ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 7, July 2015

New Approach to Schedule Bag of Tasks (BoTs) with Increase in Parallel Process

Shekhar H.Pingale, Prof.V.Mogal

Department of Computer Engineering, RMD Sinhgad School of Engineering, Savitribai Phule Pune University, India

Assistant Professor, Department of Computer Engineering, RMD Sinhgad School of Engineering, Savitribai Phule Pune University, India

ABSTRACT: In case of NP-complete problem, it is curious issue to schedule large scale parallel computing applications on heterogeneous systems like Hybrid cloud . End user want to meet Quality of Service requirement (QoS).To process huge number of Bag of-Task (BOT) concurrently in such environment with QOS is a Big problem. For that it needs a exact solution.It addresses the scheduling problem of large-scale applications inspired from real-world, characterized by a huge number of homogeneous and concurrent bags-of-tasks that are the main sources of bottlenecks but open great potential for optimization. Here method proposes Multiobjective scheduling algorithm to schedule BOTs.Algorithm optimizes objectives such as, Execution time where two constraints to consider are Network Bandwidth and Storage requirement .

NP Complete problem like scientific application which takes more time to find the results.Here, MOGTScheduling algorithm proposed in such a way that it optimize the schedule i.e computation time while dividing the Bag of task(BoTs) into Sub Bag of Task(SubBoTs)means it results in increase in parallel computation and find the result in limited number of steps.

KEYWORDS: Multi-objective scheduling , Bags-of-tasks , Hybrid clouds, NP Complete.

I. INTRODUCTION

Use of Multiobjective game theoretic Scheduling algorithm for execution of SubBoT takes less time as compared to BoT.Here it consider the objective as computation time while considering constraint as Storage and Bandwidth Increase in parallelism after dividing BoTs in SubBoTs gives faster updates which could be possible by Scheduling algorithm and depend on those update scheduler can take a decision in less time. In this scheduling ,it executes the tasks faster where the processor utilization will be more.

More will be the parallel computation more resource power it will need. For that it requires the hybrid cloud resources to have a better performance for NP Complete application. computing systems such as clouds have evolved over decade to support various types of scientific applications with dependable, consistent,pervasive and inexpensive access to geographically-distributed high-end computational capabilities. To program such a large and scalable infrastructure like hybrid clouds,loosely coupled-based coordination models of legacy software components such as bags-of-tasks (BoT) and workflows have emerged as one of the most successful programming paradigms in the scientific community.

II. RELATED WORK

Genetic algorithms (GAs) deliver strong search techniques that allow a best-quality clarification to be derived from a big search space in P-time. A GAs pools the exploitation of best solutions from past searches with the search of new regions of the solution space. An individual represent problem Solution of the search space. These algorithms preserves a population of individuals that changes over generations. Fitness function determine the quality of an individual in the population .[2]

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 7, July 2015

Most common way has been to analytically combine the multiple objectives into a single aggregate objective function by assigning weights to different objectives. Purpose of this research is to maximizing the performance and minimize the makespan.It Uses Multi-Objective Evolutionary Algorithms such as Multiobjective Genetic Algorithm (MOGA) and Multi objective Evolutionary Programming (MOEP). [3]

It propose a workflow planning method, which considers simultaneously optimizing multiple objectives. In this paper, it investigate the trade-off between two conflicting objectives execution time and cost-while meeting the users' maximum deadline and budget requirements. Paper work believe that workflow planning approach can be easily extended to support more objectives.[4]

PSO has become popular due to its simplicity and its effectiveness in wide range of application with less execution cost. Some of the real applications that have used PSO are: the data mining , chemical engineering , pattern recognition and environmental engineering . The PSO has also been applied to solve NP-Hard problems like Scheduling and task allocation.[5]

The fairness of allocation is an important factor in modern distributed systems and our scheme will be suitable for systems in which the fair treatment of the users' jobs is as important as other performance characteristics. It show that our cooperative load balancing scheme not only provides fairness but also provides a Pareto optimal operating point for the entire system.[6]

Contribution :

1)Here Multiobjective scheduling algorithm focuses on constraints such as execution time,Storage and Bandwidth requirement.It derives the minimum time of execution and allocates required bandwidth and storage,does not waste space and network bandwidth.

2)Existing algorithm divides task into BOTs and execute it in parallel.
3)Here we propose model which reduce execution time. It divides BOTs in SubBOTs which will improve execution performance and gives result in less time.

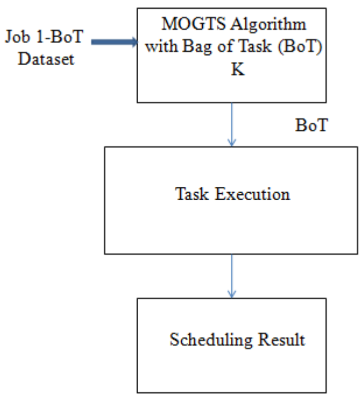

Figure 1: Existing System

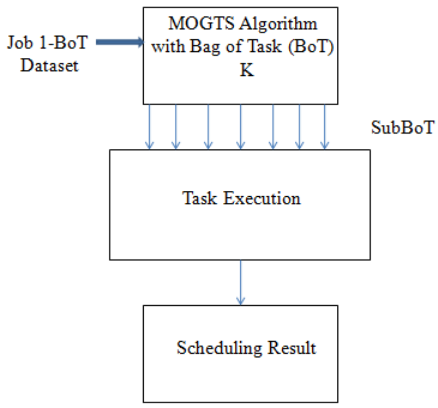

Figure 2: Proposed System

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 7, July 2015

III. PROPOSED ALGORITHM

Multiobjective Scheduling Algorithm for SubBoT:

This algorithm is proposed to work on NP complete problems given as a input. It focuses mainly on constraints as,Execution Time,Bandwidth and Storage. Due to generation of SubBoTs from BoTs the master will get results in less time which indirectly reduces execution time and also affects other factors with better results.

Mathematical Model

In Multi Objective Game Theory Scheduling for SubBoT,
Here it Propose a reliable System which gives a best solution.
It focuses mainly on four Constraints such as,

1. Execution Time
2. Network Bandwidth
3. Storage Requirement

Set Theory:
S=I,O
Where I is the Input contains,
$I=(AS,K,M,m_i,pk_i,\delta_k,\varepsilon)$
Require: b_i,br_k
Compute: $\delta^{s(l)},\Theta^{s(l)}$
Where O is the Output;
O=(Result of NP Complete Problem)

- n=Application,
- AS=Set of Applications,
- K =No.of BoTs,
- M =No.of Sites,
- m_i=No.of Processors on site $s_i,i\in[1,k]$
- pk_i (m x n):ETC matrix
- δ_k =number of tasks of BoT k($k\in[1,M]$)
- br_k =Bandwidth Requirement of BoT
- bl_i =bandwidth limit of site $s_i,(i\in[1,M])$
- T_k =($k\in[1,K]$) ; Bag of Task
- $\delta^{s(l)}$= Task Distribution Matrix
- $\Theta^{s(l)}$=Resource Allocation matrix

Ai=Makespan of application where Ai, $i\in[1,n]$ is maximum completion time of it's BoT
Here,
1]It states that, to Minimize Execution Time of all application F(x),
Minimize $F(x)=(f(x))$,
$s.t.h_i(x)\leq\lambda_{(x,i)},i\in[1,M]$
$g_i(x)\leq sl_i,i\in[1,M]$ where $x\in S$

2]Matrix which delivers the expected execution time P_{ki} of task in each BOT $k\in[1,K]$

$P_{ki}=pc_{ki},pc_{ki}\geq po_{ki};$

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 7, July 2015

$P_{ki}=pc_{ki}+(po_{ki}-pc_{ki})=po_{ki},pc_{ki}<po_{ki})$

3]Data bandwidth can be calculated as, $\lambda_{x,i}\leq \sum_{k=1}^{K}\theta_{ki}\cdot b_{ki}=\sum_{k=1}^{K}\frac{\theta_{ki}\cdot d_{ki}}{po_{ki}}$

4]Resource allocation of BOTs on S_i, $\theta_{ki}=m_i\cdot\frac{\delta_{ki}\cdot p_{ki}\cdot\omega_{ki}}{\sum_{k=1}^{K}\delta_{ki}\cdot p_{ki}\cdot\omega_{ki}}$

5]Makespan can be expressed as,aggregated execution time divided by number of processors; argmin ($\sum_{k=1}^{K}f_k(\Delta)$)

6]For initial step of matrix, BoT considers bandwidth and processors available,

$\delta_{ki}=\delta_k\cdot\frac{\frac{m_i}{p_{ki}}}{\sum_{i=1}^{M}\frac{m_i}{p_{ki}}}$ and $b_{ki}=\lambda_{x,i}\cdot\frac{d_{ki}}{\sum_{i=1}^{K}d_{ki}}$

7]After initial step this matrix will continue, $\Theta^{S(l)}=\Theta(\Delta^{S(l-1)})$, $\Delta^{S(l)}=\Delta(\Theta^{S(l)})$
8]Here,for each BoT which is scheduled and need to execute in parallel manner is get divided into SuBBoT as,
Number of Tasks in the Bag are K which now individually get executed on resource where,
Number of SubBoT K_{sub} = Number of Tasks K in Bag,$K_{sub}\in$ [1,K]

IV. PSEUDO CODE

Phase 1:

a) Initialize δ^0 and weight of the matrix ;$W\in AS$.
b) Iterate the loop and Do Scheduling of BOTs.
Divide the BOTs in SubBOTs.
Iteration Continues till the step limit.
Go for not scheduled BOTs ,repeat(b)
c)calculate $pw_{ki},sw_{ki},\delta_{ki}$ for each site s_i

Phase 2 :

Search the final task distribution $\Delta^{s(l)}$ and resource allocation $\Theta^{s(l)}$i.e. Scheduling And appropriate resource allocation for particular job.
$\Theta^{s(l)}$:Multiobjective Schedule(θ,δ,m,b,λ,p);

Phase 3:
Eliminate completed Bag of tasks, free up resources and Begin new game by iterating Phase 1 and Phase 2.
END.

IV. EXPERIMENTAL RESULTS

Here to analyze and find results of given NP Complete problem,use Hybrid Cloud environment.Each node is having different configuration.

To schedule NP Complete input we need better infrastructure i.e Hybrid Cloud.Such kind of problems take lot of iteration to compute and give approx results which is closer to actual.More is the iteration better will be the result.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 7, July 2015

Table 1:Platform Basis Comparison

	No. of Jobs	Existing System Time (msec)	Proposed System Time (msec)
Makespan on Cloud	2	854	1309
	4	22372	11179
	6	43031	21761
	8	64209	32330
	10	85305	42868
Makespan on Personal System			
	2	51877	8585
	4	154853	15538
	6	161395	50509
	8	199220	70240
	10	246354	106507

Personal System Based Evaluation:
Here the system configuration is 1TB Hard disk,8 GB RAM,Intel i3 (2 core)1.70 GHz processor;giving experimental result as,

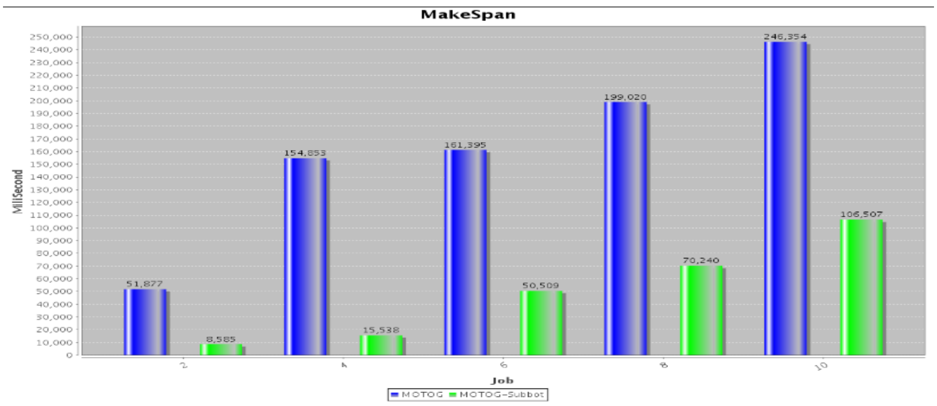

Figure 3:Makespan

Cloud Based Evaluation:
Here the IBM Cloud Configuration is Windows server 2012,32 GB RAM,12 core, Intel Xenon processor(3.45 GHz)

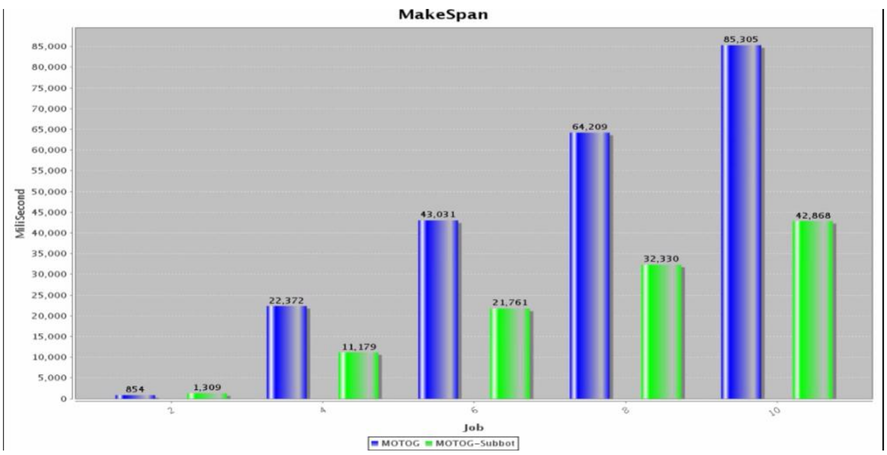

Figure 4:Makespan

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 7, July 2015

V. CONCLUSION AND FUTURE WORK

In system it uses Game Theoretic Scheduling algorithm which efficiently schedules the BoTs over a hybrid cloud and works for NP Complete problem also achieve good solution for it. It work on all the parameter as makespan , storage and bandwidth which gives aggregate better results, we can say exact solution for each BoTs to execute on specific site which is most suitable for them.

As a future scope we propose a system which divides BoTs into a sub-BoTs and we search for new appropriate weights for better performance.We can implement this model for Astrological problems like scientific study of Galaxy or Scientific problems in Physics which takes lot of time for computation.

REFERENCES

[1] Rubing Duan, Radu Prodan and Xiaorong Li, Member, IEEE Computer Society,Multi-Objective Game Theoretic Scheduling of Bag-of-Tasks Workflows on Hybrid Clouds, IEEE transaction on cloud computing, vol .2 , no. 1, January-March 2014.
[2] J. Yu and R. Buyya, "A Budget Constrained Scheduling of Workflow Applications on Utility Grids Using Genetic Algorithms," Proc. First Workshop on Workflows in Support of Large-Scale Science, June 2006.
[3] V.Vasudevan, A Reliable Schedule with Budget Constraints in Grid Computing International Journal of Computer Applications (0975 – 8887) Volume 64– No.3, February 2013
[4] H.M. Fard, R. Prodan, J.J.D. Barrionuevo, and T. Fahringer, "A Multi-Objective Approach for Workflow Scheduling in Heterogeneous Environments," Proc. 12th IEEE/ACM Int'l Symp. Cluster, Cloud and Grid Computing (CCGrid '12), 2012.
[5] S. Pandey, L. Wu, S.M. Guru, and R. Buyya, "A Particle Swarm Optimization-Based Heuristic for Scheduling Workflow Applications in Cloud Computing Environments," Proc. 24th IEEE Int'l Conf. Advanced Information Networking and Applications, pp. 400-407, 2010.
[6] L.F. Bittencourt and E.R.M. Madeira, "Towards the Scheduling of Multiple Workflows on Computational Grids," J. Grid Computing, vol. 8, no. 3, pp. 419-441, Sept. 2010.
[7] S. Penmatsa and A.T. Chronopoulos, "Cooperative Load Balancing for a Network of Heterogeneous Computers," Proc. 21st IEEE Int'l Parallel and Distributed Processing Symp. (IPDPS '06), Apr.2006.
[8] Y. Kwok, S. Song, and K. Hwang, "Selfish Grid Computing: Game-Theoretic Modeling and Nas Performance Results," Proc. Fifth IEEE/ACM Int'l Symp. Cluster, Cloud and Grid Computing (CCGrid '05), citeseer.ist.psu.edu/kwok05selfish.html, 2005.
[9] P. Ghosh, K. Basu, and S.K. Das, "A Game Theory-Based Pricing Strategy to Support Single/Multiclass Job Allocation Schemes for Bandwidth-Constrained Distributed Computing Systems," IEEE Trans. Parallel Distributed System, Mar.2007.
[10] Christopher Dabrowski, Reliability in grid computing System Concurrency computation practice experience, 2009.
[11] T. Eilam et al.,Using a utility computing framework to develop utility systems, IBM System Journal,43:97-120,2004
[12] Rajkumar Buyya,James Broberg,Andrzej Goscinski,"Cloud Computing-Principles and Paradigm"reference book,Wiley publication.

BIOGRAPHY

Shekhar Pingale Reseach Scholar at RMD Sinhgad School of Engineering, Savitribai Phule Pune University. He has received B.E. in Computer Engineering from Savitribai Phule Pune University, Pune. Currently he is pursuing M.E. in Computer Engineering from RMD Sinhgad School of Engineering, Savitribai Phule Pune University, Pune .

Prof. V. Mogal received the B.E. and M.E. Degrees in Computer engineering. He is working as Assistant Professor in Department of Computer Engineering, RMD Sinhgad School of Engineering Pune, India.