

# Deadline Constrained Task Scheduling with a Sequential Game in Cloud

**Sunaina Sangeet Thottan**

*PG Student*

*Department of Computer Science & Engineering  
Jawaharlal College of Engineering and Technology,  
Palakkad, Kerala*

**Shyji M B**

*Assistant Professor*

*Department of Computer Science & Engineering  
Jawaharlal College of Engineering and Technology,  
Palakkad, Kerala*

**Ambikadevi Amma T**

*Professor*

*Department of Computer Science & Engineering  
Jawaharlal College of Engineering and Technology, Palakkad, Kerala*

## Abstract

Developing and manipulating a virtual environment like Cloud Computing for processing tasks running on different applications in different systems is very much popular now a days. Task Scheduling from number of applications is becoming a complicated issue while dealing with this scenario. In most cases a single objective is calculated to schedule the task to the available sites with respect to the convenience of cloud provider. Cost, processing rate, resource utilization, performance, overhead storage and bandwidth are the most commonly selected objectives for scheduling heterogeneous tasks to the available virtual sites. This scenario generates a condition in which satisfaction of one objective leading to serious performance outflow in another objective. Introducing a scheduling technique considering multiple objectives a new method including a sequential game with number of stages for scheduling heterogeneous tasks running on different applications for the performance enhancement is formulated in this work. Allocation of tasks to available site is processed after computing the minimum cost and execution time requirements, considering storage and bandwidth utilization in sites within the deadline and grouping tasks in to different Bag of Tasks (BoT).

**Keywords: Deadline, Bag of Task (BoT), Heterogeneous Tasks, Multiple Objectives, Sequential Game, Task Scheduling**

## I. INTRODUCTION

Improved accessibility of resources with more flexibility and quick deployment facilities in cloud computing attracted big and small industries in last decades. Support for running applications simultaneously in number of system distributed over the network and co-ordination of these applications effectively turned out to be an important issue while dealing with cloud systems. The issue becomes much more complicated and worse when it came to heterogeneous application competing for available resources in cloud sites.

User perspective of cloud utilization is to minimize the cost of application and reduce the makespan. At the same time, efficient utilization of storage space and bandwidth is the necessity to enhance the system level efficiency. Many schemes evaluating a single objective for scheduling tasks in cloud sites were proposed, but there are only a very few methods available for scheduling heterogeneous application to different cloud sites considering multiple objectives. In dealing with single objective scheduling, satisfaction of one objective in most cases leads to dissatisfaction of other objectives which seriously affect the quality of services.

A new scheme were the managers of the heterogeneous application cooperate in a sequential game to schedule the given tasks inside the applications is introduced in this paper. Multiple objective computational tasks is improved such that user perspective and system perspective is utilized with efficiency. The homogeneous tasks in the heterogeneous applications are grouped for the better performance. Including the deadline constrain cloud users are given priority. The role of users in a cloud application is given higher priority with the scheduling strategy used considering the deadline in calculating multiple objective. Section II in the paper is the close evaluation of various schemes that are adapted for scheduling the resources to the required tasks. Section III explains the methodology proposed in the work, steps and process by which the task scheduling using deadline based multiple objective is scheduling strategy works is formulated. Result of the work and conclusion of the new method for scheduling in cloud is explained in section IV and section V respectively.

## II. RELATED WORK

Load Balancing is a serious issue in fields which require the use of most optimal resources by a large number of application. Competition among applications to get resources is most cases is satisfied by adopting many scheduling strategies. In case of

cloud environment the issue becomes much more complicated as each applications run in a virtual operating system with the resources distributed throughout in the network. There may arise cases that some applications require more number of CPU while some other may compete for storage space. Allocating tasks to virtual sites considering all the requirements satisfied is a hectic job. There are many task scheduling strategies and allocation mechanism proposed to do the job earlier, almost with their own limitations. Most common among them including static scheduling, dynamic scheduling, single objective scheduling and multiple objectives scheduling is studied in this section.

Scheduling tasks points to the assignment of task to a node for doing the required computation. Static Scheduling is the earliest scheduling strategy which assigns the task to node using perfected data. Many methods were introduced to schedule the task with static mechanism to reduce the cost, overhead and execution time. In most cases Service Level Agreements of cloud user to provider is considered for the purpose. As the basic idea of cloud computing ensures complete utilization of resource according to pay per use policies of providers and dynamic behavior of tasks, static scheduling came out to be a strategy with lot of performance issues. A system prototype Flextic is introduced that use static scheduling including time and cost constrain to schedule tasks. Based on the task duration estimate and assuming the time taken to complete each task is known an Abstract Refinement Scheduling became a breakthrough in the field. As a cloud environment is completely dynamic, static behavior require some or other way to compete with the dynamic nature of cloud computing.

Cloud provider, Resource Providers and Cloud Users dynamic interaction is essential for the performance enhancement in field of cloud computing. The optimal use of resourced can be achieved only if there is communication between the main participants in cloud. Every applications running over the network will have certain characteristics and strategies which should be satisfied for completion of processing the application efficiently. Dynamic scheduling strategies incorporates mechanism to periodically calculate the conditions which ensure quality of service. Reliability of cloud services are proposed in the approach TB scheduling, consideration of transportation cost of application the approach compute dynamic scheduling reliability in cloud. Optimum use of resources considering the current load and future load on the server is included in most of dynamic approaches. Dynamic behavior of the application processing algorithms points to performance improvement up to an extent but the time complexity issues are major concern in this situation.

Various scheduling approaches extended to improve the scheduling mechanism in cloud. Cost reduction is an important constraint which is required by user while interacting in a cloud environment. A resource management architecture which consider the objective as to minimize the cost is introduce in most single objective scheduling while a method with time optimization is suggested in Pandey's work in the context. Concern of single objective optimization is the dissatisfaction of many other objectives while focusing on meeting that particular objective in scheduling a task to the virtual sites in cloud. Many shortest path algorithms and job scheduling methods were introduced for the scheduling purpose in last decade.

As the single objective consideration was not complementing with the quality of service enhancement in all direction of scheduling, new approaches including complex objectives were introduced. In some of the method introduce including Paetro a series of objectives were considered for scheduling. Genetic Algorithms were also introduced with multiple objective to schedule the tasks. Many objective optimization developed new possibilities in cloud computing. As performance improvement and efficient scheduling of available resources are very essential in networked application multiple objective constrained scheduling became a necessity. Time, Cost, bandwidth, storage, path and processing rate became the objective to scheduling in many algorithms. While considering the set large scale application, now also there scheduling task is complicated issue. Satisfaction of some objective may lead to dissatisfaction of other objectives. The main concern is, only few systems are available that consider user perspective like deadline of a task in working out scheduling strategy.

### III. METHODOLOGY

Difficulty in scheduling applications which have different number of tasks is considered as the problem and a feasible solution with maximum performance enhancement is formulated with a sequential cooperation among the managers of each and every task taking up multiple objectives for assigning weight for each task.

#### A. Problem State:

There exist K number of applications with N number of tasks in which some are homogeneous and the problem is to allocate this tasks to Z number of sites available in the cloud system environment. Consideration of multiple objective including cost, execution time, storage availability in site and bandwidth utilization is considered giving priority to the user defined deadline to the completion of each task. More specifically, the problem state include Application={1,2,3.....K}, Site={1,2,3.....Z}, Homogeneous Tasks grouped as Bag of Tasks = {T<sub>1</sub>, T<sub>2</sub>,.....T<sub>n</sub>}

#### B. System Working:

As the problem formulates, the K applications with number of tasks are grouped as homogeneous bottle of tasks. To minimize the cost and execution time of these homogeneous tasks in cloud sites available the optimal solution is calculated for each of the objective.

For the fairness of the solution, minimization of cost and execution time is computed carefully. In cloud systems execution time calculation is a complicated task as it require communication between applications running on different sites. So execution time of a task is computed using the summation of communication time between task and completion time of each task. Execution time is denoted by  $P_{ki}$ .

Resource allocation is done in this system finalizing set of strategies agreed by each of the Bag of Task's managers. Once the site with optimum cost and time which is fulfilling the storage and bandwidth is found, the strategies are checked including the deadline. Performance, cost, storage, resource allocation and task distribution of each site is calculated using methods described below. Entire flow of the system working is illustrated in the fig 1.

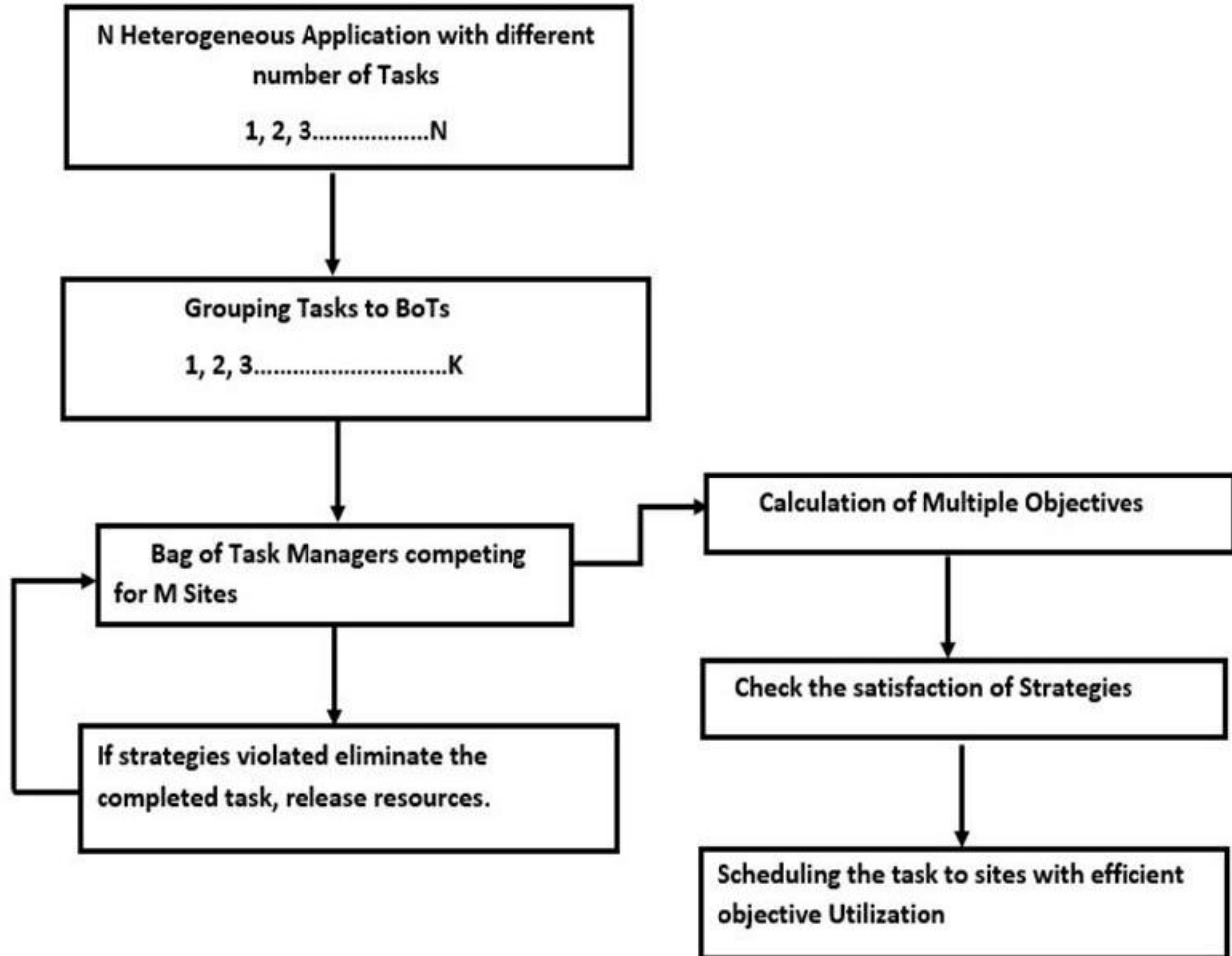


Fig. 1: Deadline Constrained Task Scheduling

Processor allocation is expressed as the product of number of processors in each site and ratio of performance, storage and cost weight of each site to the total weight of cost, storage and performance in BoT in all applications. Each of these weight calculation is done by taking the ratio in each site to summation of weight in entire application. If there is violation of any one strategy computation is repeated for optimal solution. If there is no violation, resource allocation and task distribution of each BoT in every site is checked and which meet the objectives is assigned to site. Each stage in the game is represented by the resource allocation until the violation of strategies. Whenever the strategies are violated, this procedure is repeated for new allocation. During each recalculation the task which is completed is eliminated and resources are released.

**C. Steps for Deadline constrained Multiple Object Resource Allocation:**

- 1) Step 1: Application with N tasks are given as input.
- 2) Step 2: Group similar tasks under same Bag of Tasks.
- 3) Step 3: Taking the values from input applications, compute the performance, storage, cost and task distribution on each site as described in above section.
- 4) Step 4: Calculating the resource allocation for each task in every site.
- 5) Step 5: Calculate the task distribution final matrix. Optimize the Weight and allocate site to tasks which gives most optimized solution.

- 6) Step 6: Check repeatedly for violation of set of strategies agreed by Bag of Task managers within user defined deadline.
  - If violated, eliminate the completed tasks and release resources and repeat the multi-objective calculation in Step 3 and do Step 4.
  - If not violated, wait for the entire task completion.

#### IV. RESULT

Simulation result of the deadline constrained sequential game scheduling shows an increase in performance for this work compared to other most commonly used single objective based scheduling. As the user preferred deadline is considered for allocation of resources grouped under Bag of Tasks, most accurate scheduling strategy as per the user requirement can be adapted by providers with this mechanism of scheduling while considering cloud environment. Performance enhancement is much more useful in the case of heterogeneous application as the grouping of similar tasks are done faster.

#### V. CONCLUSION

Deadline constrained task scheduling mechanism is analyzed and feasibility, resource utilization and user priority found to be well satisfied in the approach. As the multiple objective consideration of task scheduling is included in the work, performance improvement and integrity in task scheduling satisfying the cost and execution time with the optimal utilization of storage and bandwidth availability are satisfied. Including deadline for each task, user priority is highlighted in every means. Efficiency in scheduling is enhanced releasing the completed task and releasing the resources in every stage for recalculation of objectives. Each stage in the mechanism makes it closer to attain high performance in working on a cloud environment consisting of many virtual systems competing for resources to complete the task with in a deadline assigned to each task.

#### REFERENCES

- [1] Maria Alejandra Rodriguez and Rajkumar Buyya, "Deadline Based Resource Provisioning and Scheduling Algorithm for Scientific Workflows on Clouds", IEEE transactions on cloud computing, vol. 2, no. 2, April-June 2014.
- [2] Harshitha Menon, "A Distributed Dynamic Load Balancer for Iterative Applications", SC '13, November 17-21, Denver, Colorado, USA, 2013.
- [3] Laurie Young, Stephen Mc Gough and Steven Newhouse, John Darlington, "Scheduling Architecture and Algorithms within the ICENI Grid Middleware", London e-Science Centre, Imperial College London, South Kensington Campus, London, 2013 SW7 2AZ, UK.
- [4] Swachil Patel, Upendra Bhoi "Priority Based Job Scheduling Techniques In Cloud Computing: A Systematic Review", international journal of scientific & technology research volume 2, issue 11, November 2013.
- [5] Suraj Pandey, "A Particle Swarm Optimization-based Heuristic for Scheduling Workflow Applications in Cloud Computing Environments", ICT Centre Department of Computer Science and Software Engineering Hobart, Australia The University of Melbourne, Australia- 2012.
- [6] Millind Thames, "Game Theory for Security: A Real World Challenge Problem for Multiagent Systems and Beyond" Computer Science Department University of Southern California Los Angeles, CA 90089, 2011.
- [7] Rajkumar Buyya, "Scheduling Parallel Applications on Utility Grids: Time and Cost Tradeoff Management", Grid Computing and Distributed Systems Laboratory Department of Computer Science and Software Engineering The University of Melbourne Victoria 3010, Australia.
- [8] Srikumar Venugopal and Rajkumar Buyya, "A Deadline and Budget Constrained Scheduling Algorithm for eScience Applications on Data Grids", (GRIDS) Laboratory, Department of Computer Science and Software Engineering, The University of Melbourne, Australia. 2010.
- [9] Bibhudatta Sahoo, "A Genetic Algorithm Based Dynamic Load Balancing Scheme for Heterogeneous Distributed Systems", PDPTA 2008, Las Vegas, Nevada, USA, July 14-17, 2008, 2 Volumes.
- [10] Ismail Assyad, "A Bi-Criteria Scheduling Heuristic for Distributed Embedded Systems under nts", Hamoudi Kalla INRIA, 655 av. de l'Europe 3833 Saint-Ismier, Cedex - France, vol 2, 2008.
- [11] Mourad Hakem and Franck Butelle, "Reliability and Scheduling on Systems Subject to Failures", Mourad Hakem and Franck Butelle LIPN - CNRS UMR 7030 Universite' Paris Nord Av. J.B. Cle'ment 93430 Villetaneuse France, 2008.
- [12] Oscar H, "Heuristic Algorithms for Scheduling Independent Tasks on Nonidentical Processors" University of Minnesota, Minneapolis, Minnesota, 2007.