International Journal of Advanced Research in

Computer Science Engineering and Information Technology

Volume: 4 Issue: 2 02-Jan-2015,ISSN_NO: 2321-3337

RESOURCE ALLOCATION AND JOB SCHEDULING USING GENETIC ALGORITHM IN CLOUD COMPUTING ENVIRONMENT

[1]**K.Deva Prasad**, [2]**V.Samatha,** [3]**Dr.G.Sambasiva Rao**
Research Scholar, Dravidian University, Kuppam[1, 2].
Research Supervisor, Dravidian University, Kuppam[3]

ABSTRACT*—Cloud computing is an internet computing, which share resources like software, storage, data and service to computers and other devices on demand. Cloud computing is a new model for distributed computing and it is said to be the product for evolution of calculation. The technology of computing becomes widely used due to more and more researchers and applications on cloud computing. Cloud computing has a vast user group and it also deal with a large number of tasks. The main issue in cloud computation is to make a right decisions when allocating hardware resources to the tasks and also when dispatching the computing tasks to resource pool. This paper is based on the situation arises during resource allocation and job scheduling under cloud circumstance. To improve the performance some methods have been suggested with the help of dynamic resource allocation strategy based on the dynamic resource assignment and law of failure, on the basis of genetic algorithm for resource allocation, improved job scheduling and optimized genetic algorithm with dual fitness.*

Key words: Cloud Computing, Resource Allocation, Job Scheduling, Intrusion Detection, Genetic Algorithm.

1. INTRODUCTION

The Information Technology (IT) industry and academia has put forward cloud computing model [1], which achieves generalization and commercialization of previous models in some sense. Cloud computing has the potential to transform IT industry to makes software more attractive as a service and it also shapes the IT hardwares. [2]. Cloud computing can be defined as a parallel and distributed system. It consists of a collection of interconnected and virtualized computers that are dynamically provisioned and presented as one or more unified computing resources. It is based on service-level agreements established through negotiation between the service providers and consumers [1]. With the advancements 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.

The cloud computing is still at its infant stage and a very new technology for enterprises. Cloud computing in term used to describe both a platform and types of application. As a platform its supplies, configured and reconfigured server's, where servers can be physical machine or virtual machine. On the other hand, cloud computing describes application that are extended to be accessible through the internet and for this large data centers and powerful servers are used to host the web application and web services. There are some important points in the definition to be discussed regarding cloud computing. Cloud computing differs from tradition computing paradigm as it is scalable, can been capsulated as an abstract entity which provides different level of services to the clients, driven by economies of scale and the service are dynamically configurable.

As a very new technology for enterprises there are many benefits stated of cloud computing by different researchers which make it more preferable to be adopted by enterprises. Cloud computing infrastructure allows achieving more efficient use of their IT hardware and software investments. This is achieved by breaking down the physical barrier inherent in isolated systems, automating the management of the group of the systems as an entity. Cloud computing can also be describe as the ultimately virtualized system and a natural evolution for data centers which offer automated systems management.

1.1 Genetic Algorithm

Genetic Algorithm (GA) is based on biological concept of generation of the population, a rapid growing area of Artificial intelligence. GA's are inspired by Darwin's theory about evolution. According to the Darwin "Survival of the fittest", is used as the method of scheduling in which the tasks are assigned resources according to schedules in context of scheduling, which tells about which resource is to be assigned to which task. GA is based on the biological concept of population generation [3].

Initial Population

Initial population is the set of all the individuals that are used in the GA to find out the optimal solution. Every solution in the population is called as an individual. And every individual is represented as a chromosome for making it suitable for the genetic operations. From the initial population the individuals are selected and some operations are applied on those to form the next generation. The mating chromosomes are selected based on some specific criteria.

Fitness Function

The productivity of any individual depends on the fitness value. It is the measure of the superiority of an individual in the population. The fitness value shows the performance of an individual in the population. If the large fitness value, then the performance of an individual is better. Depending on the fitness or function value, the individuals survive or die out. Hence, the fitness function is the motivating factor in the GA.

Selection

Selection mechanism is used to select an intermediate solution for the next generation based on the survival of the Darwin's law. This operation is the guiding channel for G A based on the performance. There are various selection strategies to select the best chromosomes e.g. roulette wheel, Boltzmann strategy, tournament selection, selection based on rank, etc.

Crossover

Crossover/hybridizing operation can be achieved by selecting two parent individuals and then creating a new individual tree by alternating and reforming the parts of those parents. Hybridization operation is a guiding process in GA and it boosts the searching mechanism.

Mutation

After crossover mutation takes place. It is the genetic operator that introduces genetic diversity in the population. Mutation takes place whenever the population tends to become homogeneous due to repeated use of reproduction and crossover operator. It occurs during evolution according to a user-defined mutation probability, usually set to fairly low. Mutation alters one or more gene values in chromosome from its initial state. This can produce the new gene values which can be added to the gene pool. With this new gene values, the genetic algorithm may be able to produce the better solution than was previously.

2. LITERATURE SURVEY

In the year 1960, John Holland proposed an algorithm called Genetic Algorithm (GA). He studied the phenomenon of adaptation which occurs in nature and imported into computer systems. The two notable features of GA [4] are parallelism and global solution search. Map/reduce model in cloud [5], researches added fitness function to improve the GA in order to reduce the total running time and average time of task execution called dual fitness genetic algorithm (DFGA). Comparison of improved and standard genetic algorithms in cloud environment is shown in figure 1.

Chromosome coding has many ways namely direct coding and indirect coding. When the task execution status code is directly used then it is named as direct coding and if it uses resources-task indirect coding

method then it is named as indirect coding. DFGA algorithm uses the indirect encoding method where the length of chromosome is the number of sub tasks and resource number is the value of each gene on the chromosome which is allocated to the sub task.

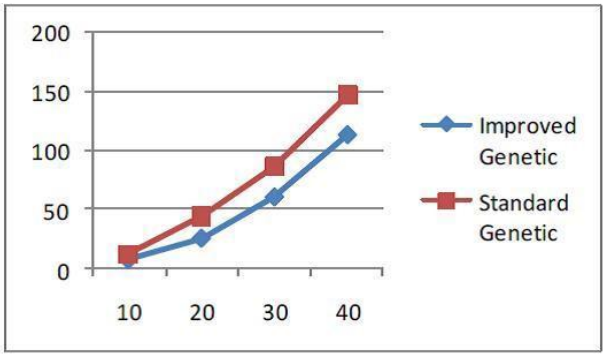

Figure 1. Comparison of Improved and standard genetic algorithms in cloud environment

Initialization is the first step in GA used to generate a SCALE number of chromosome, which has a length (M), and the value range of gene is a random number in [1, WORKER]. Here WORKER is the number of resources and M stands for the total number of sub task. The second step is to calculate two fitness functions they are, total time and average time of job running on all virtual machines. Both the fitness function values should be short. The third step is the selection operator which uses Roulette method.

Li's paper [4] depicts the comparison between DFGA and Adaptive Genetic Algorithm (AGA) using Gridsim simulator under a local environment of cloud computing. The result shows that the average time and the total time of tasks execution using DFGA tasks are significantly superior to AGA after many generations of evolution. Thus DFGA is an effective job scheduling algorithm.

3. PROBLEM STATEMENT

This work mainly emphasis about genetic algorithm approach to resource-constrained scheduling by applying a direct, time-based representation .This document describes a genetic algorithm for finding optimal solutions to dynamic resource constrained scheduling problems. The genetic algorithm approach was applied to more than 1000 small job shop and project scheduling problems (3-10 resource types ,10-300 activities).The algorithm performs well on a spacious variety of problems. Genetic algorithms operate on a population of solutions and employ heuristics such as selection, crossover, and mutation to evolve better solutions [6]. This algorithm is very beneficial also. It is used to generate practicable solutions to

optimization and search problems. The genetic algorithm performed well on some problems that were very difficult for the branch and bound techniques (i.e. the branch and bound method took a long time to and the optimal solution). The genetic algorithm did not perform well in some problems where the resources are tightly constrained. This comes as little surprise since there is presentation that applies the genetic algorithm to search for resource- feasibility, and tightly constrained resources mean fewer resource-feasible solutions the genetic algorithm will not work well on job shop problem. GA are used in various applications they are

- Mathematics
- Physics
- Chemistry
- Phylogenetic
- Bioinformatics
- Engineering
- Computational science
- Economics
- Manufacturing

Two basic metrics were collected to compare the performance of the algorithms:

Average Response Time: This the sum of the wait time plus the execution time of a job, averaged over all jobs. This metric captures the ability of the scheduling algorithm to service the smaller, shorter jobs. Since the job stream contains a large number of smaller short jobs, the metric provides a good measure of the quality of the scheduling algorithm.

Weighted Average Response Time: The weight of a job is the product of its resource requirement and execution time. The weighted average response time is then the product of the job weight and the job response time, averaged over all jobs. This is basically a measure of how well the scheduling system provides progress to the large jobs. It is also a measure of overall resource utilization over time.

4. PROPOSED APPROACH

Scheduling by employing GA is proposed as a solution for Multi-objective optimization for virtual resources. When one request is made for any resource then the virtual resources scheduling is mapped onto physical resources with proper load balancing which is very complex to achieve. This algorithm is in comparison with rank, random and static algorithm. The layer of virtualization occurs between users and physical layer and it has three characteristics usability, safety and moving. They come from independency of virtualization. The virtual resources are abstracted by making number of instances of actual physical

resource nodes with attributes. This algorithm is considered to be heuristic so it contains object functions, code and searching method. Object functions are there for load balancing calculation. NDSA II includes firstly non dominated sorting (set finding the lowest value of object function) and then crowd degree (lower). The GA has selection, crossover and mutation. In this algorithm the selection is tournament selection, crossover is two point crossovers and in mutation if the random number being chosen is the original gene is replaced by randomly generated one. By taking into consideration the CPU usage, memory and bandwidth the NDSA II comes out to be a better algorithm then rank, random and static algorithm as it provides many choices by running just once efficiently [7].

5. METHODOLOGY

Today in the fast growing era of computing and supporting software for parallelization, buying latest configuration PC or high performance system will be obsolete in a year or two. Further, it is not possible to invest in all latest computing infrastructures and software's. Therefore, cloud computing (use of computing resources such as hardware and software, which are delivered as a service over a network) is a good solution for on demand computing power (PaaS- Platform as a Service) and on demand software support (SaaS – Software as a Service).

The availability of computing infrastructure and software on cloud is quite easily available. Information about the GA based security to the Cloud computing environment shown in Table 1. Due to computational intensity and parallel nature of PGAs, it is hard to implement and optimize the performance on cloud.

Developing PGA on cloud will require knowledge of GA as well as of cloud. Many issues related to cloud like security, bandwidth etc. are still under discussion. Zhao et al., (2011) implemented PGA on prototype of cloud called Hadoop. Hadoop is an open source Cloud computing. The Function optimization using PGA on Cloud Computing gives speed-up but with short communication delay and it is suitable to solve large dimension problems. Comparing traditional PGAs on cloud with - HPC, cluster, Grid - PGA on cloud is simple, easy to implement and easy to extend to solve large-scale problems [8]. PGA on cloud will not be effective because of its distributed environment, as performance depends on network delay (Internet bandwidth). To solve the computational intensive optimization problems without hiring the computing infrastructure, cloud is good option. Some of optimization problems which are solved using PGAs are Resource Scheduling [9], Scheduling HPC Applications [10], Task Scheduling[11], [12], [13], Performance Improvement of Cloud Storage [14], Power Management in Cloud [15], Clustering composite SaaS components [16] etc.

Table.1: Information about the GA based security to the Cloud computing environment

S.no	Authors	Year And Reference	Technique	Performance
1.	C. Wang, Q. Wang, K. Ren, and W. Lou	2010.	Privacy-Preserving Public Auditing for Storage Security in Cloud	Storage Security in Cloud Computing
2.	P. Mell and T. Grance,	2009	Draft NIST Working Definition of Cloud Computing	Draft NIST Working Definition of Cloud Computing
3.	M. Armbrust, A. Fox, R.Griffith, A.D. Joseph, R.H. Katz, A. Konwinski, G. Lee, D.A. Patterson, A.	2009	Above the Clouds: A Berkeley View of Cloud Computing	Technical Report UCB-EECS-2009-28,Univ. of California, Berkeley, Feb.2009
4.	www.cloudsecurityallianc e.org	2010	Top Threats to Cloud Computing	Cloud Security Alliance
5.	M. Arrington	2006	Gmail Disaster: Reports of Mass Email Deletions	Disaster reports of- mass-email-deletions
6.	J. Kincaid	2008	"MediaMax/TheLin kup Closes Its Doors."	Media max the linkup-close sits- doors
7.	Amazon.com	2008.	Amazon s3 Availability Event	Cloud Security
8.	Q. Wang, C. Wang, K. Ren, W. Lou, and J. Li	2011	Enabling Public Auditability and Data Dynamics for Storage Security in Cloud	Enabling Public Auditability and Data Dynamics for Storage Security in Cloud Computing
9.	G. Ateniese, R. Burns, R. Curtmola, J. Herring, L. Kissner, Z.Peterson, and	2007	Provable Data Possession at Untrusted Stores	Cloud Security - Provable Data Possession at Untrusted Stores
10.	M.A. Shah, R. Swaminathan, and M.	2008	Privacy-Preserving Audit and Extraction of Digital	Cryptology ePrint Archive, Report
11.	A. Juels and J. Burton, S. Kaliski.	2007.	PORs: Proofs of Retrievability for Large	Proofs of Retrievability
12.	Cloud Security Alliance	2009	Security Guidance for Critical Areas of Focus in Cloud Computing	Security Guidance for Critical Areas of Focus in Cloud Computing

13.	H. Shacham and B. Waters,	2008	"Compact Proofs of Retrievability," Proc. Int'l Conf. Theory and Application of Cryptology and	Information Security: Advances in Cryptology
14.	C. Wang, K. Ren, W. Lou, and J. Li,	2010	Towards Publicly Auditable Secure Cloud Data Storage Services	Secure Cloud Data Storage Services
15.	M.A. Shah, M. Baker, J.C. Mogul, and R.	2007	Auditing to Keep Online Storage Services Honest	Online Storage Services
16.	R. Curtmola, O. Khan, and R. Burns,	2008	Robust Remote Data Checking," Proc. Fourth ACM Int'l Workshop	Storage Security and Survivability
17.	K.D. Bowers, A. Juels, and A. Oprea,	2009	Proofs of Retrievability: Theory and	Cloud Computing Security
18.	D. Boneh, B. Lynn, and H.Shacham,	2004	Short Signatures from the Weil Pairing	Information Security: Advances in Cryptology
19.	A.L. Ferrara, M. Green, S. Hohenberger, and M.	2009	Practical Short Signature Batch Verification	Cryptographers' Track at the RSA
20.	G. Ateniese, R.D. Pietro, L.V. Mancini, and G.	2008	Scalable and Efficient Provable Data Possession	Security and Privacy in Comm. Networks
21.	C. Wang, Q. Wang, K. Ren, and W. Lou,	2012	Towards Secure and Dependable Storage Services in Cloud	Service Computing
22.	C. Erway, A. Kupcu, C. Tamassia,	2009	Dynamic Provable Proc Data Possession	Computer and Comm. Security
23.	R.C. Merkle	1980	Protocols for Public Key Cryptosystems	Security and Privacy
24.	G. Ateniese, S. Kamara, andJ. Katz	2009	Proofs of Storage from Homomorphic Identification Protocols	Application of Cryptology and Information Security: Advances in Cryptology
25.	M. Bellare and G. Neven	2006	Multi-Signatures in the Plain Public- Key Model and a General Forking	Computer and Comm. Security
26.	Amazon.com	2009	Amazon Elastic Compute Cloud	Elastic Compute Cloud
27.	Y. Zhu, H. Wang, Z. Hu, G.-J. Ahn, H. Hu, and S.	2010.	Efficient Provable Data Possession for Hybrid	Hybrid Clouds

28.	Y. Dodis, S.P. Vadhan, and D. Wichs,	2009	Proofs of Retrievability via Hardness	Theory of Cryptography (TCC)
29.	F. Sebe, J. Domingo-Ferrer, A. Martı´nez-Balleste, Y. Deswarte, and	2008	Efficient Remote Data Possession Checking in Critical Information	Remote Data Possession Checking
30.	T. Schwarz and E.L. Miller	2006	Store, Forget, and Check: Using Algebraic Signatures to Check	Distributed Computing Systems
31.	R. Curtmola, O. Khan, R.Burns, and G. Ateniese,	2008	MR-PDP: Multiple-Replica Provable Data	Distributed Computing Systems
32.	K.D. Bowers, A. Juels, and A. Oprea		HAIL: A High-Availability and Integrity	Computer and Comm. Security

6. CONCLUSIONS

This paper analyzes the issues that occur in job scheduling and resource allocation under cloud environment and it also describes interrelated solution which is proposed by the researchers. The improvement has been made in the existing technique, which improve the efficiency in the usage of resources in cloud. In the current situation, the application of cloud computing becomes wider therefore the resource allocation and job scheduling algorithm will be improved to a variety of particular application environments.

REFERENCE

[1]. Foster, Y Zhao, I. Raicu, and S. Lu, “Cloud Computing and Grid Computing 360-degreecompared[C]”, inGridComputing Environments Workshop, 2008, pp. 1-10.

[2]. Michael Arm brust, Armando Fox, Rean Griffith, Anthony D. Joseph, Randy H. Katz, Andrew Konwinski, Gunho Lee, David A. Patterson, Ariel Rabkin, Ion Stoica, Matei Zaharia, “Above the Clouds: A Berkeley View of Cloud Computing”, Technical Report No. UCB/EECS-2009-28, 2009.

[3]. Rajkumar Buyya, Rajiv Ranjan, Rodrigo N. Calheiros,“Modeling and Simulation of Scalable CloudComputingEnvironments and the CloudSim Toolkit: Challenges andOpportunities”, in The 2009 International Conference onHigh Performance Computing and Simulation, HPCS2009, pp:1-11.

[4]. Li Jianfeng, Peng Jian. Task scheduling algorithm based on improved genetic algorithm in cloud computing environment[J]. Journal of Computer Applications, 2011, 1 (31): 184~186.

[5]. DEAN J, GHEMAWATS. MapReduce simplified data processing on large clusters[C]. Proceedings of the 6th Symposium on Operating System Design and Implementation. New York ACM, 2004, pp137~150.

[6]. Philippe Baptiste, Marek Chrobak ,Christoph D¨urr1,“ Polynomial Time Algorithms for Minimum Energy Scheduling”, ieee 2010

[7]. Chun-Wei Tsai, and Joel J. P. C. Rodrigues, “Metaheuristic Scheduling for Cloud: A Survey”, IEEE Systems Journal, 10.1109/JSYST.2013.2256731.

[8]. J. F. Zhao, W. H. Zeng, G. M. Li and M. Liu, “Simple Parallel Genetic Algorithm Using Cloud Computing”, Applied Mechanics and Materials, Vol. 121–126, pp. 4151-4155, 2011.

[9]. Z. Zheng, R. Wang, H. Zhong and X. Zhang, “An Approach for Cloud Resource Scheduling Based on Parallel Genetic Algorithm”, Proceedings of 3rd International Conference on Computer Research and Development, Vol. 2, pp. 444–447, 2011.

[10]. Y. Kessaci, N. Melab and E. Talbi, “A Pareto-based GA for Scheduling HPC Applications on Distributed Cloud Infrastructures”, Proceedings of International conference on High Performance Computing and Simulation, pp. 456 – 462, 2011.

[11]. E. Maria Mocanu, M. Florea, M.I. Andreica and N. Ţăpuş, “Cloud Computing – Task Scheduling based on Genetic Algorithms”, Proceedings of IEEE International Systems Conference, pp. 1-6, 2012.

[12]. S. Kaur and A. Verma, “An Efficient Approach to Genetic Algorithm for Task Scheduling in Cloud Computing Environment”, International Journal of Information Technology and Computer Science, Vol. 4, No. 10, pp. 74-79, 2012.

[13]. Y. Chang-tian and Y. Jiong, “Energy-aware Genetic Algorithms for Task Scheduling in Cloud Computing”, Proceedings of Seventh ChinaGrid Annual Conference, pp. 43-48, 2012.

[14]. K. Jindarak and P. Uthayopas, “Performance Improvement of Cloud Storage using a Genetic Algorithm based Placement”, Proceedings of Eighth International Joint Conference on Computer Science and Software Engineering, pp. 54-57, 2011.

[15]. Z. Xiong, Z. Zhang, H. Kong and D. Zou, “Genetic Algorithm-based Power Management in Cloud Platform”, Proceedings of International Conference on Internet Technology and Applications, pp. 1-4, 2011.

[16]. M. Yusoh, Z. Izzah and T. Maolin, “Clustering composite SaaS components in cloud computing using a grouping genetic algorithm”, IEEE Congress on Evolutionary Computation, pp. 1-8, 2012.