] Uma Somani, "Implementing Digital Signature with RSA Encryption Algorithm to Enhance the Data Security of Cloud in Cloud Computing," 2010 1st International Conference on Parallel, Distributed and Grid Computing (PDGC - 2010).

[10] RuWei Huang, Si Yu, Wei Zhuang and Xiao Lin Gui, "Design of Privacy-Preserving Cloud Storage Framework" 2010 Ninth International Conference on Grid and Cloud Computing.

[11] RuWei Huang, Si Yu, Wei Zhuang and Xiao Lin Gui, "Research on Privacy-Preserving Cloud Storage Framework Supporting Cipher text

Retrieval" 2011 Ninth International Conference on Grid and Cloud Computing.

[12] Farzad Sabahi, "Cloud Computing Security Threats and Responses," IEEE Trans. on Cloud Computing., vol. 11, no. 6, pp. 670 { 684, 2002.}

[13] Cong Wang, Qian Wang, Kui Ren and Wenjing Lou "Ensuring Data Storage Security in Cloud Computing." IEEE 2009.

[14] Mehmet Yildiz, Jemal Abawajy, Tuncay Ercan and Andrew Bernoth "A Layered Security Approach for Cloud Computing Infrastructure" 2009 10th International Symposium on Pervasive Systems, Algorithms, and Networks. © 2009 IEEE.

[15] Jianfeng Yang, Zhibin Chen "Cloud Computing Research and Security Issues" Vol. 978-1-4244-5392-4/10/$26.00 ©2010 IEEE.

Author Profile

Kamaljeet Mangat- Assistant professor in PURCITM, Mohali and currently pursuing PhD from Punjabi University, Patiala.

Guntaas Kaur- Currently pursuing M.Tech in computer science from PURCITM, Mohali.